

LINEAR RELATIONSHIPS TO ASSIST IN PARAMETER SELECTION OF A CUSUM CONTROL CHART

**Michael Anthony Arnold
 Cox¹⁾**

¹⁾ Newcastle University
 United Kingdom
 email: mike.cox@ncl.ac.uk

Abstract: The cumulative sum (CUSUM) chart is widely employed in quality control. These charts are designed to exhibit acceptable average run lengths (ARL) both when the process is in and out of control. For such a chart, for specified ARLs, this paper introduces a technique employing a simple linear approximation for parameter selection. Graphs are produced to provide the coefficients for the fit. The simplicity of this approach should encourage its adoption as an additional tool in the statistical process control kit.

Keywords: ARL; average run length; cumulative sum chart; CUSUM; SPC; statistical process control;

1. INTRODUCTION

An effective way of monitoring a quantitative measure of a continuously operating manufacturing system is to employ a CUSUM control chart. Given an acceptable process mean m_a , the aim is to detect rapidly significant deviations to a barely tolerable mean m_r , any worse being rejected, where $m_r > m_a$. The suffix “a” denotes acceptance, while “r” denotes rejection, this notation follows that of Kemp (1962). This paper has the following structure. Having outlined the problem the bounds on the ARLs associated with the acceptance/rejection cases are found. A special case is employed as a motivating example, linear fits being derived for the parameters. These are extended by employing graphical estimation for the linear parameters of the fits. A detailed comparison is made to alternate methods that are more cumbersome. Finally additional examples are presented and possible extensions considered.

2. BACKGROUND DISCUSSION

This problem has been previously outlined (Cox, 2003/4) which contains a general introduction and motivation for the problem. Those interested in details of the numerical integration employing Gaussian knot points see Cox (2010). The notation previously used is employed, only key terms being redefined below. The CUSUM at time t is written following Woodall (1986) the two-sided chart may be expressed as two one-sided charts. For standardised variables,

$$U_t = \max(0, U_{t-1} + z_t - k),$$

$$L_t = \min(0, L_{t-1} + z_t + k).$$

An out of control signal is produced if $U_t \geq h$, or if $L_t \leq -h$. Where k is referred to as the reference value and h as the threshold. This paper employs the one-

sided CUSUM chart approach. The average number of samples tested before a change is detected, or before a false alarm is sent if the process has not changed, is termed the ARL. A high ARL (L_a) is desired when the process mean is at the acceptable level m_a , and a low ARL (L_r) when the process mean moves off to a rejectable level m_r . The intention of the chart is to detect a shift of D , where $D = m_r - m_a$. L_a is usually fixed at several hundreds, while L_r is generally between 3 and 10.

Bounds On L_r Given L_a

The original nomogram (Kemp 1962) or an alternative approach (Cox 2003/4) may be employed to provide acceptable ranges for L_a and L_r . In other words for given L_a what range of values could L_r achieve? The procedure to find estimates of k and h for given L_r and L_a is fully described in the previous paper. It is the polynomials from Tables 3, 4, 5 and 6 (Cox 2003/4) that are employed here. An exhaustive approach is adopted, given L_r and L_a if h and k can be found then the ARLs are acceptable. The resulting acceptable range for L_r given L_a is presented in Figure 1.

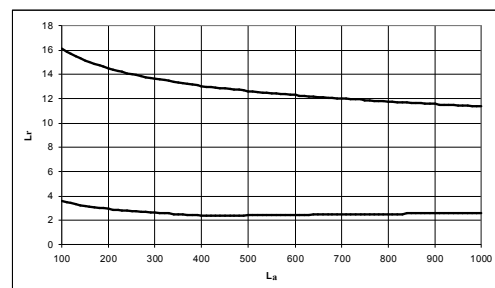


Figure 1-Bounds On L_r Given L_a

So, for example if L_a is 500, then L_r must lie in the range [2.42, 12.63] more sensibly rounded to [3, 12]. For practical purposes the smooth curves will usually be rounded to integer values contained between the two curves. This is for convenience, run lengths are integers but their average need not be. It is only the L_r, L_a pairs within the bounds that are acceptable.

Within the constraints provided by the bounds in Figure 1 it is desired to obtain parameter estimates consistent with given ARLs. To investigate the relationship between k and the ARL a series of solutions were found for $2 \leq h \leq 5$ in 101 steps, where the range selected duplicates that of the popular nomogram presented by Kemp (1962). L_a takes values from 100 to 1000, integer values of L_r consistent with the above plot, are examined.

