CUSUM tests based on grouped observations

Karl-Heinz Eger
Chemnitz University of Technology
Department of Mathematics
Chemnitz, Germany
eger@mathematik.tu-chemnitz.de

Evgeni Borisovich Tsyo
Novosibirsk State Technical University
Department of Applied Mathematics
Novosibirsk, Russia
ebcui@nstu.ru

Abstract—This paper deals with CUSUM tests based on grouped or classified observations. The computation of average run length is reduced to that of solving of a system of simultaneous linear equations. Moreover a corresponding approximation for the mean of a normal distribution based on F-optimal grouping schemes. The considered example demonstrates that hight efficient CUSUM tests can be obtained for F-optimal grouping schemes already with a small number of groups.

Index Terms—CUSUM, continuous inspection schemes, average run length, grouped observations, classified observations, sequential tests, sequential analysis.

I. INTRODUCTION

We consider a technological, chemical or environmental process which is observed by a sequence \( X_1, X_2, \ldots \) of independent random variables with a density \( f_{\theta}(x) \) with respect to some measure \( \mu \). Here \( \theta \) denotes a parameter with values in a parameter space \( \Theta \). The process is said to be under control if to a given parameter value \( \theta^* \in \Theta \) the true parameter value \( \theta \) satisfies \( \theta \leq \theta^* \).

We assume, that at a random time point \( T \) the parameter \( \theta \) changes from \( \theta \leq \theta^* \) to \( \theta > \theta^* \). Then we are interested in a sampling scheme that detects this parameter change as with a delay as small as possible (so-called change point problem).

Based on Wald’s sequential likelihood ratio test (SLRT) Page [7] developed in this context his famous cumulative sum test (CUSUM test). This test or procedure consists of sequences of SLRTs for

\[
H_0 : \theta = \theta_0 < \theta^* \quad \text{against} \quad H_1 : \theta = \theta_1 > \theta^*.
\]

The acceptance of \( H_0 \) by a test of this sequence is interpreted as a confirmation of \( \theta < \theta^* \), otherwise acceptance of \( H_1 \) is considered as an alert signal for changing of parameter value from a value \( \theta < \theta^* \) to a value \( \theta > \theta^* \). That means, sampling is continued as long as the tests accept hypothesis \( H_0 \) and is stopped if for the first time a test decides for \( H_1 \). For a survey on CUSUM procedures we refer to [5]. Certain aspects of optimality of CUSUM tests are considered by [8] and [6].

I. INTRODUCTION

We consider a technological, chemical or environmental process which is observed by a sequence \( X_1, X_2, \ldots \) of independent random variables with a density \( f_{\theta}(x) \) with respect to some measure \( \mu \). Here \( \theta \) denotes a parameter with values in a parameter space \( \Theta \). The process is said to be under control if to a given parameter value \( \theta^* \in \Theta \) the true parameter value \( \theta \) satisfies \( \theta \leq \theta^* \).

We assume, that at a random time point \( T \) the parameter \( \theta \) changes from \( \theta \leq \theta^* \) to \( \theta > \theta^* \). Then we are interested in a sampling scheme that detects this parameter change as with a delay as small as possible (so-called change point problem).

Based on Wald’s sequential likelihood ratio test (SLRT) Page [7] developed in this context his famous cumulative sum test (CUSUM test). This test or procedure consists of sequences of SLRTs for

\[
H_0 : \theta = \theta_0 < \theta^* \quad \text{against} \quad H_1 : \theta = \theta_1 > \theta^*.
\]

The acceptance of \( H_0 \) by a test of this sequence is interpreted as a confirmation of \( \theta < \theta^* \), otherwise acceptance of \( H_1 \) is considered as an alert signal for changing of parameter value from a value \( \theta < \theta^* \) to a value \( \theta > \theta^* \). That means, sampling is continued as long as the tests accept hypothesis \( H_0 \) and is stopped if for the first time a test decides for \( H_1 \). For a survey on CUSUM procedures we refer to [5]. Certain aspects of optimality of CUSUM tests are considered by [8] and [6].

