IJPAS ONLINE

ISSN:1597-9261
FACULTY OF SCIENCE
BAYERO UNIVERSITY,
KANO - NIGERIA

Vol 2, No 3 (2008)

Table of Contents

Articles

8 pgs

Int. Jor. P. App. Scs., 2(3):76-83, 2008 *www.ijpas.com*

IJPAS ONLINE

ISSN: 1597-9261
FACULTY OF SCIENCE
BAYERO UNIVERSITY,
KANO - NIGERIA

INTERNATIONAL JOURNAL OF PURE AND APPLIED SCIENCES

FULL LENGTH RESEARCH PAPER

Retrospective Analysis and Design of Cusum Control Chart for Detection of Outbreaks of Infectious Diseases

R. A. Adeyemi

Department of Crop production, School of Agriculture and Agricultural Technology, Federal University of Technology, P.M.B. 65, MINNA, NIGER STATE, NIGERIA

Correspondence author email: adeyemira@yahoo.ca, Mobile: +234-08035861447

Received: July 2008 *Accepted: August 2008*

ABSTRACT

Cumulative sum (CUSUM) quality control charts are widely used in manufacturing industries to understand process performance, set up standard for the operational parameters, maintaining the standard and continual monitoring for improvement. Its application in non- manufacturing sectors (service industries, quality health care delivery etc), do not differ significantly. CUSUM charts which use all the available information has the advantage over Shewart chart in overcoming the problem of individual plot and can be used to capture a small shift from the reference value. In this paper, a retrospective analysis was first carried out, to provide diagnostic information about the process to be monitored. CUSUM scheme was then designed and implemented on hospital data obtained from Research and Record Unit of Ahmadu Bello Teaching Hospital (ABUTH), Kaduna, Nigeria. The applicability of this technique on infectious disease data further confirms a vital role quality control charts can play in medical process as well.

KEYWORDS: *CUSUM, Average run length , average quality level, average rejection level, infectious Disease*

INTRODUCTION

In process monitoring, quality practitioners frequently use Shewhart control chart (e.g. X, R, P, C-chart e.t.c.) so named after the pioneering work of Shewhart (1931). It can be shown that if there are sharp, intermittent changes in a process, these types of charts are highly effective in detecting them. However, if one is interested in small sustained shift in a process, other types of control chart may be preferred, for example Cumulative Sum (CUSUM) chart and for exponential increase in quality characteristics the Exponential Weighted Moving Average chart may be appropriate, Brower (2000).

The Cumulative sum (CUSUM) control chart attempts to device a simple procedure that will make use of all available information. It was introduced by Page (1954), who provided integral equations for approximating the average run length (ARL). Many authors (e.g. Ewan and Kemp.,1960; Johnson and Leone,,1983 ; Montgomery., 2000; and Lucas, 1982) have worked on the CUSUM and they proposed it as an alternative to Shewhart Control Charts. A major advantage of this chart over the ordinary Shewhart Control is that it is very effective in detecting relatively small shifts. They are more meaningful graphically, as process shifts are often easy to detect and points of change can be easily located.

Lucas (1985) gave a detailed procedure for designing a counted data CUSUM chart based on a Poison counts and implemented it on the occurrence of industrial accidents. Osanaiye and Talabi (1989) also considered a versatile dimension on the application of CUSUM chart in a non-manufacturing sector and implemented the approach on cases of diabetes patients. The monitoring of congenital malformations discussed by Chen (1978) and the detection of outbreaks of infectious disease by Farrington et.al.,(1996) are others of possible applications.

More so, in medical context rather, CUSUM charts have been proposed to monitor procedures in clinical chemistry, Nix A.B et.al (1986) and to monitor rare congenital malformations, Gallus G. et.al (1986). Application of the CUSUM to monitoring surgical performance was first proposed by Williams et. al., (1992).

The principal feature of the CUSUM chart is that successive observations (count/ week), say x_i values, of a variable are compared with predetermined target or reference value, *k* . The cumulative sum of deviations from *k*, i.e.

$$S_n = \sum_{i=1}^{n}(X_i - k)$$

is plotted on a chart , where k is taken to be target value and S_0 is the starting value for the CUSUM sequence often set at zero.