The numerical procedure is extremely time consuming. Given required values of L_a and L_r an estimate of k ($k_a = k$ and $k_r = -k$) corresponding to the initial value of h ($h = 2$) is provided. The non-linear equation corresponding to the ARLs are solved successively using a modification of the Powell (1970) hybrid method to obtain estimates of k_a and k_r . It should be recalled that each functional evaluation involves inverting a 150x150 matrix! The current estimates of k_a and k_r are fed forward as starting estimates corresponding to the next value of h . The final estimates provided the values for the plots and tables presented below. As an illustration a special case is now considered.

Parameter Estimation For $L_a=100$ and $L_r=10$

Consider $L_a = 100$ and $L_r = 10$, these values may be achieved with a number of different parameter estimates (k_a, k_r, h). Certain linear relationships become very clear in the following graphs (Figures 2, 3 and 4) that summarise the solutions.

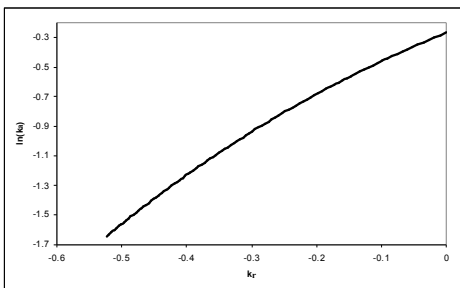


Figure 2- For $L_a=100$ And $L_r=10$ The Relationship Between $\ln(k_a)$ And k_r For h In [2,5]

It is important to assess the linearity of this plot, linear regression is employed and R^2 reported. R^2 is called the coefficient of determination; a value of 100% indicates a perfect linear fit. For Figure 2 the results are $\ln(k_a) = -0.176 + 2.66 k_r$ and $R^2 = 99.2\%$.

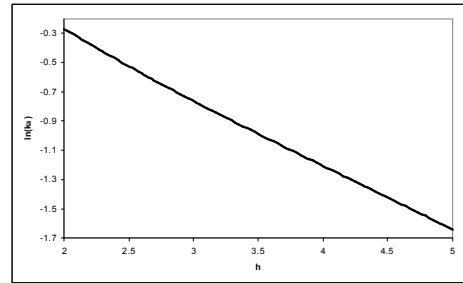


Figure 3 - For $L_a=100$ And $L_r=10$ The Relationship Between $\ln(k_a)$ And h

For Figure 3 the results are $\ln(k_a) = 0.602 - 0.451 h$ and $R^2 = 99.9\%$.

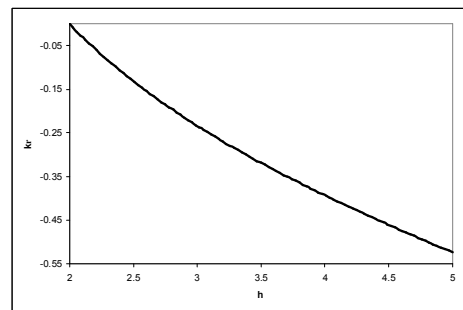


Figure 4 - For $L_a=100$ And $L_r=10$ The Relationship Between k_r And h

For Figure 4 the results are $k_r = 0.286 - 0.167 h$ and $R^2 = 98.6\%$. Thus in these cases it is reasonable to assume a linear relationship between $(k_r, \ln(k_a))$, $(h, \ln(k_a))$ and (h, k_r) . But is this linearity extendable?

It is of interest to see how these parameters are affected by changing L_a and L_r . For $L_a = 100$ and $L_r = 4$ the corresponding results are $\ln(k_a) = 0.625 + 1.50 k_r, R^2 = 100.0\%$, $\ln(k_a) = 0.602 - 0.451 h, R^2 = 99.9\%$, $k_r = -0.0156 - 0.300 h, R^2 = 99.9\%$.

While for $L_a = 1000$ and $L_r = 10$ the corresponding results are $\ln(k_a) = 0.301 + 1.84 k_r, R^2 = 99.9\%$, $\ln(k_a) = 0.832 - 0.310 h, R^2 = 99.3\%$, $k_r = 0.286 - 0.167 h, R^2 = 98.6\%$.

While the linearity is preserved the relationship between the parameters (gradient and intercept) exhibits no obvious consistency on adjusting the ARLs. Extending the analysis to additional ARLs, more detailed results may be produced.