Let \( l \) be the random number of observations until the first acceptance of \( H_1 \). Then the most important statistical characteristic of a CUSUM test is the average run length (ARL) \( E_{\theta}L \), \( \theta \in \Theta \). This function describes the average number of observations until an alert depending on parameter \( \theta \in \Theta \). An ‘ideal’ ARL should have the following property:

\[
E_{\theta}L = \begin{cases} 
\infty & \text{for } \theta \leq \theta^* \\
1 & \text{for } \theta > \theta^*
\end{cases}.
\]

It is evident that this property can be realized only approximately, e.g., as follows. Let \( l_0 \) and \( l_1 \), \( l_0 > l_1 \), given bounds for the ARL for \( \theta = \theta_0 \) and \( \theta = \theta_1 \), \( \theta_0 < \theta^* < \theta_1 \). Then a side condition for a CUSUM test could be

\[
E_{\theta}L \geq l_0 \quad \text{for } \theta = \theta_0 \quad \text{and} \quad E_{\theta}L \leq l_1 \quad \text{for } \theta = \theta_1.
\]

Numerical studies show that even such a side condition is yet to strong. An other aspect could be: \( E_{\theta}L = l^* \) for \( \theta = \theta^* \) and minimisation of \( E_{\theta}L \) for \( \theta = \theta_1 \). As a rule one has to consider the whole behaviour of ARL for \( \theta \in \Theta \) to decide whether a CUSUM test is useful in a particular situation. Hence very important are methods for computation of ARL of CUSUM test.

In this paper we will present a method for design of CUSUM tests if the random variables \( X_1, X_2, \ldots \) can be observed only in a restricted manner as so-called grouped or classified observation variables. The computation of ARL is reduced to that of solving of a system of simultaneous linear equations. As example we consider CUSUM tests for mean of normal distribution with known variance based on \( F \)-optimal grouping schemes. The influence of grouping to ARL is discussed. The examples show that efficient CUSUM tests can be obtained already with small numbers of groups.

II. SEQUENTIAL TESTS BASED ON GROUPED OBSERVATIONS

CUSUM tests are sequences of sequential likelihood ratio tests (SLRTs). Preliminary some remarks to SLRTs based on grouped observations.

Let \( X_1, X_2, \ldots \) be a sequence of independent random variables with density function \( f_{\theta}(x) \), \( \theta \in \Theta \). Let \( \mathcal{X} \) be the codomain of \( X_1 \) and let \( \mathcal{X}_1, \ldots, \mathcal{X}_m \), \( m \geq 2 \), be a partition of \( \mathcal{X} \) such that \( \bigcup_{i=1}^{m} \mathcal{X}_i = \mathcal{X} \) and \( \mathcal{X}_i \cap \mathcal{X}_j = \emptyset \), \( i \neq j \), \( i,j = 1, \ldots, m \). Such a partition is called a grouping or classification. For instance, the groups \( \mathcal{X}_i \) can be chosen as

\[
\text{See also: Proceedings of the 2009 International Forum on Strategic Technologies - IFOST 2009, Vol. 'Automation and Mechatronics', Ho Chi Minh City University of Technology, October 21-23, 2009, Ho Chi Minh City, Vietnam, pp. 71-76.}
intervals $X_i = [x_{i-1}^*, x_i^*], i = 1, ..., m, x_0^* < x_1^* < ... < x_m^*$, then we have a so-called interval grouping (classification).

Now we assume that the random variables $X_1, X_2, ...$ can be observed in context of a given grouping only in a restricted manner as follows. Instead of $X_1, X_2, ...$ we observe discrete random variables $X_1^G, X_2^G, ...$, the so-called group numbers, where

$$ X_i^G = k \iff X_i \in \mathcal{X}_k, \quad k = 1, ..., m, $$

for $i = 1, 2, ...$ holds. Let $p_0^G(k)$ be the probability mass function of $X_i^G$ given by

$$ p_0^G(k) = P_0(X_i^G = k) = P_0(X_i \in \mathcal{X}_k), \quad k = 1, ..., m, \theta \in \Theta. $$

