Efficient Design and Sensitivity Analysis of Control Charts Using Monte Carlo Simulation

Michael C. Fu • Jian-Qiang Hu
The Robert H. Smith School of Business, Institute for Systems Research, University of Maryland, College Park, Maryland 20742-1815
mfu@umd5.umd.edu
Department of Manufacturing Engineering, Boston University, Boston, Massachusetts 02215
hu@enga.bu.edu

The design of control charts in statistical quality control addresses the optimal selection of the design parameters (such as the sampling frequency and the control limits) and includes sensitivity analysis with respect to system parameters (such as the various process parameters and the economic costs of sampling). The advent of more complicated control chart schemes has necessitated the use of Monte Carlo simulation in the design process, especially in the evaluation of performance measures such as average run length. In this paper, we apply two gradient estimation procedures—perturbation analysis and the likelihood ratio/score function method—to derive estimators that can be used in gradient-based optimization algorithms and in sensitivity analysis when Monte Carlo simulation is employed. We illustrate the techniques on a general control chart that includes the Shewhart chart and the exponentially-weighted moving average chart as special cases. Simulation examples comparing the estimators with each other and with "brute force" finite differences demonstrate the possibility of significant variance reduction in settings of practical interest.
(*Statistical Quality Control*; *Control Charts*; *Average Run Length*; *Sensitivity Analysis*; *Economic Design Problem*; *Monte Carlo Simulation*; *Perturbation Analysis*; *Likelihood Ratio/Score Function Method*)

1. Introduction

Two critical issues in the design of control charts in statistical quality control are the *optimal* selection of the design parameters—such as sample size, sampling frequency, and control limits—and *sensitivity analysis* with respect to system parameters such as the economic costs of sampling and the characteristics of the potential process shifts (cf. Montgomery 1996). Depending on the design approach, the performance measures of interest fall into two main types: average run lengths or expected economic costs. The increasing complexity of many of the recently proposed control charts in the research literature has led to analytically intractable models, so Monte Carlo simulation is routinely used to estimate performance. Examples include Grimshaw and Alt (1997), where control charts for quantile function values are proposed; Albin et al. (1997), where a number of different control charts are compared in their average run length to false alarms and to detection of process mean and standard deviation shifts; and Baxley (1995), where variable sampling interval control charts are applied.

0025-1909/99/4503/0395$05.00

The generality of Monte Carlo simulation makes it a popular tool, since it allows the modeller to be quite flexible. Among the clearest advantages are the following.

• Assumptions on process characteristics can be relaxed (e.g., normality, independence, and stationarity assumptions). For example, in Grimshaw and Alt (1997), a control chart is derived under a large sample approximation invoking the Central Limit Theorem. However, they point out that, in practice, relatively small sample sizes are used. In their small test example, the simulation estimate of in-control average run length was 147.6, compared with the large sample approximation of 200.

• Any control chart can be handled, including Shewhart, Cumulative Sum (CUSUM), Exponentially Weighted Moving Averages (EWMA), and Bayesian. In comparing control charts, Albin et al. (1997) "chose to use simulation. Essentially one program (with less than fifty lines of code) is used for all combinations of charts and run rules."

• An economic cost model can be made as general as desired. Barish and Hauser (1963) applied Monte Carlo simulation to test various combinations of parameters in an economic cost control chart design.

On the other hand, when it comes to the design and sensitivity analysis of control charts, the use of Monte Carlo simulation has been limited to "brute force" application. In other words, sensitivity analysis is conducted by changing the value of the parameter of interest and rerunning the simulation (e.g., Albin et al. 1997 for different process shift amounts). Optimization is carried out in a somewhat ad hoc trial-and-error manner, i.e., no formal optimization techniques are employed (e.g., Barish and Hauser 1963). In a recent paper by Lele (1997), derivative estimates play a crucial role in the control chart design problem, but the solution proposed there uses brute-force finite differences.

