STAPRO 3782
pp: 1--10 (col.fig.: Nil)
PROD. TYPE: COM
ED: Mangala
PAGN: Anitha – SCAN: Raghwho
ARTICLE IN PRESS
Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT
Statistics & Probability Letters ▮▮▮ (▮▮▮▮) ▮▮▮–▮▮▮
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

A unified Markov chain approach for computing the run length distribution in control charts with simple or compound rules☆

James C. Fu[a], Galit Shmueli[b,*], Y.M. Chang[a]

[a] *Department of Statistics, University of Manitoba, Manitoba, Canada*
[b] *Department of Decision & Information Technologies, Robert H. Smith School of Business, University of Maryland, 4361 Van Munching Hall, College Park, MD 20742-1815, USA*

Abstract

We introduce a general theoretical framework, based on the Markov chain imbedding approach, that leads to the run-length distribution for a multitude of control charts that are based on a simple rule or on a compound set of rules.

Keywords: Control charts; Markov chain; Transition matrix; Discretization

1. Introduction

The multitude of control charts for monitoring various process parameters (such as the mean, variance, and proportion) exists due to the multiple types of shifts that can occur in that parameter. There is no single chart that is optimal in detecting all types of shifts. Sometimes several charts are used simultaneously, while in other cases new combined charts are used (e.g., robust Cusum charts and Shewhart charts with runs rules).

Various control charts have been investigated yielding the run-length distribution for a given pattern (usually a step function), using various methods. Comparisons between some of these charts have also been done using different methods: from simulations (Roberts, 1959), through numerical analyses (Robinson and Ho, 1978; Luceno and Puig-Pey, 2000; Rao et al., 2001) to theoretical approximations and exact derivations.

☆ This work was supported in part by Nature Science and Engineering Research Council of Canada Grant A-9216.
* Corresponding author. Tel.: +1-301-405-9679; +1-301-405-8655.
E-mail address: gshmueli@rhsmith.umd.edu (G. Shmueli).

0167-7152/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.spl.2003.10.004

Among the theoretical derivations, many authors have used the Markov chain approach (Champ and Woodall, 1987; Lucas and Saccucci, 1990; Lucas and Crosier, 1982b), which was introduced by Brook and Evans (1972). Since each author focused on one or more charts, the different Markov chain applications were tailored to each case. For instance, the state space of the Markov chain has been formulated differently by different authors (Champ and Woodall, 1987; Lucas and Crosier, 1982a), and in several cases it was not specified at all.

We introduce a general unified framework that is based on discretization and the use of a finite Markov chain imbedding technique for run statistics (Fu, 1996). The method can be applied to any control scheme that is based either on a simple boundary crossing rule, or on a compound rule based on run or scan statistics that include several criteria. Some known results can be viewed as special cases of our general method.

Our unified approach sheds light on the relation between different types of monitoring schemes, and on their performance in the presence of different data structures. It also enables a straightforward comparison of performance of various schemes, thus being important from an applied point of view as well.

The paper is laid out as follows: Section 2, describes the general framework, which is based on a Markov chain imbedding formulation. We show how some well-known charts can be formulated in this framework. In Section 3, we provide a detailed numerical example of compound rule based on a run statistic and Cusum to demonstrate how to imbed it as a Markov chain. In Section 4, we discuss the implication and possible extensions of this general approach.

2. The Markov chain approach

Brook and Evans (1972) introduced a Markov chain representation for computing the run-length distribution of a Cusum chart. Their basic idea for a discrete monitoring statistic (e.g., a count) is to treat the m values that the monitoring statistic can obtain within the control limits as states of a Markov chain, and all the values that exceed the limits as an absorbing state. If the monitoring statistic is continuous, the same method is used, after discretizing the area of the control chart into m regions within the control limits and one region that exceeds the limits (the absorbing state).