We start with an SLRT for $H_0 : \theta = \theta_0$ against

$$ H_1 : \theta = \theta_1, \quad \theta_0 < \theta < \theta_1 $$

based on sequence $\{X_i^G\}_{i=1}^\infty$. Let $\{L_n^G\}_{n=1}^\infty$ be the corresponding sequence of likelihood ratios

$$ L_n^G = \frac{p_0^G(Y_i^G)}{p_0^G(X_i^G)}, \quad n = 1, 2, ... $$

Then to given stopping bounds $B$ and $A$, $0 < B < A < \infty$, the sample size $N^G$ and the terminal decision rule $\delta^G$ are defined by

$$ N^G = \min\{n \geq 1 : L_n^G \notin (B, A)\} $$

and

$$ \delta^G = 1_{\{L_{\delta}^G \leq B\}}. $$

That means we continue sampling as long as for $n = 1, 2, ...$ the critical inequalities

$$ B < L_n^G < A $$

hold. If on stage $n$ for the first time $L_n^G \leq B$ or $L_n^G \geq A$ holds we stop sampling and decide for $H_0$ or $H_1$, respectively. We denote this test by $(N^G, \delta^G)$.

For further considerations it will be more convenient if we switch from likelihood ratios $L_n^G$ to

$$ Z_n^G = \ln L_n^G = \sum_{i=1}^n Y_i^G $$

with

$$ Y_i^G = \ln \frac{p_0^G(X_i^G)}{p_0^G(X_i^G)}, \quad i = 1, ..., n, $$

$n=1,2,...$ If stopping bounds $B$ and $A$ are transformed correspondingly by

$$ b = \ln B \quad \text{and} \quad a = \ln A $$

the critical inequality (1) becomes to

$$ b < \sum_{i=1}^n Y_i^G < a. $$

We now consider a discrete version of test $(N^G, \delta^G)$ which will serve as basic test for our CUSUM test here. We multiply critical inequality (1) by a constant $\gamma, \gamma > 0$. Then we get instead of (1)

$$ \gamma b < \sum_{i=1}^n \gamma Y_i^G < \gamma a $$

or

$$ 0 < -\gamma b + \sum_{i=1}^n \gamma Y_i^G < \gamma a - \gamma b. $$

By rounding according

$$ \hat{c} = \text{round}(-\gamma b), \quad s = \text{round}(\gamma(a - b)) $$

and

$$ \hat{Y}_i^G = \text{round}(\gamma Y_i^G), \quad i = 1, 2, .. $$

we obtain then a discrete or integer valued version of our test $(N^G, \delta^G)$ by

$$ \hat{N}^G = \min\{n \geq 1 : \hat{c} + \sum_{i=1}^n \hat{Y}_i^G \notin (0, s)\} $$

and

$$ \hat{\delta}^G = 1_{\{c + \sum_{i=1}^n \hat{Y}_i^G \leq 0\}}. $$

If $s$ is sufficiently large we get for the OC-functions $Q^G(\theta) = E_\theta \delta^G$ and $Q^G(\theta) = E_\theta \delta^G$ and the average sample number functions (ASN-functions) $E_\theta N^G$ and $E_\theta N^G, \theta \in \Theta$, of $(N^G, \delta^G)$ and $(\hat{N}^G, \delta^G)$

$$ Q^G(\theta) \approx \bar{Q}^G(\theta) \quad \text{and} \quad E_\theta N^G \approx E_\theta \hat{N}^G, $$

respectively.

More generally considered test $(\hat{N}^G, \hat{\delta}^G)$ can be regarded as a test which starts on stage 0 with start value $\tilde{c}$. In this sense we can get generalised tests $(\hat{N}^G, \hat{\delta}^G)$ for $k = 1, ..., s - 1$ which start on stage 0 with a start value $k, k = 1, ..., s - 1$. We have for $k = 1, ..., s - 1$

$$ \tilde{N}_k^G = \min\{n \geq 1 : k + \sum_{i=1}^n \tilde{Y}_i^G \notin (0, s)\} $$

and

$$ \tilde{\delta}_k^G = 1_{\{k + \sum_{i=1}^n \tilde{Y}_i^G \leq 0\}}. $$

The OC-functions $q_k^G = E_\theta \delta_k^G$ and ASN-functions $\delta_k^G = E_\theta \delta_k^G, k = 1, ..., s - 1$, of these tests can be computed to given $\theta \in \Theta$ by solving of systems of simultaneous linear equations. For more details we refer to [1],[3],[4] as well as [2].