Deriving The Linear Fit Parameters

To investigate the general problem, following the results obtained above, tables were constructed. An exhaustive analysis is not essential. Sufficient data is only required to derive descriptive coefficients. Points were chosen to provide paths of equal values of L_a/L_r through the tables, the results are presented in Tables 1,...,6.

Table 1 - The Constant In The Linear Fit $\ln(ka)=\text{constant} + \text{gradient } h$

| | L_a | | | | | | | | |
|---------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 100 | 175 | 250 | 375 | 500 | 625 | 750 | 875 | 1000 |
| 3 | | | 0.658 | | 0.744 | | 0.796 | | 0.832 |
| 4 | 0.602 | | | | | | | | |
| 5 | | | 0.658 | | 0.744 | | 0.796 | | 0.832 |
| 7 | | | | | | | 0.796 | 0.815 | 0.832 |
| 8 | | | | 0.707 | 0.744 | 0.772 | | | |
| L_r 9 | | 0.621 | 0.658 | | | | | | |
| 10 | 0.602 | | | | 0.744 | | | | 0.832 |
| 11 | | | | | | | 0.796 | | 0.832 |
| 12 | | | | | 0.744 | | | | |
| 14 | | | 0.658 | | | | | | |
| 16 | 0.602 | | | | | | | | |
| Average | 0.602 | 0.621 | 0.658 | 0.707 | 0.744 | 0.772 | 0.796 | 0.815 | 0.832 |

Table 2- The Gradient In The Linear Fit $\ln(ka)=\text{constant} + \text{gradient } h$

| | L_a | | | | | | | | |
|---------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| | 100 | 175 | 250 | 375 | 500 | 625 | 750 | 875 | 1000 |
| 3 | | | -0.359 | | -0.328 | | -0.317 | | -0.310 |
| 4 | -0.451 | | | | | | | | |
| 5 | | | -0.359 | | -0.328 | | -0.317 | | -0.310 |
| 7 | | | | | | | -0.317 | -0.313 | -0.310 |
| 8 | | | | -0.339 | -0.328 | -0.321 | | | |
| L_r 9 | | -0.384 | -0.359 | | | | | | |
| 10 | -0.451 | | | | -0.328 | | | | -0.310 |
| 11 | | | | | | | -0.317 | | -0.310 |
| 12 | | | | | -0.328 | | | | |
| 14 | | | -0.359 | | | | | | |
| 16 | -0.451 | | | | | | | | |
| Average | -0.451 | -0.384 | -0.359 | -0.339 | -0.328 | -0.321 | -0.317 | -0.313 | -0.310 |

Table 3- R2 For The Linear Fit $\ln(ka)=\text{constant} + \text{gradient } h$

| | L_a | | | | | | | | |
|---------|-------|------|------|------|------|------|------|------|------|
| | 100 | 175 | 250 | 375 | 500 | 625 | 750 | 875 | 1000 |
| 3 | | | 99.6 | | 99.5 | | 99.4 | | 99.3 |
| 4 | 99.9 | | | | | | | | |
| 5 | | | 99.6 | | 99.5 | | 99.4 | | 99.3 |
| 7 | | | | | | | 99.4 | 99.4 | 99.3 |
| 8 | | | | 99.5 | 99.5 | 99.4 | | | |
| L_r 9 | | 99.7 | 99.6 | | | | | | |
| 10 | 99.9 | | | | 99.5 | | | | 99.3 |
| 11 | | | | | | | 99.4 | | 99.3 |
| 12 | | | | | 99.5 | | | | |
| 14 | | | 99.6 | | | | | | |
| 16 | 99.9 | | | | | | | | |

Table 4- The Constant In The Linear Fit $kr = \text{constant} + \text{gradient } h$

| | L_a | | | | | | | | | Average |
|---------|--------|-------|--------|-------|--------|-------|--------|-------|--------|---------|
| | 100 | 175 | 250 | 375 | 500 | 625 | 750 | 875 | 1000 | |
| 3 | | | -0.111 | | -0.111 | | -0.111 | | -0.111 | -0.111 |
| 4 | -0.016 | | | | | | | | | -0.016 |
| 5 | | | 0.058 | | 0.058 | | 0.058 | | 0.058 | 0.058 |
| 7 | | | | | | | 0.167 | 0.167 | 0.167 | 0.167 |
| 8 | | | | 0.211 | 0.211 | 0.211 | | | | 0.211 |
| L_r 9 | | 0.250 | 0.250 | | | | | | | 0.250 |
| 10 | 0.286 | | | | 0.286 | | | | 0.286 | 0.286 |
| 11 | | | | | | | 0.318 | | 0.318 | 0.318 |
| 12 | | | | | 0.347 | | | | | 0.347 |
| 14 | | | 0.399 | | | | | | | 0.399 |
| 16 | 0.445 | | | | | | | | | 0.445 |