We formalize and generalize this method as follows: for a control scheme that involves a compound decision rule, the run length is imbedded into a finite Markov chain Y_t, according to the rules applied (denoted by $\phi_1, \ldots, \phi_l$). Y_t then has a state space Ω and transition matrix $\mathbf{M}_t$ of the form

$$\mathbf{M}_t = \left[\begin{array}{c|c} \mathbf{N}_t & \mathbf{C}_t \\ \hline \mathbf{0} & \mathbf{I} \end{array}\right]. \tag{1}$$

The run-length probability is then given by

$$P(RL = n) = \pi_0 \left(\prod_{t=1}^{n-1} \mathbf{N}_t \right) (\mathbf{I} - \mathbf{N}_n)\mathbf{1}', \tag{2}$$

where $\pi_0 = (1, 0, \ldots, 0)$ and $\mathbf{1} = (1, 1, \ldots, 1)$.

We denote a simple rule by ϕ and the monitoring statistic by $W_t(k) = [W_{t-k+1}, \ldots, W_t]$, where k is the length of the "history" that is retained in order to reach a decision at time t.

A compound rule dictates that the chart signals an alarm at time t if any one rule $\phi_i(W_t(k))$ based on the monitoring statistic $W_t(k) = [W_{t-k+1}, \ldots, W_t]$ exceeds some limit at time t (or falls within a certain area on the chart).

For this general pattern we define imbedded Y_t as

$$Y_t(W_t(k); \phi_1, \ldots, \phi_l) = H(\phi_i(W_t(k)); \; i = 1, \ldots, l), \tag{3}$$

where H is a function that combines the information from the l different rules. In general, the state space Ω of the imbedded Markov chain Y_t is induced by ϕ_i and the vector $W_t(k)$.

2.1. Discretizing W_t

Although all different uses of the Markov chain approach rely on discretizing a continuous variable, we make an important distinction between two types of discretization: natural vs. artificial. Discretizing a random variable can either arise naturally from the monitoring scheme, or else it is artificially imposed. Examples where discretization arises naturally are charts with discrete monitoring statistics (such as counts) and charts with a continuous monitoring statistic which effectively divide the values of the statistic into two or more regions (e.g., Shewhart charts with or without runs rules).

Examples where discretization is carried out artificially are Cusum and EWMA charts (Brook and Evans, 1972; Fu et al., 2001). The reason for imposing an artificial discretization is to simplify the calculation of a complicated probability. For example, in Cusum and EWMA schemes it is hard to calculate the probability that the monitoring statistic exceeds the boundary. The alternative, which is based on discretizing the continuous measurements, is described in Section 2.2.2.

For W_t, the monitoring statistic at time t, we use $R(W_t)$ to denote a natural discretization of W_t and $D(W_t)$ an artificial discretization of W_t. A rule $\phi(W_t(k))$ determines which type of discretization is used, and the number of last values k that should be retained.

2.2. Imbedding well-known monitoring schemes

The Markov chain approach which was introduced by Brook and Evans (1972) was used by several authors to derive the average run length (ARL) or the entire run length distribution for various control charts. We show how the different results can be formalized as described above, and how the run length is imbedded into a finite Markov chain. We describe rules that lead to natural discretization, rules that require artificial discretization, and rules that involve both types of discretization.

2.2.1. Class 1: rules that lead to natural discretization

The class of monitoring schemes where the decision rule leads to a natural discretization of W_t includes schemes that are based on a discrete statistic and schemes where the control chart is divided into discrete regions.

The “history” that is retained in this case has the form $[R(W_{t-k+1}), \ldots, R(W_t)]$, where k is the number of previous values that must be tracked in order to reach a decision. Examples of charts with naturally discretizing rules and different values of k are:

- A Shewhart chart, where we only retain information of X_t at time t, i.e., $R(X_t)$.
- A discrete Cusum chart such as a Poisson Cusum, where the monitoring statistic is a function of the observation X_t and of the cumulative sum S_{t-1}. In such cases we retain information on times $t-1$ and t, i.e., $[R(S_{t-1}), R(X_t)]$.
- A Shewhart chart with Western Electric rules, where the decision rule is based on information of the locations of the k previous statistics. The following set of Western Electric rules (see Montgomery, 2001) are widely applied: Signal an alarm if
  1. One or more points exceed the 3-sigma control limits,
  2. Two of three-consecutive point fall beyond the 2-sigma limits,
  3. Four of five-consecutive points fall beyond the 1-sigma limits,
  4. Eight consecutive points fall on one side of the center line.