III. THE CUSUM TEST

The SLRTs $(\hat{N}^G_k, \hat{\delta}_k^G), k = 1, ..., s - 1$, form the base for our CUSUM test. We start sampling with SLRT $(\hat{N}_k^G, \hat{\delta}_k^G)$ to any given start value $k \in \{1, ..., s - 1\}$. If this test accepts hypothesis $H_0$ we start a further SLRT with a new start value $k_0 \in \{1, ..., s - 1\}$. If this test again decides for $H_0$ we repeat this test. Hence we get a sequence of SLRTs

$$(\hat{N}_k^G, \hat{\delta}_k^G), (\hat{N}_{k_0}^G, \hat{\delta}_{k_0}^G), (\hat{N}_{k_0}^G, \hat{\delta}_{k_0}^G), ...$$
The procedure is finished if for the first time a test of this sequence leads to acceptance of $H_1$. Then sampling is stopped and this event is interpreted as an alert signal that parameter value $\theta$ has changed to a value $\theta > \theta^*$.

We denote this CUSUM test by $C(k, k_0)$.

Let $\hat{L}_k^G(k_0)$ be the random number of observations until an alert by our CUSUM test $C(k, k_0)$. This random sample size of a CUSUM test is denoted as so-called run length. Then, to given values $k, k_0 \in \{1, \ldots, s-1\}$ and $s$ the run time $\hat{L}_k^G(k_0)$ is given as follows. Let $\hat{W}_n^G(k, k_0)$ be random variables defined for $n = 0, 1, \ldots$ by

$$
\hat{W}_0^G(k, k_0) = k, \\
\hat{W}_n^G(k, k_0) = \begin{cases} k_0 & \text{if } \hat{W}_{n-1}^G(k, k_0) + \tilde{Y}_n^G \leq 0, \\
\hat{W}_{n-1}^G(k, k_0) + Y_n^G & \text{else.}
\end{cases}
$$

Then we get for the run length $\hat{L}_k^G(k_0)$ of our CUSUM test $C(k, k_0)$

$$
\hat{L}_k^G(k_0) = \min\{n \geq 1 : \hat{W}_n^G(k, k_0) \geq s\}.
$$

The most important characteristics in view of the statistical properties of a CUSUM test are the average run lengths (ARLs)

$$
\hat{L}_k^G(k_0) = E_\theta \hat{L}_k^G(k_0), \quad k = 1, \ldots, s-1, \quad \theta \in \Theta.
$$

This is the average numbers of observations until test variable $\hat{L}_k^G(k_0)$ reaches or exceeds the threshold $s$. This event is interpreted as a hint or an alert signal that parameter $\theta$ has changed from from a value $\theta \leq \theta^*$ to a value $\theta > \theta^*$.

If $\theta \in \Theta$ is the true parameter value and if $D_{\theta} X_1^G > 0$ then we have

$$
P_\theta(\hat{L}_k^G(k_0) < \infty) = 1 \quad \text{and} \quad E_\theta \hat{L}_k^G(k_0) < \infty.
$$

That means, even in case of $\theta \leq \theta^*$ if $D_{\theta} X_1^G > 0$, we will get with probability one an alert signal, a so-called false alert. False alerts are unavoidable here, but should occur as rarely as possible.

A. Direct computation of average run length

Analogously to SLRTs based on grouped observations (see [3],[2]) the ARLs $\tilde{L}_k^G(k_0), \ldots, \tilde{L}_{s-1}^G(k_0)$ can be obtained as solutions of a system of simultaneous linear equations which can be written as

$$
\tilde{L}_k^G(k_0) = 1 + \sum_{j=1}^{s-1} c_{kj} \tilde{L}_j^G(k_0) + a_{0,k} \tilde{G}_k(k_0), \quad k = 1, \ldots, s-1,
$$

see Appendix A. The coefficients $c_{kj}$ and $a_{0,k}$ are defined as follows: $c_{kj}$ is the transition probability from 'point' $k$ to 'point' $j$ during one observation step, that means