In summary, the coefficients of determination (Tables 3 and 6) support the assumption of a linear relationship. It is also clear that $k_a = k_a(L_a, h)$ (Tables 1 and 2) being independent of L_r , while $k_r = k_r(L_r, h)$ (Tables 4 and 5) being independent of L_a . This independence

will go some way to indicating why numerical procedures experience problems when solving the problem. Simple interpolation may be employed to derive estimates of the coefficients not included in the tables.

Table 5 - The Gradient In The Linear Fit $kr = \text{constant} + \text{gradient } h$

| | L_a | | | | | | | | | Average |
|---------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---------|
| | 100 | 175 | 250 | 375 | 500 | 625 | 750 | 875 | 1000 | |
| 3 | | | -0.397 | | -0.397 | | -0.397 | | -0.397 | -0.397 |
| 4 | -0.300 | | | | | | | | | -0.3 |
| 5 | | | -0.249 | | -0.249 | | -0.249 | | -0.249 | -0.249 |
| 7 | | | | | | | -0.198 | -0.198 | -0.198 | -0.198 |
| 8 | | | | -0.184 | -0.184 | -0.184 | | | | -0.184 |
| L_r 9 | | -0.174 | -0.174 | | | | | | | -0.174 |
| 10 | -0.167 | | | | -0.167 | | | | -0.167 | -0.167 |
| 11 | | | | | | | -0.162 | | -0.162 | -0.162 |
| 12 | | | | | -0.159 | | | | | -0.159 |
| 14 | | | -0.154 | | | | | | | -0.154 |
| 16 | -0.151 | | | | | | | | | -0.151 |

Table 6 - R^2 For The Linear Fit $kr = \text{constant} + \text{gradient } h$

| | L_a | | | | | | | | |
|---------|-------|------|-------|------|-------|------|-------|------|-------|
| | 100 | 175 | 250 | 375 | 500 | 625 | 750 | 875 | 1000 |
| 3 | | | 100.0 | | 100.0 | | 100.0 | | 100.0 |
| 4 | 99.9 | | | | | | | | |
| 5 | | | 99.8 | | 99.8 | | 99.8 | | 99.8 |
| 7 | | | | | | | 99.4 | 99.4 | 99.4 |
| 8 | | | | 99.1 | 99.1 | 99.1 | | | |
| L_r 9 | | 98.9 | 98.9 | | | | | | |
| 10 | 98.6 | | | | 98.6 | | | | 98.6 |
| 11 | | | | | | | 98.4 | | 98.4 |
| 12 | | | | | 98.2 | | | | |
| 14 | | | 97.8 | | | | | | |
| 16 | 97.5 | | | | | | | | |

It is necessary to generate linear coefficients for all acceptable ARLs. This is done employing the graphs (Figures 5,...,9) below, which summarise the results in the previous tables. The figures provide a natural aid for

interpolation.

Graphical Estimation Of The Linear Parameters

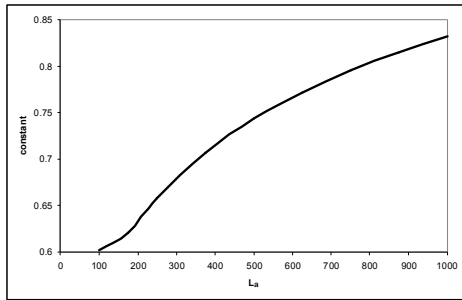


Figure 5 - The Relationship Between L_a And The Constant In The Linear Fit

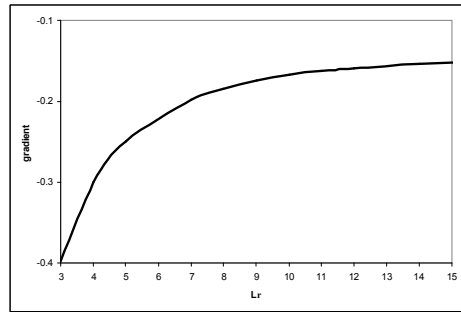
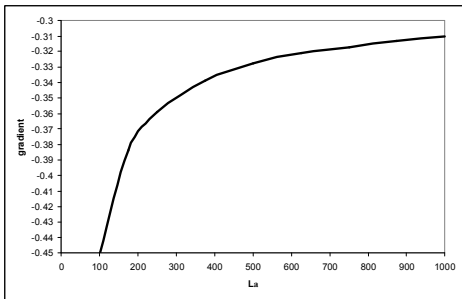