If all the four rules are combined then we must retain the locations of the last 8 points, i.e. $[R(X_{t-7}), R(X_{t-6}), \ldots, R(X_t)]$.

According to the rules ϕ_i, $i=1,\ldots,l$ (with or without runs rules), the statistic $W_t(k)$ is imbedded into a Markov chain $Y_t(W_t(k);\phi_1,\ldots,\phi_l)$. A simple example is a Shewhart chart where X_t is naturally discretized into an indicator taking two possible levels: α (outside the limits) and r (within the limits). The state space for $Y_t(W_t(1);\phi)=R(X_t)$ is $\Omega=\{\emptyset, r, \alpha\}$, and the corresponding transition matrix is given by

$$\mathbf{M} = \left[\begin{array}{cc|c} 0 & p_r & 1-p_r \\ 0 & p_r & 1-p_r \\ \hline 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{c|c} \mathbf{N} & \mathbf{C} \\ \hline \mathbf{0} & 1 \end{array}\right], \tag{4}$$

where $\emptyset$ is the dummy initial state for the imbedded Markov chain and $p_r = P(\mathrm{LCL} < X_t < \mathrm{UCL})$, under a given value of the monitored parameter.

2.2.2. Class 2: rules that require artificial discretization

When the monitoring statistic is continuous there are cases where it is too complicated to compute the run-length distribution directly. Two such examples are the Cusum and EWMA charts. An alternative is to discretize X_t artificially. This leads to a discrete Cusum/EWMA statistic W_t, and we retain the last k discrete values $[D(W_{t-k+1}),\ldots,D(W_t)]$.

For example, a one-sided Cusum chart can be discretized such that the Cusum statistic S_t obtains $m+2$ values (or $2(m+1)+1$ values for a two-sided chart): The in-control area $[0,h)$ is divided into m equally-sized regions and the out-of-control area $[h,\infty)$ is the $m+1$ region (as described in Brook and Evans, 1972; Fu et al., 2001). For both the Cusum and EWMA schemes we retain information on the accumulating statistic (denoted by S) at time $t-1$ and on the accumulated statistic (denoted by X) at time t, i.e. $k=2$ and we retain 2 last discrete values $[D(S_{t-1}), D(X_t)]$. This information is imbedded into a Markov chain $Y_t(W_t(2);\phi) = H(\phi(W_t(2))) = D(S_t)$.

The transition matrix includes $m+3$ states. By selecting m to be large enough, the run length of the discretized statistic will approach that of the continuous one.

2.2.3. Class 3: rules that involve natural and artificial discretization

This class is the most general. It includes decision rules that require monitoring a naturally discrete statistic and a continuous one. Two examples are the "robust Cusum chart" (Lucas and Crosier, 1982a) and the "robust EWMA chart" (Lucas and Saccucci, 1990) which are combinations of a Shewhart chart (with 4 or 5-sigma limits) and a Cusum or EWMA chart. Two statistics are tracked simultaneously: The Cusum/EWMA continuous statistic (S_t) and the naturally-discretized Shewhart statistic (X_t).

The decision rule is compound: signal an alarm if the continuous (Cusum or EWMA) statistic exceeds the (Cusum/EWMA) limits, or if two consecutive Shewhart statistics exceed the Shewhart chart limits. The information that is needed in order to reach a decision is based on the last Cusum/EWMA value and the last two Shewhart statistics: $W_t(2)=[S_{t-1},X_t]$. Here the Cusum/EWMA statistic is artificially discretized, and the Shewhart statistic is naturally binary:

$$\phi_1(W_t(2)) = D(S_t) \quad \text{and} \quad \phi_2(W_t(2)) = R(X_t). \tag{5}$$

These two rules are then combined and imbedded into a Markov chain Y_t.