The CUSUM charts are often used to detect an upward or a downward shift in process quality (one – sided CUSUM chart) or shift is in both directions (two-sided CUSUM chart). To monitor a positive shift from the goal value, the CUSUM Statistic

$$S_H = max(0, X_i - k + S_{i-1}) \qquad (2)$$

is used to detect positive shift and

$$S_L = min(0, X_i - k + S_{i-1}) \qquad (3)$$

is used to detect negative shift . The process is taken to be out of control if $S_H \geq h$ for an upward shift or $S_L \leq -h$ for a downward shift. In this work the starting value S_0 is taken to be zero and h is the decision limit, the procedure to obtain h value is given in the next section.

The main objectives of this paper are to carry out a retrospective analysis on the outbreak of some diseases and to design CUSUM charts for early detection of the outbreaks of epidemics. To achieve this objective, this paper has five sections, first gives the introduction, section 2 describes the methodology, section 3 presents the results while section 4 and 5 give a summary of the result and conclusion respectively.

RESTROSPECTIVE ANALYSIS AND DESIGN METHODS

DESIGN METHODS

For a counted data CUSUM, the parameters to be determined are the goal or target value for the process, denoted by *k*, and the decision limit value, denoted by *h*. The value of k can be described as the goal value for the process, which is usually chosen between the acceptable process mean value (μ_a) and the mean level that the CUSUM scheme is intended to detect quickly; (μ_d) otherwise known as the rejectable mean value. The values μ_a and μ_d are mean numbers of counts per sampling interval.

The reference value *k* for the counted CUSUM should be chosen close to

$$k = \frac{\mu_d - \mu_a}{\ln \mu_d - \ln \mu_a} \qquad (4)$$

The procedures for determination of the CUSUM *k* and *h* are discussed extensively in the literature (Page.,1954; Lucas., 1985; Osanaiye and Talabi.,1989 and British Standards Institution (1982) .The reference value *k* is the same as the reference value for on SPRT testing the null hypothesis that mean is μ_a and the alternative hypothesis that is the mean μ_d, when $k \geq 1$, the k value will usually be rounded to the nearest integer. Thus, the countable CUSUM computations require only integer arithmetic.

SPAN TEST OF SIGNIFICANCE

The "span" test is used in this work for the test of significance of segments and the basis for the segmentation is the occurrençe of local maxima and minima which can be observed from the CUSUM chart.

Let T be the target or goal value from which departure in average value requires to be detected by the CUSUM scheme; C be the Cumulative sum of deviation from target value; V_{max} be maximum vertical distance of a CUSUM path from a chord joining the starting and ending points of a segment, σ be standard deviation of the sampled process; and r be the sample number corresponding to a suspected change point in the sequence i+1 to j, such that i, r, j are successive "corners" on the CUSUM chart.

The value of V_{max}/σ is referred to the monograph for appropriate span test. The p – scale on the monograph measures the probability under the null hypothesis of V_{max}/σ against span length m, where

$$V_{max} = C_r - C_i - \frac{(r-i)}{j-i}(C_j - C_i) \quad (5)$$

$$\text{and } m = j - 1 \text{ is the span length.}$$

It should be noted that in carrying out the retrospective analysis of long series of observations, each of the series is subjected to a type I error risk which is the probability of concluding that a difference exists between adjacent segments when no real change has occurred. It is therefore suggested that on testing a difference between two segments of total length m in a series of n observations at required significant level , individual tests are modified to $\frac{m}{n}(\alpha)$. This minimizes excessive type I error.

ESTIMATION OF LOCAL MEANS

The average departure from target value(T) over a segment is the average contribution of the sample values constituting the segments to the CUSUM deviation and the local mean over (i + 1, j) is therefore is calculated as;

$$\bar{x}_{i+1,j} = T + \frac{C_j - C_i}{j-i} \quad (6)$$

where T is the Target value and C_i and C_j are CUSUM values for starting and ending points of a segment.