Figure 9 -The Relationship Between L_r And The Gradient In The Linear Fit



To aid the estimation process, a *log/log* plot is produced of Figure 5.

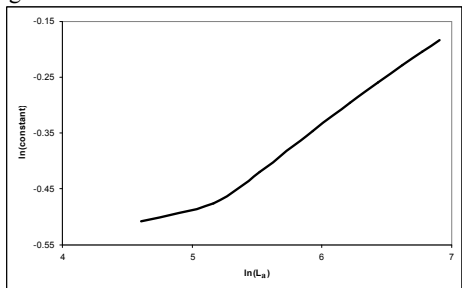


Figure 7 - A Logarithmic Plot Of The Relationship Between L_a And The Constant In The Linear Fit

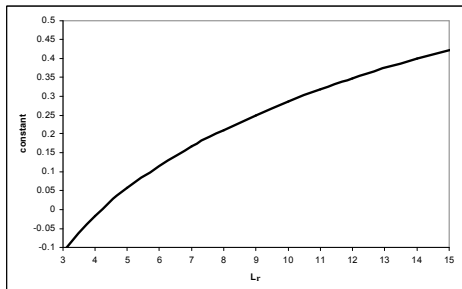


Figure 8 - The Relationship Between L_r And The Constant In The Linear Fit

A numerical example is employed to illustrate the procedure.

3. EXAMPLE

Recall the estimates obtained for $L_a = 100$ and $L_r = 10$.

$\ln(k_a) = 0.602 - 0.451 h$, $R^2 = 99.9\%$ (from Tables 1, 2 and 3 or Figures 5 and 6 or 7 for the parameters) $k_r = 0.286 - 0.167 h$, $R^2 = 98.6\%$ (from Tables 4, 5 and 6 or Figures 8 and 9 for the parameters) On eliminating h it is found that $\ln(k_a) = -0.170 + 2.70 k_r$ remarkably close to the line produced on deriving a direct fit $\ln(k_a) = -0.176 + 2.66 k_r$, (Figure 2) so the procedure is reliable. Given L_a and L_r a relationship can be derived between k_a and k_r using Tables 1, 2, 4 and 5 or Figures 5, 6, 8 and 9. This may be solved for the optimal situation ($k_a = k_r$) or to derive k_a and k_r for given h . In this case the optimal situation corresponds to $k = 0.338$, thus $h = 3.738$ where the problem ($\ln(k_a) = -0.170 + 2.70 k_r$, $k_a = k$ and $k_r = -k$) may be solved iteratively, for instance using the solve option within Excel (available from the author), or graphically. This should be compared with $k = 0.35$ and $h = 3.73$ from Cox (2003/4) and may be checked using available spreadsheet code (Cox 1999) giving $L_a = 101$ and $L_r = 10$, an acceptable agreement. A comparison with the previous procedure (Cox 2003/4) revealed at most a 2% variation in estimates of k and 3% for h . The current approach has a greater range of applicability as demonstrated by the following examples.

The following examples exhibit reasonable results beyond the scope of Cox (2003/4). For $L_a = 100$ and $L_r = 3$, the estimates obtained were $k = 0.818$ and $h = 1.781$, on checking the ARL values were $L_a = 85$ and $L_r = 3$. For $L_a = 375$ and $L_r = 14$, the estimates obtained were $k = 0.372$ and $h = 5.005$, on

checking the ARL values were $L_a = 335$ and $L_r = 13$, a reasonable performance.

4. EXTENSIONS

It is important to see if the approach has a wider range of application. Consider extending the range for h to $[2,10]$. The same ARLs adopted for Figures 2, 3 and 4, $L_a = 100$ $L_r = 10$, are presented here in Figures 10, 11 and 12.