Thus, the primary goal of this paper can be stated as follows: to investigate efficiency improvement of Monte Carlo simulation in the design and analysis of control charts by introducing *gradient estimation techniques*.

Gradient estimation techniques generally require little additional overhead in the simulation and usually result in significant computational savings over the multiple runs needed to construct finite difference estimates. Moreover, the optimal economic design problem can be addressed using gradient-based algorithms (cf. Fu 1994). For problems of the type considered in Lele (1996, 1997), where derivative estimates are required, our numerical results indicate that the gradient estimators derived here are superior to finite differences.

In this paper, we consider the two most commonly used gradient estimation techniques: perturbation analysis (PA) and the likelihood ratio (LR) method (also known as the score function method). Monographs for the former are Ho and Cao (1991), Glasserman (1991), Cao (1994), and Fu and Hu (1997a), and for the latter is Rubinstein and Shapiro (1993). These methods have been applied predominantly to queueing and inventory models. This work represents the first application to statistical quality control.

Specifically, in this paper we derive sensitivity estimates for average run lengths with respect to different types of parameters: the control limits, the sampling frequency, and various process shift parameters. In §2, we introduce the problem setting with the requisite notation and briefly discuss the gradient estimation technique to be applied. In §3, we consider the standard in-control and out-of-control average run length performance measures and present sensitivity estimators with respect to the control limits for a fairly general control chart that includes the classical Shewhart and EWMA charts as special cases. The detailed derivations and technical proofs of unbiasedness have been placed in the Appendix and in Fu and Hu (1997b). In §4, we incorporate the dynamics of the process shift. In addition to considering the control limit parameters, we also derive estimators with respect to the sampling frequency and various process shift parameters. In §5, we concisely summarize, compare, and contrast the practical implementation of the various estimators. Section 6 contains simulation results comparing the numerical properties of the estimators with each other and with "brute force" finite differences. Section 7 concludes with a summary and a discussion of avenues for further research.

2. Problem Setting

We consider the standard control chart setting involving a single measurable process variable with two distinct states called "in control" and "out of control" (cf. Montgomery 1996). Samples of the process are taken at regularly spaced intervals and a test statistic generated (possibly based on past samples, as well). The test statistic is compared with control limits (that may vary as a function of time, as well) to declare the process in control or out of control. We begin by defining the following notation:

h = sampling interval, i.e., samples are taken every h time units;

n = sample size;

F_0 = sampling process c.d.f. (with p.d.f. f_0) when in control;

F_1 = sampling process c.d.f. (with p.d.f. f_1) when out of control;

μ_0 = in-control process parameter (usually the mean);

μ_1 = out-of-control process parameter;

X_i = output from the ith sample, i.i.d. F_0 or F_1 (X indicates "generic" version);

Y_i = test statistic after ith sample;

LCL_i = lower control limit for the ith test statistic;

UCL_i = upper control limit for the ith test statistic.

In words, samples of size n are taken every h units of time to generate $\{X_i\}$, from which the test statistic sequence $\{Y_i\}$ is derived. An out-of-control signal is declared if the test statistic Y_i falls outside of the interval defined by the lower and upper control limits $[\mathrm{LCL}_i, \mathrm{UCL}_i]$. The underlying process has an in-control c.d.f. F_0 and an out-of-control c.d.f. F_1. The sampling distributions F_0 and F_1 can be quite general, with standard distributions assumed (invoking the central limit theorem) being the normal distribution or the chi-squared distribution.

In this paper, we will assume that the test statistic generated by the control chart at the ith sampling has the following general form:

$$Y_i = \psi(X_i, Y_{i-1}), \quad i > 1, \tag{1}$$

where ψ is a function independent of other system parameters, and $Y_1 = X_1$ is specified as the initial condition. For notational convenience in the analysis in the next section, we will write

$$Y_i = \psi_i(X_i, Y_{i-1}), \quad i \geq 1, \tag{2}$$

by taking $\psi_1(x, y) = x$ and $\psi_i = \psi$ for $i > 1$.