DATA APPLICATION AND RESULTS

The data used for illustration was obtained from Research and Record Unit of Ahmadu Bello Teaching Hospital (ABUTH), Kaduna, Nigeria. It consists of 285 weekly sequences of reported cases of five infectious diseases (Malaria, Measles and pneumonia) between 1st January 1991 and 11th June 1996. These are extracted from a form containing 34 infectious diseases (including Food Poison and HIV/AID) tabulated by age and sex of which these diseases are epidemiological importance (Adeyemi., 2006).

Figure 1and figure2 below represents the CUSUM charts for Malaria and Tuberculosis respectively. The corresponding tabulated CUSUM values are used for arithmetic convenience and computations.

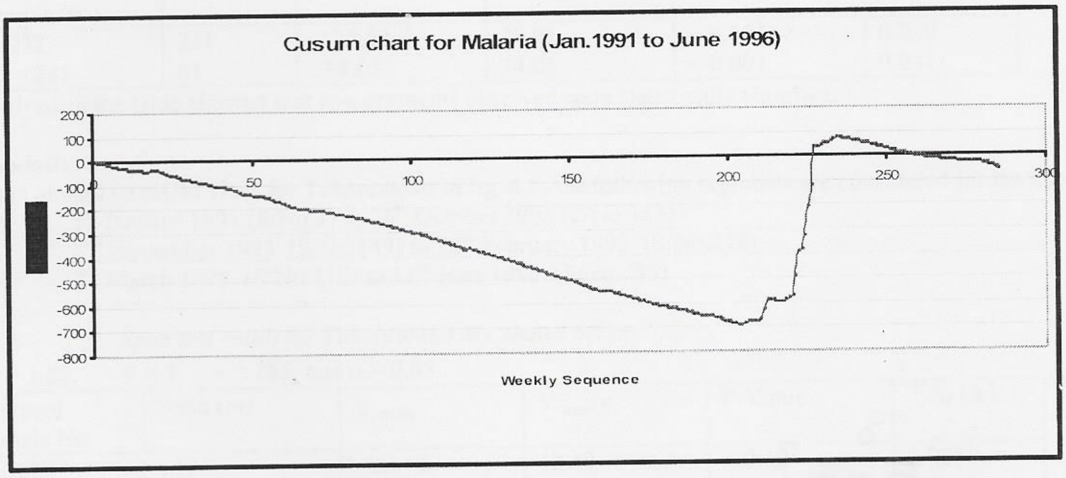

Figure 1: CUSUM chart for Malaria

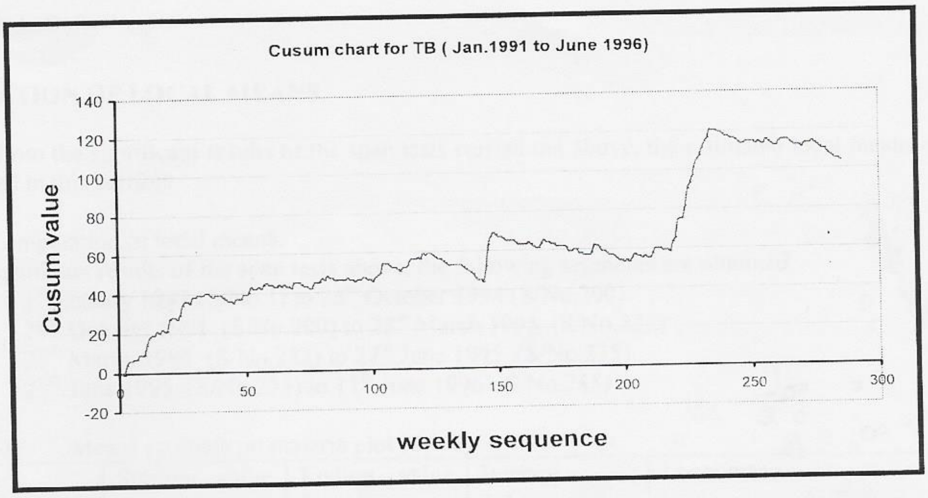

Figure 2: **CUSUM chart for Tuberculosis**

RETROSPECTIVE ANALYSIS

To carry out retrospective analysis on each disease, we use the CUSUM plots above for Malaria and Tuberculosis above.