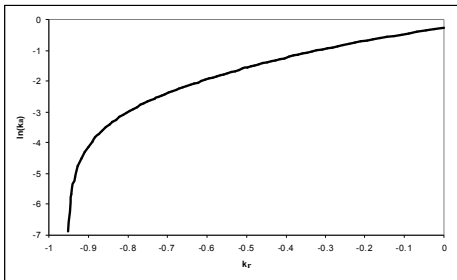


Figure 10 - For $L_a=100$ And $L_r=10$ The Relationship Between $\ln(ka)$ And kr For h In $[2,10]$

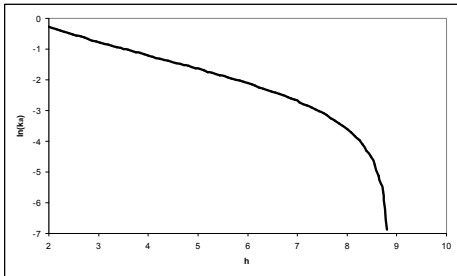


Figure 11 - For $L_a=100$ And $L_r=10$ The Relationship Between $\ln(ka)$ And h

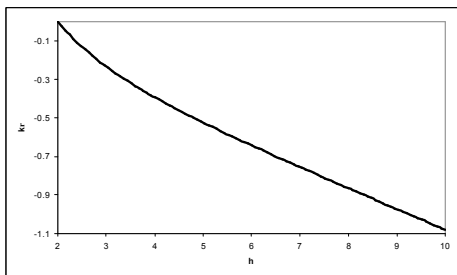


Figure 12 - For $L_a=100$ And $L_r=10$ The Relationship Between kr And h

Statistical analysis is not necessary; the non-linearity of Figures 10 and 11 is self-evident. Was the range adopted for the original nomograms (Kemp 1962) fortuitous?

A comparison to an alternate approach is now

presented.

5. COMPARISON TO OTHER APPROACHES

In Rogerson (2006) very complex formulas were developed to solve the equivalent problem. The exact formula for k involved the LambertW function (see Rogerson 2006), which gives technical details plus some practical references). A greatly simplified approximation is also given

$$k \approx \sqrt{\frac{\ln(-\ln(v)) - \ln(v)}{2L_r} - \frac{1}{2L_a}}$$

where

$$v = \frac{L_r}{L_a} e^{1-(L_r/L_a)}$$

then

$$h \approx \frac{2k^2 L_a + 2}{2k^2 L_a + 1} \times \frac{\ln(1 + 2k^2 L_a)}{2k} - 1.166$$

The following examples are the two employed by Rogerson (2006), who also reports the estimates of Rossi et al. (1999). For the first example if $L_a = 500$ and $L_r = 7$.

- $k = 0.5634$ $h = 3.96$ $L_a = 492$ Rogerson (2006)
- $k = -0.5634$ $h = 3.96$ $L_r = 8$ Rogerson (2006)
- $k = 0.6$ $h = 3.8$ $L_a = 521$ Rossi et al. (1999)
- $k = -0.6$ $h = 3.8$ $L_r = 7$ Rossi et al. (1999)
- $k = 0.60$ $h = 3.85$ $L_a = 554$ the methods proposed here
- $k = -0.60$ $h = 3.85$ $L_r = 7$ the methods proposed here

The calculated ARLs were found using published procedures (Cox 1999). Extending the precision of the results reported here gives $k = 0.6062$ and $h = 3.7310$ as the values, corresponding to the desired ARL values. It should be noted that estimated ARL values are highly dependent on the precise values of the parameters selected, for example $k = 0.59$ and $h = 3.72$ gives $L_a = 443$ while $k = 0.61$ and $h = 3.74$ gives $L_a = 518$. As a rule of thumb it is suggested that the selected parameters lie in the range $k \pm 0.015k$ and $h \pm 0.015h$ so $\pm 1.5\%$ in each case. The estimate of L_r is fairly robust to these changes, adopting these slight adjustments is reflected in the estimate of L_a .

- For the second example if $L_a = 500$ and $L_r = 3$.
- $k = 0.96$ $h = 2.39$ $L_a = 470$ Rogerson (2006)
- $k = -0.96$ $h = 2.39$ $L_r = 3$ Rogerson (2006)
- $k = 1.04$ $h = 2.26$ $L_a = 531$ Rossi et al. (1999)
- $k = -1.04$ $h = 2.26$ $L_r = 3$ Rossi et al. (1999)
- $k = 1.01$ $h = 2.25$ $L_a = 451$ the methods proposed here
- $k = -1.01$ $h = 2.25$ $L_r = 3$ the methods proposed here

The calculated ARLs were found using published procedures (Cox 1999). Extending the precision of the results reported here gives $k = 1.01$ and $h = 2.30$ as the values corresponding to the desired ARL values.