$$
c_{kj} = P_\theta(k + \tilde{Y}_j^G = j), \quad k, j = 1, \ldots, s-1.
$$

$a_{0,k}$ is the probability of acceptance of $H_0$ by SLRT $(\tilde{N}_k^G, \tilde{Y}_0^G)$ after next step or observation. In this case we continue sampling with test $(\tilde{N}_{k_0}^G, \tilde{Y}_{k_0}^G)$. It holds

$$
a_{k,0} = P_\theta(k + \tilde{Y}_j^G \leq 0), \quad k = 1, \ldots, s-1.
$$

Equation system (2) can be written in matrix form as

$$
(E - C - C_0)\tilde{L}_k^G(k_0) = \tilde{1}
$$

with $\tilde{L}_k^G(k_0) = \{\tilde{L}_j^G(k_0)\}_{j=1}^{s-1}$, $C = \{c_{ij}\}_{i=1, j=1}$ and $C_0 = (\tilde{0}, \tilde{0}, \tilde{0}, \ldots, \tilde{0})$.

where vector $\tilde{a}_0 = \{a_{0,k}\}_{k=1}^{s-1}$ forms the $k_0$-th column of matrix $C_0$, $\tilde{0} = \{0\}_{k=1}^{s-1}$, $\tilde{1} = \{1\}_{k=1}$ and $E$ denotes a suitable unit matrix.

Beside direct computation of ARL by solving of equation system (2) ARLs $\{\tilde{L}_k^G(k_0)\}_{k=1}^{s-1}$ can be obtained alternatively by means of operating characteristic functions (OC-functions) and average sample number functions (ASN-functions) of tests $(\tilde{N}_k^G, \tilde{Y}_0^G), k = 1, \ldots, s-1$. Denote by

$$
\tilde{q}_k^G = P_\theta(\text{Acceptance of } H_0 \text{ by } (\tilde{N}_k^G, \tilde{Y}_{k_0}^G)) = E_\theta \tilde{G}_k
$$

and

$$
\tilde{e}_k^G = E_\theta \tilde{N}_k^G
$$

the OC- and ASN-function of test $(\tilde{N}_k^G, \tilde{Y}_0^G), k = 1, \ldots, s-1, \theta \in \Theta$. Then we get analogously to (2) linear equation systems

$$
(E - C) \tilde{q} = \tilde{a}_0
$$

and

$$
(E - C) \tilde{e} = \tilde{1}
$$

with $\tilde{q} = \{\tilde{q}_k^G\}_{k=1}^{s-1}$ and $\tilde{e} = \{\tilde{e}_k^G\}_{k=1}^{s-1}$. Moreover, it holds for $k = 1, \ldots, s-1$ for the ARLs $\tilde{L}_k^G(k)$ of CUSUM tests $C(k, k)$:

$$
\tilde{L}_k^G(k) = \frac{\tilde{e}_k^G}{1 - \tilde{q}_k^G},
$$

see e.g. [4].

B. An approximation for average run length

By means of relation (4) we can obtain approximations for ARLs $\tilde{L}_k^G(k)$ based on the Wald approximations for OC- and ASN-function of test $(\tilde{N}_k^G, \tilde{Y}_0^G).

Let $(N_k^G, \delta_k^G)$ be a Wald’s SLRT for $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$ based on sequence $(X_i^G)_{i=1}^\infty$. If then stopping bounds $B$ and $A$ of test $(N_k^G, \delta_k^G)$ are chosen by

$$
B = \exp(-k/\gamma) \quad \text{and} \quad A = \exp((s-k)/\gamma)
$$

then we have for the OC-function $Q_k^G(\theta) = E_\theta \delta_k^G$ of test $(N_k^G, \delta_k^G)$

$$
Q_k^G(\theta_0) \approx 1 - \alpha \quad \text{and} \quad Q_k^G(\theta_1) \approx \beta
$$

with

$$
\alpha = \frac{1 - B}{A - B} \quad \text{and} \quad \beta = BA - 1.
$$

This implies with (4)

$$
E_\theta \tilde{L}_k^G(k) \approx \frac{(1 - \alpha)\ln B + \alpha \ln A}{\alpha} \cdot \frac{1}{E_\theta Y_i^G}
$$

and

$$
E_{\bar{\theta}} \tilde{L}_k^G(k) \approx \frac{\beta \ln B + (1 - \beta) \ln A}{1 - \beta} \cdot \frac{1}{E_{\bar{\theta}} Y_i^G}
$$