Malaria

From the CUSUM chart for malaria in fig.1, the following segments are observed for carrying out span test

(i) 1st January 1991 (S/ No.1) to 28th March 1995 (S/No.222)

(ii) 28th January 1995 (S/No.222) 11th June 1996 (S/No.285)

Table1: The span test results are shown below for CUSUM segments for malaria

$\sigma = 16.09$ $T = 6$, $n = 285$, and $\alpha = 0.05$(level of significance)

Segments (Sample No)	Span (m)	V_{max}	V_{max}/σ	P-value	$m/n\ (\alpha)$
1 – 222	221	-322.02	20.01	< 0.001	0.039
222 – 285	63	398.05	24.05	< 0.001	0.011

The analysis in the table showed that two segments observed were statistically significant.

Tuberculosis

From the plotted CUSUM chart for Tuberculosis in fig. 4.5, the following segments are considered for the Span test

(i) 1st January 1991 (S/No.1) to 26th October 1993 (S/No.148)

(ii) 2nd November 1993 (S/No.149) to 28th February 1995 (S/No.218)

(iii) 7th March 1995 (S/No.219) to 11th June 1996 (S/No.285)

Table 2: *Span test result for Tuberculosis are shown below*

$\sigma = 1.86$, $T = 3$, $n = 285$, and $\alpha = 0.05$

Segment Sample No.	Span (m)	V_{max}	V_{max}/σ	P-Value	$m/n\ (\alpha)$
1 – 148	147	- 23.12	12.50	< 0.10	0.0260
149 – 218	69	- 3.26	1.76	> 0.50	0.0156
218 – 285	66	45.62	24.66	< 0.001	0.012

NB: It requires computing a single trend value (local mean) for segments as 1 – 218.

Analysis in Table 2 above shows that it was statistically significant for only one segment, which include samples between 218 to 285.

COMPUTATION OF LOCAL MEANS

Following from the significant results of the span tests carried out above, the estimated local means of the segments are presented in this section.

Malaria: Computation of local means.
From the significant results of the span tests above, the following segments are obtained

(i) 1st January 1991 (S/No.1) to 25th October 1994 (S/No.200)
(ii) 25th October 1991 (S/No.200) to 28th March 1995 (S/No.222)
(iii) 28th March 1995 (S/No.222) to 27th June 1995 (S/No.235)
(iv) 27th June 1995 (S/No.235) to 11th June 1996 (S/No.285)

Table 3: Means estimate on malaria plot (T = 6)

Segment	Starting value C_i	Ending value C_i	Interval J -I	Local mean $\bar{x}_{i-j} = T + \frac{c_j - c_i}{j - i}$
1 – 200	-5	-681	199	3
200 – 222	-681	- 372	22	21
222 – 235	- 372	73	13	41
235 – 285	73	-58	50	4

Table 3 confirms that CUSUM chart has captured the changes in the process mean level. Examining the last column, the variation in means level can be observed in column (5).

Tuberculosis: Computation of local means.
The table below gives estimates of local means of segments obtained from CUSUM chart for tuberculosis

(i) 1st January 1991 (S/No.1) to 28th February 1995 (S/No.104)
(ii) 28th February 1995 (S/No.104) to 18th July 1995 (S/No.238)
(ii) 18th July 1995 (S/No.238) to 11th June 1996 (S/No.285)

Table 4: Means estimate on Tuberculosis plot (T = 2).

Segment (Sample No.)	Starting value C_i	Ending value C_j	Interval J –I	Local mean $\bar{x}_{i-j} = T + \frac{c_j - c_i}{j - i}$
1 – 218	10	58	217	3
218 – 238	58	119	20	6
238 – 285	119	104	47	2

Table 10 presents the estimates of local means for the TB segments, which revealed differences in values of local means taking values 2,3 and 6 counts per week.