Example 1. Shewhart $\bar{X}$-chart:

$$Y_i = X_i \quad \text{for all } i \geq 1.$$

Example 2. EWMA chart:

$$Y_i = \alpha X_i + (1 - \alpha) Y_{i-1}$$

$$\text{for all } i \geq 2,\ 0 < \alpha < 1,\ Y_1 = X_1.$$

As described above, an out-of-control signal is declared when the test statistic falls outside of the specified control limits. The corresponding sample number is defined as the run length, which is the performance measure of interest:

$$L = \min\{i : Y_i \notin [\mathrm{LCL}_i, \mathrm{UCL}_i]\}. \tag{3}$$

The expectation of this stopping time for $\{Y_i\}$ is what is commonly known as the average run length (ARL). We will consider three types of the ARL performance measure: the in-control and out-of-control ARLs, and the ARL under process shift dynamics. The in-control (out-of-control) ARL assumes that the process is in control (respectively, out of control) the entire time. Ideally, one wants long in-control run lengths and short out-of-control run lengths. The third type of ARL assumes that the process starts in control, but goes out of control at some later time. It is usually considered in the context of the economic design problem, where different costs are assigned to the in-control and out-of-control periods. The following quantities are introduced to characterize the process shift dynamics:

T = (r.v.) time to go from F_0 to F_1;

F = c.d.f. (with p.d.f. f) for T, parametrized by γ (e.g., the mean);

G_i = c.d.f. (with p.d.f. g_i) for $X_i \in \{F_0, F_1\}$.

Starting from a new in-control state, the process will go out of control after T units of time, where T is a random variable independent of $\{X_i\}$ with c.d.f. F, p.d.f. f, and parameter γ (e.g., the mean). The event $\{hL < T\}$ indicates a false alarm. The underlying dynamics that drive the process out of control may be quite general, since Monte Carlo simulation is to be employed, e.g., T need not be exponentially distributed, as is assumed in most analytical models. The sampling distribution sequence $\{G_i\}$ forms a discrete-time stochastic process which takes on the "value" F_0 or F_1, depending on whether or not the process is in control.

We introduce the following notation, which conditions on the value of the process shift time T and the initial test statistic Y_1:

$$J(t, y) = E[L \mid T = t, Y_1 = y], \tag{4}$$

$$J(t) = E[J(t, X_1)], \tag{5}$$

where the second definition is a slight abuse of notation used for the convenience of representing the "standard" initial starting condition of $Y_1 = X_1$ (expectation taken only with respect to second argument). Using this notation, the three performance measures of interest can be expressed as follows:

$$\mathrm{ARL}_T = E[E[L \mid T, Y_1 = X_1]] = E[J(T)],$$

$$\mathrm{ARL}_0 = E[L \mid T = \infty, Y_1 = X_1] = J(\infty),$$

$$\mathrm{ARL}_1 = E[L \mid T = 0, Y_1 = X_1] = J(0).$$

ARL_0 is the in-control (on-target) ARL, and ARL_1 is the out-of-control (off-target) ARL. ARL_T is the ARL for a process that starts in control, but shifts out of control at time T. ARL_T is useful for control chart design based on economic costs; when there is no confusion, the T argument (or subscript) will be dropped. Furthermore, except where specified otherwise, the initial condition $Y_1 = X_1$ is implicit throughout.

We will derive estimators for sensitivities of the average run length

$$\frac{d\mathrm{ARL}_0}{d\theta}, \quad \frac{d\mathrm{ARL}_1}{d\theta}, \quad \text{and} \quad \frac{d\mathrm{ARL}_T}{d\theta},$$

where θ will represent different types of parameters: the sampling frequency, control limit parameters, and various process parameters. The easiest to use gradient estimation techniques are the straightforward likelihood ratio method (LRM) and infinitesimal perturbation analysis (IPA). They are based on the following respective interchanges of expectation and differentiation:

$$\frac{dE[L(T, X_1, X_2, \ldots; \theta)]}{d\theta} = E\left[L \cdot \frac{d \ln f_L(T, X_1, X_2, \ldots; \theta)}{d\theta}\right], \tag{6}$$

$$\frac{dE[L(T, X_1, X_2, \ldots; \theta)]}{d\theta} = E\left[\frac{dL(T, X_1, X_2, \ldots; \theta)}{d\theta}\right], \tag{7}$$

where f_L represents the joint p.d.f. for the input r.v.s that determine L. The straightforward LRM in (6) is applicable only to distributional parameters, although the so-called "push out" method can be used to extend the approach to some structural parameters (cf. Rubinstein and Shapiro 1993). However, in this paper, we will consider only the straightforward LRM.

Unfortunately, IPA is not applicable to our problem setting, because L is a discrete random variable, taking on integer values. As a consequence, L as a function of the parameter will be piecewise constant with jumps, and thus the resulting IPA estimator $dL/d\theta$ in (7) will be 0. We will apply an extension of IPA introduced by Gong and Ho (1987) that employs conditional Monte Carlo, known as smoothed perturbation analysis (SPA). The particular approach taken here is based on the general framework of Fu and Hu (1992), as presented in Fu and Hu (1997a). The general form of the estimator contains two parts: the IPA estimator, zero in this case; and conditional contributions representing the product of a probability jump rate and the resulting jump in the performance measure. In the following two sections, we present the various estimators, leaving the detailed derivations and unbiasedness proofs to the Appendix.

3. In-Control and Out-of-Control ARLs

For the in-control and out-of-control ARLs, we consider derivative estimates with respect to the control limits. Throughout this section, we treat the two cases simultaneously by defining the two constants, $\tau_0 = \infty$, $\tau_1 = 0$, to correspond to the in-control and out-of-control cases, respectively. We can then use $T = \tau_j$, $j = 0, 1$ to refer to the two cases concurrently, where $\{X_i\}$ would have corresponding p.d.f. f_j and c.d.f. F_j. For example, under this notation, we have $\text{ARL}_j = J(\tau_j)$.

3.1. Control Limits

We consider the constant control limits case first:

$$l = \text{LCL}_i, \quad u = \text{UCL}_i.$$

If θ is a parameter in both LCL and UCL, as is typically the case, then the chain rule can be applied to obtain the sensitivities

$$\frac{dE[L]}{d\theta} = \frac{dE[L]}{du}\frac{du}{d\theta} + \frac{dE[L]}{dl}\frac{dl}{d\theta}.$$

Since the control limits are not distributional parameters, the straightforward LRM cannot be applied, although it is possible that the "push out" method could be applicable (cf. Rubinstein and Shapiro 1993). For example, in the Shewhart chart case, the control limits are usually of the form $\mu \pm k\sigma/\sqrt{n}$, so that derivatives with respect to the control limits are related to other derivatives.

Right-hand and left-hand derivative PA estimators can be derived by considering $\Delta u > 0$ and $\Delta u < 0$, respectively. In the former case, the analysis centers on the possibility that an increase in the upper control limit could cause the run length to be extended beyond the current stopping time (see Figure 1), whereas in the latter case, a decrease in the upper control limit could cause the run length to be shortened to any of the earlier samples (see Figure 2). In the Appendix, we provide the detailed derivation and establish unbiasedness for the following PA estimators:

$$\left(\frac{d\text{ARL}_j}{du}\right)_{\text{PA},R} = \frac{f_j(\psi_L^{-1}(u, Y_{L-1}))}{1 - F_j(\psi_L^{-1}(u, Y_{L-1}))}\frac{d\psi_L^{-1}(u, Y_{L-1})}{du} \times J(\tau_j, \psi(X, u))\mathbf{1}\{Y_L > u\}, \tag{8}$$

Figure 1 Extension of Run Length Caused by Positive Perturbation in Upper Control Limit