73
Fig. 1. Average run lengths \( \tilde{l}_{G_{k_0}}(k_0) = E_\theta \tilde{L}_{G_{k_0}}(k_0) \) for \( \theta = \theta_0 = -0.25, m = 2, 3, 5, 10 \) and \( s = 500 \).

with

\[
E_\theta Y^G_1 = \sum_{k=1}^{m} \ln \frac{p_{G_\theta}^G(k)}{p_{G_\theta_0}^G(k)}, \quad \theta \in \Theta.
\]

The expectation values \( E_{\theta_0} Y^G_1 \) and \( E_{\theta_1} Y^G_1 \) depend on grouping \( G \). Hence, this relations can be used to describe or to estimate how a given grouping \( G \) changes the ARL.

IV. EXAMPLE

We consider a CUSUM test for the mean of a normal distribution with known variance \( \sigma^2 = 1 \). The process be under control as long as for the mean \( \theta \) relation

\[
\theta \leq \theta^* = 0
\]

holds.

We assume, that instead of a sequence \( \{X_i\}_{i=1}^{\infty} \) of independent, \( N(\theta, \sigma^2) \)-distributed random observation variables only grouped observations \( \{X^G_i\}_{i=1}^{\infty} \) are available to a so-called \( F \)-optimal interval grouping scheme given by

\[
X'_1 = (-\infty, x^*_1), X'_2 = [x^*_2, x^*_3), ..., X'_m = [x^*_m-1, \infty).
\]

An interval grouping is said to be \( F \)-optimal (FISHER-optimal) for a given parameter value \( \theta \in \Theta \), if the group bounds \( x^*_1, ..., x^*_m-1, m \geq 2 \), maximise the FISHER information

\[
I^G_\theta(\theta) = E_\theta \left( \frac{\partial \ln p^G_\theta(X^G_1)}{\partial \theta} \right)^2.
\]

For more details of \( F \)-optimal grouping schemes we refer to [3] and [1].

Table I presents the \( F \)-optimal interval group bounds for the mean of a normal distribution for \( \theta = 0, \sigma^2 = 1 \) and \( m = 2, 3, 5, 10 \) and \( s = 500 \).

As basic test for our CUSUM procedure we consider an SLRT based on \( \{X^G_i\}_{i=1}^{\infty} \) for the mean of a normal distribution with hypotheses

\[
H_0 : \theta = \theta_0 = -0.25 \quad \text{and} \quad H_1 : \theta = \theta_1 = 0.25.
\]
Figure 1 shows the ARLs $I_{k0}^G(k_0) = E_{k0}L_{k0}^G(k_0)$ for $-0.25 \leq \theta \leq 0.25$, $m = 2, 3, 5, 10$ and $s = 500$. The discretisation parameters $\gamma$ and $s$ are

$$\gamma = 84.905818 \quad \text{and} \quad s = 500.$$  

Figure 1 shows the ARLs $I_{k0}^G(k_0) = E_{k0}L_{k0}^G(k_0)$, $k_0 = 1, \ldots, s - 1$, of a corresponding CUSUM test for $\theta = \theta_0 = -0.25$ and $m = 2, 3, 5, 10$. Figures 2 and 3 present the corresponding ARLs for $\theta = 0$ and $\theta = \theta_1 = 0.25$.

The efficiency of a CUSUM test can be evaluated beside ARLs $E_{k0}L_{k0}^G(k_0)$ and $E_{k0}L_{k0}^G(k_0)$ by the differences

$$E_{k0}L_{k0}^G(k_0) - E_{k0}L_{k0}^G(k_0)$$

for $k_0 = 1, \ldots, s - 1$. Figures 1 and 3 show that this difference is maximised for $k_0 = 1$. This recommends the value $k_0 = 1$ as start or recycling value for CUSUM tests considered here.