DESIGN CUSUM CONTROL CHART FOR DETECTION OF OUTBREAKS

Malaria (Case)

The mean level of occurrence of malaria when the process is regarded as in control state is known as the acceptable quality level (AQL), which is chosen nearer to the current mean level (3.34). It is denoted by $\mu_a = 4$ (approximately) with standard deviation $\sigma = 3.09$.

Suppose the administration or authority demands a shift of 1.5σ from μ_a to be the rejectable level, then a shift of 1.5σ in the positive direction yields $\mu_d = 9$ (approximately).
The values μ_a , μ_d , L_a and L_d are subject to continual review from time to time in accordance with the administration's desire or predefined goal.

Therefore, for $\mu_a = 4$ and $\mu_d = 9$, by applying equation (6)

$$k = \frac{\mu_d - \mu_a}{\ln\mu_d - \ln\mu_a}$$

i.e. k = 7 (approximately).

The corresponding decision limit, h is obtained from table to be 7.
Thus, hypothesis is stated as:
Ho: $\mu_a = 4$ versus H_1: $\mu_d = 9$,

For k = 7 and h = 7, we obtained the ARL at AQL, L_a = ARL (μ_a) = 95.5 and the corresponding ARL at RQL, ARL(μ_d) =3.14 (Lucas ;1985).

This implies that an out of control signal will be indicated whenever $S_i = \max(0, X_i - 7 + S_{i-1}) \geq 7$ for a case of Malaria. The scheme (k, h) = (7, 7) implemented on 4th quarters 1994 through 1st of quarters of 1995 yields the results in the table below

Count/week X_i	$(X_i - k)$	$\sum(X_i - k)$	$S_H \geq 7$
1	-6	0	0
0	-7	-7	0
2	-5	-12	0
1	-6	-18	0
1	-6	-24	0
1	-6	-30	0
2	-5	-35	0
18	11	-24	0
17	10	-14	0
5	-2	-16	0
4	-3	-19	0
4	-3	-22	0
15	8	-14	0
47	40	26	26*
43	36	52	52*
6	-1	51	51*

*This shows that an out of control signal occurred on 10th of January 1995.

Tuberculosis (Case)

Hypothesis is stated as:
Ho: $\mu_a = 4$ versus H_1 : $\mu_d = 6$, with a standard deviation of $\sigma = 1.85$, a 1.5σ shift in positive direction of : μ_a yields $\mu_d = 6$ (approx.) and obtain k = 5 and h = 5, with the ARL at AQL = 41.1 and ARL at RQL = 3.09. This implies that an out of control signal will be issued whenever, $S_i = \mathrm{Max}(0, X_i - 5 + S_{i-1}) \geq 5$ for TB case.

The scheme (k, h) = (5, 5) implemented on second quarter of 1995 for TB yields the result in the table below

Count/week X_t	$X_t - k$	$\Sigma(X_t - k)$	$S_H \geq 5$
2	-3	-3	0
2	-3	-6	0
1	-4	-10	0
7	2	-8	0
4	-1	-9	0
11	6	-3	0
3	-2	-5	0
2	--3	-8	0
8	3	-5	0
11	6	1	1
5	0	1	0
6	1	2	0
8	3	5	5*

* This shows that an out of control signal occurred on 9th of May 1995.

SUMMARY OF RESULTS

From the retrospective analysis, it was found that the CUSUM charts plotted for these diseases are capable of detecting small shift from mean level. More so, the points of changes in the process were clearly identified and the points at which these changes occurred easily located on the CUSUM plots in figures 1 and 2. Thus, CUSUM plot provides a visual aids diagnosis and identifies situation required urgent attention.

The span tests provided the statistical basis for recognizing significant changing point in the mean level and the point which the change occurred. In this work, it is meant to test the significance in the process changes between the adjacent segments along the CUSUM path, as evidenced in Tables 1 and Table2 for malaria and Tuberculosis respectively. Thus, the retrospective analysis gave a good understanding of the process being monitored to enable the SPC users plan for future.