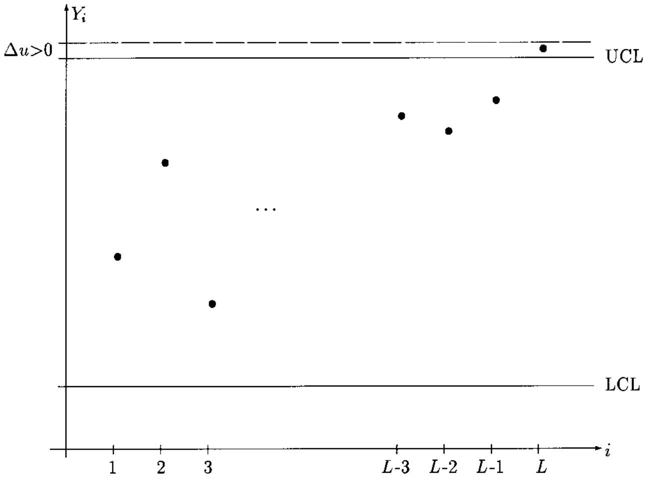
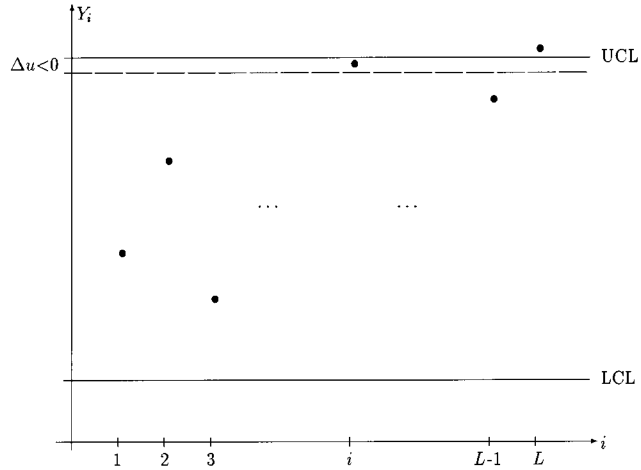


Figure 2 Shortening of Run Length Caused by Negative Perturbation in Upper Control Limit



$$\left(\frac{d\text{ARL}_j}{du}\right)_{\text{PA},L} = \sum_{i<L} \frac{f_j(\psi_i^{-1}(u, Y_{i-1}))}{F_j(\psi_i^{-1}(u, Y_{i-1})) - F_j(\psi_i^{-1}(l, Y_{i-1}))} \times \frac{d\psi_i^{-1}(u, Y_{i-1})}{du} J(\tau_j, \psi(X, u)), \tag{9}$$

where $X \sim F_j$ and $\psi_i^{-1}(\cdot, \cdot)$ denotes the inverse with respect to the first argument, so that

$$X_i = \psi^{-1}(Y_i, Y_{i-1}), \quad i > 1, \quad \psi_1^{-1}(w, y) = w.$$

For example, we have for the EWMA control chart:

$$\psi_i^{-1}(w, y) = \begin{cases} w & i = 1, \\ (w - (1 - \alpha)y)/\alpha & i > 1, \end{cases} \tag{10}$$

$$\frac{d\psi_i^{-1}(w, y)}{dw} = \alpha^{-\mathbf{1}\{i>1\}} = \begin{cases} 1 & i = 1, \\ 1/\alpha & i > 1. \end{cases} \tag{11}$$

In (8), the first two terms represent the probability rate that the run length at L will be extended, with the J term giving the expected length of the extension and the indicator term imposing the condition that this extension can only take place if the out-of-control signal was a violation of the upper control limit. In (9), the first two terms represent the probability rate that the run length at L will be shortened, with the J term again giving the expected amount shortened at sample i, and the summation over $i < L$ indicating that the shortening can take place at any of the earlier samples. In terms of practical implementation, the estimators may require an additional simulation to estimate the term $J(\tau_j, \psi(X, u))$.