Figure 4 presents the ARLs $E_{k0}L_{k0}^G(k_0)$ for $k_0 = 1$ and $0.25 \leq \theta \leq 0.25$. Figure 5 presents the ARLs for $0 \leq \theta \leq 0.25$, where the CUSUM test should stop as soon as possible. It can be seen, how grouping increases ARL. Simultaneously this figure shows that influence of grouping is quite small even for small group numbers. Figure 5 contains also the ARL for an $F$-optimal grouping with $m = 50$ groups (dotted line). There is practically no essential difference in comparison with ARL for $m = 10$. The example underlines that hight efficient CUSUM tests can be obtained for $F$-optimal grouping schemes already with small number of groups. Finally, Table A presents some numerical values for the CUSUM test considered here.

**Appendix A**

**Proof of Equation (2)**

Consider ARLs $I_{k_0}^G(k_0) = E_{k0}L_{k0}^G(k_0)$ of CUSUM test $CS(k_0, k_0)$ for $k_0 = 1, \ldots, s - 1$. We get

$$I_{k_0}^G(k_0) = E_{k0}L_{k0}^G(k_0) = E_{k0}E_{k0}(L_{k0}^G(k_0)|k + \tilde{Y}_1^G)$$

$$= \sum_{j=-\infty}^{\infty} E_{k0}(\tilde{L}_{k0}^G(k_0)|k + \tilde{Y}_1^G = j)P_{k0}(k + \tilde{Y}_1^G = j)$$

$$= \sum_{j=-\infty}^{\infty} E_{k0}(\tilde{L}_{k0}^G(k_0)|C_{kj})c_{kj}$$

(5)

with $C_{kj} = \{k + \tilde{Y}_1^G = j\}$ and $c_{kj} = P_{k0}(C_{kj}) = P_{k0}(k + \tilde{Y}_1^G = j)$.

Now we consider three cases.

(i) $j = -\infty, \ldots, 0$: In this case the CUSUM test is recycled and we continue sampling by $CS(k_0, k_0)$ based on $\{\tilde{Y}_i^G\}_{i=2}^{\infty}$. Let $L_{k0}^G(k_0)$ be the corresponding run length. Then, because of the iid-property of $\{\tilde{Y}_i^G\}_{i=2}^{\infty}$ we get $E_{k0}L_{k0}^G(k_0) = E_{k0}L_{k0}^G(k_0) = L_{k0}^G(k_0)$ and we have

$$\sum_{j=-\infty}^{0} E_{k0}(L_{k0}^G(k_0)|C_{kj})c_{kj} = \sum_{j=-\infty}^{0} (1 + E_{k0}L_{k0}^G(k_0)))c_{kj}$$

$$= \sum_{j=-\infty}^{0} c_{kj} + L_{k0}^G(k_0)a_{0,k}$$

(6)
with

\[ a_{k,0} = \sum_{j=-\infty}^{0} c_{kj} = \sum_{j=-\infty}^{0} P_\theta(k+\bar{Y}_{1}^G = j) = P_\theta(\bar{Y}_{1}^G \leq j-k). \]

(ii) \( j = 1, \ldots, s-1 \): Here we continue sampling with CS \((j, k_0)\) based on \( \{\bar{Y}_{1}^G\}_{j=2}^{\infty} \). Let \( \bar{L}_j(k_0) \) be again the further run length, then under condition \( C_{kj} \) and by iid-property we get

\[ E_\theta \bar{L}_j(k_0) = E_\theta \bar{L}_j^G(k_0) = E_\theta^G \bar{L}_j^G(k_0) \]

and

\[ \sum_{j=1}^{s-1} E_\theta(\bar{L}_j^G(k_0))C_{kj} c_{kj} = \sum_{j=1}^{s-1} (1 + E_\theta \bar{L}_j(k_0))c_{kj} = \sum_{j=1}^{s-1} c_{kj} + \sum_{j=1}^{s-1} \bar{G}_j(k_0)c_{kj}. \] 

(iii) \( j = s, \ldots, \infty \): In this case threshold \( s \) is reached or overcrossed and our CUSUM test is stopped. This implies

\[ E_\theta(\bar{L}_j^G(k_0)C_{kj})c_{kj} = 1 \quad \text{and} \quad \sum_{j=s}^{\infty} E_\theta(\bar{L}_j^G(k_0))C_{kj} c_{kj} = \sum_{j=s}^{\infty} c_{kj}. \]

Then, (6), (7) and (8) together with (5) and \( \sum_{j=-\infty}^{\infty} c_{kj} = 1 \) complete the proof.