Other hypothetical decision rules that fall into this category would be Cusum or EWMA schemes with runs rules.

3. An example of a compound rule that involves natural and artificial discretization

To illustrate this general case consider a hypothetical one-sided Cusum chart, which in addition to the upper control limit has a "warning limit". The compound decision rule is to signal at time t if the Cusum statistic, given by $S_t = \max(0, S_{t-1} + X_t)$ exceeds the upper control limit at time t; or if two within three consecutive Cusum statistics fall in the interval between the warning and control limits. We assume that X_t are i.i.d. $N(0,1)$.

As in the ordinary Cusum chart, for computational reasons the accumulated statistic X is artificially discretized with step Δ:

$$D(X) = i\Delta, \quad i = 0, \pm 1, \pm 2, \ldots, \pm (m+1). \tag{6}$$

We define $p_i = P(D(X) = i\Delta)$ and $F(i) = P(D(X) \leqslant i\Delta)$ as follows: for $i = 0, \pm 1, \ldots, \pm m$

$$p_i = \int_{(i-.5)\Delta}^{(i+.5)\Delta} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-(1/2)x^2}\,\mathrm{d}x, \qquad p_{m+1} = \int_{[(m+1)-.5]\Delta}^{\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-(1/2)x^2}\,\mathrm{d}x,$$

$$p_{-(m+1)} = \int_{-\infty}^{[-(m+1)+.5]\Delta} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-(1/2)x^2}\,\mathrm{d}x, \qquad F(i) = \sum_{j=-(m+1)}^{i} p_j.$$

This results in a discrete Cusum statistic S. We denote the upper control limit by $h = (m+1)\Delta$ and the warning limit by $h^* = m^*\Delta$, $m^* \leqslant m$. Then S_t can obtain the values $i\Delta$, $i = 0, 1, \ldots, m+1$ in the interval $[0, h]$.

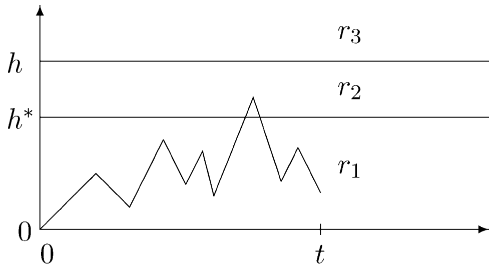

Fig. 1. An illustration of the three-region chart resulting from a one-sided Cusum with a warning limit.

In addition, the chart in this example is naturally divided by the second rule into three regions (see Fig. 1)

$$R(S_t) = \begin{cases} r_1 & \text{if } 0 \leqslant S_t < h^*, \\ r_2 & \text{if } h^* \leqslant S_t < h, \\ r_3 & \text{if } S_t \geqslant h. \end{cases} \tag{7}$$

The "history" that is required in this scheme in order to reach a decision is $k = 3$, with $W_t(3) = [S_{t-2}, S_{t-1}, S_t]$. We can thus write the two rules as:

$$\phi_1(W_t(3)) = S_t \quad \text{and} \quad \phi_2(W_t(3)) = [R(S_{t-1}), R(S_t)]. \tag{8}$$

The combined rules induce the state space of a Markov chain of the form:

$$Y_t(W_t(3); \phi_1, \phi_2) = H(\phi_1(W_t(3)),\ \phi_2(W_t(3))) = [R(S_{t-1}), S_t], \tag{9}$$

where $R(S_t) = r_1, r_2$ or r_3 and $S_t = i\Delta, i = 0, 1, \ldots, m+1$. For simplicity we write $S_t = i$ to denote $S_t = i\Delta$. States that include r_3 or $m+1$ as one of their coordinates can be combined into α, the absorbing state. Hence, the state space is:

$$\Omega = \{(\emptyset, \emptyset), (\emptyset, 0), \ldots, (\emptyset, m), (r_1, 0), \ldots, (r_1, m), (r_2, 0), \ldots, (r_2, m^* - 1), \alpha\} \tag{10}$$

with $2 + 2(m+1) + m^*$ states.