In a similar fashion, the following unbiased estimators can be derived:

$$\left(\frac{d\mathrm{ARL}_j}{dl}\right)_{\mathrm{PA},R} = -\sum_{i<L} \frac{f_j(\psi_i^{-1}(l, Y_{i-1}))}{F_j(\psi_i^{-1}(u, Y_{i-1})) - F_j(\psi_i^{-1}(l, Y_{i-1}))} \times \frac{d\psi_i^{-1}(l, Y_{i-1})}{dl} J(\tau_j, \psi(X, l)), \tag{12}$$

$$\left(\frac{d\mathrm{ARL}_j}{dl}\right)_{\mathrm{PA},L} = -\frac{f_j(\psi_L^{-1}(l, Y_{L-1}))}{F_j(\psi_L^{-1}(l, Y_{L-1}))} \frac{d\psi_L^{-1}(l, Y_{L-1})}{du} \times J(\tau_j, \psi(X, l))\mathbf{1}\{Y_L < l\}. \tag{13}$$

EXAMPLE 1. Shewhart chart: $J(\tau_j, \psi(X, u)) = J(\tau_j, \psi(X, l)) = E[L] = E[(L - i) \mid L > i]$

$$\left(\frac{d\mathrm{ARL}_j}{du}\right)_{\mathrm{PA},R} = \frac{f_j(u)}{1 - F_j(u)} \mathbf{1}\{X_L > u\} \cdot L,$$

$$\left(\frac{d\mathrm{ARL}_j}{du}\right)_{\mathrm{PA},L} = \sum_{i<L} \frac{f_j(u)}{F_j(u) - F_j(l)} (L - i) = \frac{f_j(u)}{F_j(u) - F_j(l)} \frac{L(L-1)}{2},$$

$$\left(\frac{d\mathrm{ARL}_j}{dl}\right)_{\mathrm{PA},R} = -\sum_{i<L} \frac{f_j(l)}{F_j(u) - F_j(l)} (L - i) = -\frac{f_j(l)}{F_j(u) - F_j(l)} \frac{L(L-1)}{2},$$

$$\left(\frac{d\mathrm{ARL}_j}{dl}\right)_{\mathrm{PA},L} = -\frac{f_j(l)}{F_j(l)} \mathbf{1}\{X_L < l\} \cdot L.$$

In this case, no additional simulation is required for any of the four estimators.

EXAMPLE 2. EWMA chart:

$$\left(\frac{d\mathrm{ARL}_j}{du}\right)_{\mathrm{PA},R} = \frac{f_j(\psi_L^{-1}(u, Y_{L-1}))}{1 - F_j(\psi_L^{-1}(u, Y_{L-1}))} \alpha^{-\mathbf{1}\{L>1\}} \times J(\tau_j, \alpha X + (1 - \alpha)u)\mathbf{1}\{Y_L > u\},$$

$$\left(\frac{d\mathrm{ARL}_j}{du}\right)_{\mathrm{PA},L} = \sum_{i<L} \frac{f_j(\psi_i^{-1}(u, Y_{i-1}))}{F_j(\psi_i^{-1}(u, Y_{i-1})) - F_j(\psi_i^{-1}(l, Y_{i-1}))} \alpha^{-\mathbf{1}\{i>1\}} \times J(\tau_j, \alpha X + (1 - \alpha)u),$$

$$\left(\frac{d\mathrm{ARL}_j}{dl}\right)_{\mathrm{PA},R} = -\sum_{i<L} \frac{f_j(\psi_i^{-1}(l, Y_{i-1}))}{F_j(\psi_i^{-1}(u, Y_{i-1})) - F_j(\psi_i^{-1}(l, Y_{i-1}))} \alpha^{-\mathbf{1}\{i>1\}} \times J(\tau_j, \alpha X + (1 - \alpha)l),$$

$$\left(\frac{d\mathrm{ARL}_j}{dl}\right)_{\mathrm{PA},L} = -\frac{f_j(\psi_L^{-1}(l, Y_{L-1}))}{F_j(\psi_L^{-1}(l, Y_{L-1}))} \alpha^{-\mathbf{1}\{L>1\}} \times J(\tau_j, \alpha X + (1 - \alpha)l)\mathbf{1}\{Y_L < l\},$$

where ψ_i^{-1} is given by (10). In this case, additional simulation will be required for all of the estimators. The additional simulation is used to estimate ARL_j for the cases where the initial condition is given by $Y_1 = \alpha X_1 + (1 - \alpha)u$ and $Y_1 = \alpha X_1 + (1 - \alpha)l$, instead of the usual $Y_1 = X_1$.