### TABLE I

<table>
<thead>
<tr>
<th>m</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_1 )</td>
<td>0.0000</td>
<td>-0.6120</td>
<td>-0.9816</td>
<td>-1.2444</td>
<td>-1.4469</td>
</tr>
<tr>
<td>( x_2 )</td>
<td>0.6120</td>
<td>0.0000</td>
<td>-0.3823</td>
<td>-0.6589</td>
<td>-0.9656</td>
</tr>
<tr>
<td>( x_3 )</td>
<td>0.9816</td>
<td>0.3823</td>
<td>0.0000</td>
<td>0.1837</td>
<td>0.4047</td>
</tr>
<tr>
<td>( x_4 )</td>
<td>1.2444</td>
<td>0.6589</td>
<td>0.1837</td>
<td></td>
<td></td>
</tr>
<tr>
<td>%</td>
<td>63.66</td>
<td>80.98</td>
<td>88.25</td>
<td>92.01</td>
<td>94.20</td>
</tr>
</tbody>
</table>

### TABLE II

<table>
<thead>
<tr>
<th>( \theta \backslash m )</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>10</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.30</td>
<td>4963</td>
<td>15251</td>
<td>14581</td>
<td>2096</td>
<td>2118</td>
<td>2186</td>
</tr>
<tr>
<td>-0.25</td>
<td>6114</td>
<td>6015</td>
<td>5979</td>
<td>5686</td>
<td>4964</td>
<td>5052</td>
</tr>
<tr>
<td>-0.20</td>
<td>2672</td>
<td>2535</td>
<td>2350</td>
<td>2368</td>
<td>2096</td>
<td>2118</td>
</tr>
<tr>
<td>-0.15</td>
<td>1267</td>
<td>1160</td>
<td>1072</td>
<td>1071</td>
<td>960.4</td>
<td>964.1</td>
</tr>
<tr>
<td>-0.10</td>
<td>659.5</td>
<td>585.2</td>
<td>539.4</td>
<td>534.7</td>
<td>484.6</td>
<td>483.7</td>
</tr>
<tr>
<td>-0.05</td>
<td>379.1</td>
<td>327.8</td>
<td>301.8</td>
<td>297.1</td>
<td>271.6</td>
<td>269.9</td>
</tr>
<tr>
<td>-0.00</td>
<td>240.0</td>
<td>203.5</td>
<td>187.3</td>
<td>183.4</td>
<td>168.7</td>
<td>167.1</td>
</tr>
<tr>
<td>0.05</td>
<td>165.5</td>
<td>138.4</td>
<td>127.4</td>
<td>124.3</td>
<td>114.8</td>
<td>113.5</td>
</tr>
<tr>
<td>0.10</td>
<td>122.4</td>
<td>101.5</td>
<td>93.4</td>
<td>90.9</td>
<td>84.2</td>
<td>83.1</td>
</tr>
<tr>
<td>0.15</td>
<td>95.7</td>
<td>78.8</td>
<td>72.6</td>
<td>70.5</td>
<td>65.4</td>
<td>64.5</td>
</tr>
<tr>
<td>0.20</td>
<td>78.0</td>
<td>64.0</td>
<td>59.0</td>
<td>57.2</td>
<td>53.0</td>
<td>52.2</td>
</tr>
<tr>
<td>0.25</td>
<td>65.7</td>
<td>53.7</td>
<td>49.5</td>
<td>47.9</td>
<td>44.4</td>
<td>43.8</td>
</tr>
<tr>
<td>0.30</td>
<td>56.7</td>
<td>46.2</td>
<td>42.6</td>
<td>41.2</td>
<td>38.2</td>
<td>37.6</td>
</tr>
</tbody>
</table>

### REFERENCES