Remark 3.1. Here we assumed that $P(X_0 = \emptyset) \equiv 1$ as initial distribution of X_0. Hence the states $(\emptyset, \emptyset), (\emptyset, 0), \ldots, (\emptyset, m)$ are required so that Y_0 and Y_1 of the imbedded Markov chain $\{Y_t\}$ can be properly defined with the initial distribution given by $P(Y_0 = (\emptyset, \emptyset)) \equiv 1$.

If Y_t is not in the absorbing state, it will *not* move into α if one of the three conditions is met: (i) $S_{t+1} < h^*$, (ii) $\{h^* \leqslant S_{t+1} < h,\ R(S_{t-1}) = \emptyset$ and $R(S_t) = r_1\}$, or (iii) $\{h^* \leqslant S_{t+1} < h,\ R(S_{t-1}) = R(S_t) = \emptyset$

or $r_1\}$. In that case,

$$P\{Y_{t+1}=[R(S_t),S_{t+1}]\,|\,Y_t=[R(S_{t-1}),S_t]\}$$
$$=\begin{cases} P(X_{t+1}=-S_t) & \text{if } S_{t+1}=0,\\ P(X_{t+1}=S_{t+1}-S_t) & \text{if } 0<S_{t+1}<h^*,\\ P(X_{t+1}=S_{t+1}-S_t) & \text{if condition (ii) or (iii) hold},\\ 0 & \text{otherwise.}\end{cases} \quad (11)$$

Since the probability of each row is 1, the probability of entering the absorbing state can be obtained by subtracting all the positive probabilities in the row from 1.

To illustrate such a transition matrix $\mathbf{M}$, we choose $m=3$, $h=4$ $(\Delta=h/(m+1)=1)$ and $h^*=2$ $(m^*=h^*/\Delta=2)$. Then, $D(X_t)=i\Delta=i$, $i=0,\pm 1,\pm 2,\ldots,\pm 4$, and S_t takes the values $i=0,1,2,3,4$ (with 4 denoting the absorbing state α). $\mathbf{M}$ is given by

	$(\emptyset,\emptyset)$	$(\emptyset,0)$	$(\emptyset,1)$	$(\emptyset,2)$	$(\emptyset,3)$	$(r_1,0)$	$(r_1,1)$	$(r_1,2)$	$(r_1,3)$	$(r_2,0)$	$(r_2,1)$	α
$(\emptyset,\emptyset)$		$F(0)$	p_1	p_2	p_3							$1-F(3)$
$(\emptyset,0)$						$F(0)$	p_1	p_2	p_3			$1-F(3)$
$(\emptyset,1)$						$F(-1)$	p_0	p_1	p_2			$1-F(2)$
$(\emptyset,2)$										$F(-2)$	p_{-1}	$1-F(-1)$
$(\emptyset,3)$										$F(-3)$	p_{-2}	$1-F(-2)$
$(r_1,0)$						$F(0)$	p_1	p_2	p_3			$1-F(3)$
$(r_1,1)$						$F(-1)$	p_0	p_1	p_2			$1-F(2)$
$(r_1,2)$										$F(-2)$	p_{-1}	$1-F(-1)$
$(r_1,3)$										$F(-3)$	p_{-2}	$1-F(-2)$
$(r_2,0)$						$F(0)$	p_1					$1-F(1)$
$(r_2,1)$						$F(-1)$	p_0					$1-F(0)$
α												1

and it can always be written in the form given in (4).

Fig. 2 gives the run-length probability distributions for the case when $h=3$ and $h^*=2$ with different values of m. The mean and standard deviation of run length are computed by using the following formulas

$$E[RL]=\pi_0(\mathbf{I}-\mathbf{N})^{-1}\mathbf{1}' \quad \text{and} \quad E[RL^2]=\pi_0(\mathbf{I}+\mathbf{N})(\mathbf{I}-\mathbf{N})^{-2}\mathbf{1}'$$

from Fu et al. (2001), Theorem 1(iii). Numerical results are given in Table 1.