A simple scatter plot suggests, in the two cases

considered here, that the estimates of Rogerson (2006) are furthest from the best estimate. In all cases the rejection run lengths are consistent, while the acceptable values differ, but are broadly similar and close to the desired values.

6. ALTERNATE APPROACHES

This section is not intended as a complete review of available software, but describes a recent development. Knoth (2009) has provided a package "SPC" which is implemented in the R (R Development Core Team 2009) statistical language. R is a language and environment for statistical computing and graphics. The SPC package provides functions for the evaluation of control charts by means of the zero-state, steady state ARL. Setting up control charts for given in-control ARL and plotting of the related figures. The control charts under consideration are one- and two-sided Exponentially Weighted Moving Average (EWMA), CUSUM, and Shiryayev-Roberts schemes for monitoring the mean of normally distributed independent data. The ARL calculation of the same set of schemes under drift is available. The approach determines the ARL by numerically solving the related ARL integral equation

by means of collocation (Chebyshev polynomials, Knoth 2005 and 2006). This approach would require downloading, installing and learning these packages. A little daunting for the occasional user.

7. CONCLUSIONS

The procedures presented here give a pair of linear equations, which quickly provide estimates of the parameters for a CUSUM chart with desired ARLs. The approach is reasonably accurate and more elegant than the traditional approach employing a nomogram. The approach exhibits similar results to those previously reported (Cox, 2003/4) with the advantage of greater ease of use and range of applicability. It is demonstrated that the range of validity adopted by the nomogram is ideal, any extension to the key parameters, k_a , k_r and h , is not practical with the technique proposed here.

A natural extension would be to examine EWMA charts whose parameters become difficult to evaluate when the distributional parameters of quality characteristic are estimated and the ideal assumption of normality is violated. These charts may also be described by the general formalism introduced earlier (Cox 1999).

REFERENCES:

- [1] Cox M.A.A., "Towards The Implementation Of A Universal Control Chart and Estimation Of Its Average Run Length Using A Spreadsheet", *Quality Engineering*, 11, 511-536, 1999.
- [2] Cox M.A.A., "Beyond The Nomogram: Rapid Selection Of Parameters For A CUSUM Chart", *Quality Engineering*, 16, 1-12, 2003/4.
- [3] Cox M.A.A., "A Numerical Method For Modelling The Parameters λ and Δ Of An EWMA Chart", *International Journal for Quality Research*, 4, 171-180, 2010.
- [4] Kemp K.W., "The Use of Cumulative Sums for Sampling Inspection Schemes", *Applied Statistics*, 11, 16-31, 1962.
- [5] Knoth S., "Accurate ARL computation for EWMA-S² control charts", *Statistics and Computing*, 15, 341-352, 2005.
- [6] Knoth S., "Computation of the ARL for CUSUM-S² schemes", *Computational Statistics and Data Analysis*, 51, 499-512, 2006.
- [7] Knoth S. "SPC" <http://cran.r-project.org/web/packages/spc/> (software) and <http://cran.r-project.org/web/packages/spc/spc.pdf> (manual) 2009. Accessed May-2012.
- [8] Powell, M.J.D., "A Hybrid Method For Non-Linear Algebraic Equations" in *Numerical Methods for Non-linear Algebraic Equations*, ed. P. Rabinowitz, London, New York, Gordon and Breach Science Publishers, 1970.
- [9] R Development Core Team, "R: A Language and Environment for Statistical Computing", R Foundation for Statistical Computing, Vienna, Austria <http://www.r-project.org/> 2009. Accessed May-2012.
- [10] Rogerson, P.A., "Formulas for the Design of CUSUM Charts", *Communications in Statistics: Theory and Methods*, 35, 373-383, 2006.
- [11] Rossi, G., Lampugnani, L. and Marchi, M., "An approximate CUSUM procedure for the surveillance of health events", *Statist. Med.*, 18, 2111-2122, 1999.
- [12] Woodall W.H., "The Design of CUSUM Quality Control Charts", *Journal of Quality Technology*, 18, 99-102, 1986

Received: 13.01.2012

Accepted: 04.04.2012

Open for discussion: 1 Year