The magnitude of the mean shift in the process is measured by the estimate of the local means as observed on the CUSUM paths are shown in Tables 3 and Table 4 for malaria and Tuberculosis respectively. This confirms ability of CUSUM charts to detect moderate mean shift from process mean level.

In section 4, CUSUM charts were designed for Malaria and Tuberculosis. A model each was generated for data series of Malaria and Tuberculosis. A CUSUM scheme run on a portion of the Malaria data indicated out of control (i.e. outbreak) on 10th of January 1995 while the other scheme run on Tuberculosis indicated out of control on 9th of May 1995.

CONCLUSION

By applying Cumulative sums (CUSUM), it has shown that quality control techniques commonly used in manufacturing sector can also be used in the medical sector for detection of increase in incidence rate of reported cases of disease outbreaks. Thus, quality control techniques are powerful tools, which can be used to obtain first hand information about an aberration in events of disease attack, disability or death. CUSUM schemes described in this work are capable of detecting relatively small shift in process not only in manufacturing process, but in medical sector as well. The illustration in this paper further confirmed the vital role Quality Control charts are capable of playing in a non-manufacturing sector as well, in particular for detection of such outbreaks in hospital data.

ACKNOWLEDGEMENT

We are highly grateful to Prof. P.A. Osanaiye, Department of Statistics, University of Ilorin, Nigeria for his contribution and comment. Also, thanks go to the Staff of Research and Record Unit, Ahmadu Bello Teaching Hospital, Kaduna Nigeria for their permission to use the data and their cooperation.

REFERENCES

Adeyemi R.A (2006) Retrospective analysis and design of CUSUM and EWMA control Charts for detection of outbreaks of disease epidemics", *Unpublished M. Sc. Thesis,* Department of Statistics, University of Ilorin, Nigeria.

Bower K.M.(2000) Using Exponential Weighted Moving Average (EWMA) charts, *Asia Pacific Process Engineer.*

British Standards Institution (1982), British Standards Guide to Data Analysis and Quality Control Using CUSUM Techniques (BS 5703, Part 4), London Author.

Chen, R. (1978) "A Surveillance System for Congenital Malformations", *Journal of the American Statistical Association.* 73, 323-327 .

Ewan, W. O. and Kemp K. W (1960) "Sampling Inspection of Continuous Processes with No Autocorrelation Between successive Results" *Biometrics.* 47, 363-380.

Farrington CP., Andrew N.J., Beale A.D., and M.A. Catchpole.(1996) A statistical algorithm for the early detection of outbreaks of infectious disease . *Journal of Royal Statistical Society, Series A:159: 547 – 63.*

Gallus G., Mandelli C., Marchi M., Radaelli G. (1986) "On Surveillance Methods for Congenital Malformations." *Statistic in Medicine.* 5, 565- 571.

Johnson , N. L and Leone , F.C (1983). "Cumulative Sum Control Charts- Mathematical Principles Applied to their Construction and use- PART III " *Industrial Quality Control,* 19.2, 22-28.

Lucas J.M .and R.B. Crosier (1982)Fast initial response for CUSUM quality control schemes. You're your CUSUM a head start. *Technometrics;*24:199- 205.

Lucas J.M. (1985). Counted data CUSUM's. *Technometrics;* 27:129-44.

Montgomery DC (2000) *Introduction to Statistic Quality Control:* Fourth Edition. New York: John Wiley and Sons.

Nix A.B., Rowlands R.J., Kemp K.W. (1986), Internal Quality Control in Clinical Chemistry: A Teaching Review," *Statistics in Medicine*, 6, 425-440.

Osanaiye P.A. and O. Talabi (1989) On some non-manufacturing applications of counted data cumulative sum (CUSUM) control schemes, *The Statistician.*

Page, E.S. (1954) "Continuous Inspection Schemes, "*Biometrika*, **41**, 100-114.

Shewhart W.A (1931) Economic control of the quality of manufactured product, Van Nostrand New York

Williams S.M., B.R. Parry, and M.M.T. Schlup (1992). Quality control: An application of the CUSUM. *British Medical Journal* 304:1359-61.