Remark 3.2. Traditionally, the mean and standard deviation of the run-length distribution are used for comparing the performance of control charts. However, in view of our numerical result that the run-length distribution for a compound control rule is rather skewed, we feel that displaying the quantiles of the distribution is more adequate. In general, the distribution resulting from a compound control rule is always highly skewed toward the right, especially when it involves several control charts.

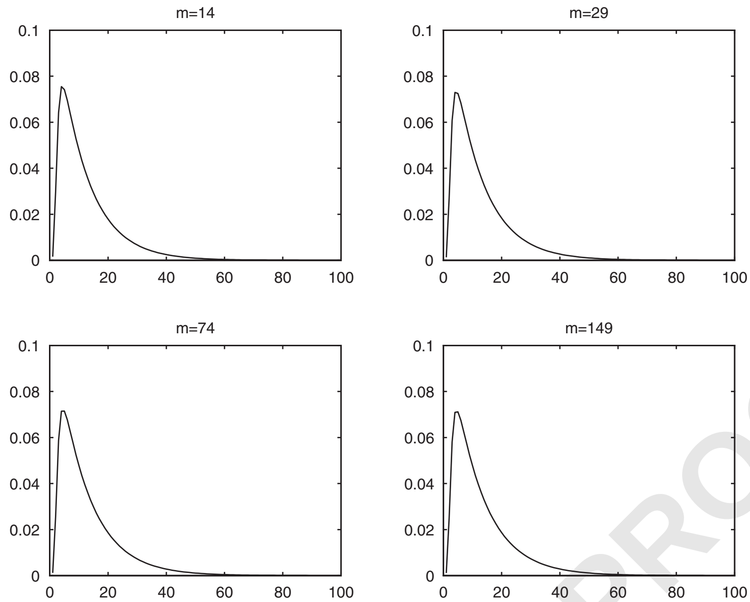

Fig. 2. Probability distributions of the run length for a one-sided Cusum with 3-sigma control limit and 2-sigma warning limit and a compound decision rule, at $m = 14, 29, 74, 149$.

Table 1
Quartiles, mean, and standard deviation of the run length for a one-sided Cusum with 3-sigma control limit and 2-sigma warning limit and a compound decision rule, at different levels of discretization (m)

m	Q_1	Q_2	Q_3	Mean	Std
5	4.626	8.444	14.904	11.739	9.386
14	5.066	9.213	16.232	12.749	10.187
29	5.223	9.486	16.701	13.103	10.473
74	5.316	9.649	16.976	13.319	10.649
149	5.348	9.703	17.075	13.392	10.709
299	5.364	9.730	17.124	13.428	10.738
749	5.373	9.747	17.154	13.450	10.756
1499	5.376	9.753	17.164	13.457	10.762
1874	5.376	9.753	17.164	13.459	10.763

4. Discussion

In many cases two or more control charts are used simultaneously, either to monitor several parameters (e.g., the mean and the standard deviation), or to be able to detect different sizes or types of shifts (e.g., a Cusum and an X-bar simultaneously). It is important to know in these cases about the signaling behavior of the joint charts.

The combined run length is the time until the first signal is raised by any one of the charts. Mathematically, it is a waiting time problem with $T = \min(T_1, \ldots, T_r)$, where $T_i,\ i = 1, \ldots, r$ are waiting times of r charts (or decision rules). When the monitoring charts are independent, the distribution of the combined run-length can be easily derived. In the dependent case, the combined run-length distribution of a compound rule is more complicated. Even in this situation, it still can be incorporated into the general Markov chain imbedding framework. For example, the additional Western Electric rule ϕ_4 "Eight consecutive points fall on one side of the centerline" can be easily incorporated into the example of the compound rule of Section 3 by using an additional coordinate with states $0, 1, \ldots, 7$ and state 8 as absorbing state. We leave the details to the reader. Further, with a simple modification of transition matrix, the results in Section 3 can be extended to the case when the sequence $\{X_t\}$ has a Markovian dependence structure.

The unified approach can be used to learn about a chart's ability to detect different types of signals. In order to study the performance of some monitoring scheme for a particular parameter pattern (e.g., a step-shift or linear trend from the target value), we can integrate a given pattern into the Markov chain and compute the run-length distribution. In comparison to the case of a constant parameter, the Markov chain is no longer time homogeneous. This means that the transition probabilities will now depend on the value of the parameter at time t. A simple example would be to incorporate a simulated pattern of the process mean μ_t into a Shewhart scheme. The transition matrix, $\mathbf{M}_t$, would be the same as (4), except that p_r would depend on time, i.e. $p_{r,t}$.

Note that in this paper we did not investigate discretization errors or error bounds. They are in general extremely complex and depend not only on m and n but also on the structure of the transition probability matrix N of the imbedded Markov chain associated with the compound control rule. From our experience in computing the distributions of RL and the ARL when $m > 500$, the numerical results are very stable. This can be seen in Fig. 2 and Table 1. In view of this, we suggest using $m = 500$ for numerical computations of distributions or ARL values. For a given large m, it can be shown by using the Perron–Frobenius Theorem (see Seneta, 1981), that

$$P(\mathrm{RL} > n) = c \exp(n \log \lambda_{[1]})(1 + \mathrm{O}(\mathrm{e}^{n \log |\lambda_{[2]}/\lambda_{[1]}|}), \tag{12}$$

where c is a constant independent of n, and $\lambda_{[1]}, \lambda_{[2]}$ are the largest and second largest eigenvalues of the transition matrix N, respectively (see also Fu and Lou, 2003).

In conclusion, the general framework described here can be used for framing a multitude of control charts into a Markov chain imbedding setting for the purpose of computing the run-length distribution. It extends to combinations of charts, to dependence, and to many other monitoring schemes both hypothetical and ones that are suggested in the control chart literature.

References

Brook, D., Evans, D.A., 1972. An approach to the probability distribution of CUSUM run length. Biometrika 59, 539–549.

Champ, C.W., Woodall, W.H., 1987. Exact results for Shewhart control charts with supplementary runs rules. Technometrics 29 (4), 393–399.

Fu, J.C., 1996. Distribution theory of runs and patterns associated with a sequence of multi-state trials. Statist. Sinica 6 (4), 957–974.

Fu, J.C., Lou, W.Y.W., 2003. Distribution Theory of Runs and Patterns and Its Applications. World Scientific, Singapore.

Fu, J.C., Spiring, F.A., Xie, H., 2001. On the average run lengths of quality control schemes using a Markov chain approach. Statist. Probab. Lett. 56, 369–380.

Lucas, J.M., Crosier, R.B., 1982a. Robust CUSUM: a robustness study for Cusum quality control schemes. Comm. Statist. — Theory Methods 11 (23), 2669–2687.
Lucas, J.M., Crosier, R.B., 1982b. Fast initial response for CUSUM quality-control schemes: give your CUSUM a head start. Technometrics 24 (3), 199–205 (reprinted in 2000).
Lucas, J.M., Saccucci, M.S., 1990. Exponentially weighted moving average control schemes: properties and enhancements. Technometrics 32 (1), 1–12.
Luceno, A., Puig-Pey, J., 2000. Evaluation of the run-length probability distribution for CUSUM charts: assessing chart performance. Technometrics 42 (4), 411–416.
Montgomery, D.C., 2001. Introduction to Statistical Quality Control, 4th Edition. Wiley, New York.
Rao, B.V., Disney, R.L., Pignatiello Jr., J.J., 2001. Uniqueness and convergence of solutions to average run length integral equation for cumulative sum and other control charts. IIE Trans. 33 (6), 463–469.
Roberts, S.W., 1959. Control chart tests based on geometric moving averages. Technometrics 1 (3), 239–250.
Robinson, P.B., Ho, T.Y., 1978. Average run lengths of geometric moving average charts by numerical methods. Technometrics 20 (1), 85–94.
Seneta, E., 1981. Non-negative Matrices and Markov Chains, 2nd Edition. Springer, New York.