Next we consider the general case where the control limits are also indexed by the sample number, i.e., we have a sequence $\{[l_i, u_i]\}$. We can use the chain rule to find sensitivities as follows:

$$\frac{\partial E[L]}{\partial \theta} = \sum_i \frac{\partial E[L]}{\partial u_i} \frac{\partial u_i}{\partial \theta} + \sum_i \frac{\partial E[L]}{\partial l_i} \frac{\partial l_i}{\partial \theta},$$

so for example this includes the constant control limit chart as a special case.

For the right-hand (left-hand) derivative w.r.t. u_i (l_i), we can only have a change if $L = i$, so there is no longer a summation. Otherwise, similar to before, our estimators are the following:

$$\left(\frac{d\mathrm{ARL}_j}{du_i}\right)_{\mathrm{PA},R} = \frac{f_j(\psi_i^{-1}(u_i, Y_{i-1}))}{1 - F_j(\psi_i^{-1}(u_i, Y_{i-1}))} \frac{d\psi_i^{-1}(u_i, Y_{i-1})}{du_i} \times J(\tau_j, \psi(X, u_i))\mathbf{1}\{Y_L > u_L\}\mathbf{1}\{L = i\},$$

$$\left(\frac{d\mathrm{ARL}_j}{du_i}\right)_{\mathrm{PA},L} = \frac{f_j(\psi_i^{-1}(u_i, Y_{i-1}))}{F_j(\psi_i^{-1}(u_i, Y_{i-1})) - F_j(\psi_i^{-1}(l_i, Y_{i-1}))} \times \frac{d\psi_i^{-1}(u_i, Y_{i-1})}{du_i} J(\tau_j, \psi(X, u_i))\mathbf{1}\{L > i\},$$

$$\left(\frac{d\mathrm{ARL}_j}{dl_i}\right)_{\mathrm{PA},R} = -\frac{f_j(\psi_i^{-1}(l_i, Y_{i-1}))}{F_j(\psi_i^{-1}(u_i, Y_{i-1})) - F_j(\psi_i^{-1}(l_i, Y_{i-1}))} \times \frac{d\psi_i^{-1}(l_i, Y_{i-1})}{dl_i} J(\tau_j, \psi(X, l_i))\mathbf{1}\{L > i\},$$

$$\left(\frac{d\mathrm{ARL}_j}{dl_i}\right)_{\mathrm{PA},L} = -\frac{f_j(\psi_i^{-1}(l_i, Y_{i-1}))}{F_j(\psi_i^{-1}(l_i, Y_{i-1}))} \frac{d\psi_i^{-1}(l_i, Y_{i-1})}{dl_i} \times J(\tau_j, \psi(X, l_i))\mathbf{1}\{Y_L < l_L\}\mathbf{1}\{L = i\}.$$

3.2. Process Distribution

Next, we consider derivatives with respect to parameters μ_j in the process sampling distributions, F_j ($j = 0, 1$). Since μ_j is a distributional parameter, straightforward LR can be applied via (6) to give the following estimator:

$$\left(\frac{d\mathrm{ARL}_j}{d\mu_j}\right)_{\mathrm{LR}} = L \cdot \sum_{i=1}^{L} \frac{d \ln f_j(X_i)}{d\mu_j}. \tag{14}$$

Note that the form of this estimator is independent of the form of the control chart, i.e., the only dependence on ψ is through L. This makes the estimator extremely easy to implement. In addition, unbiasedness will hold under standard conditions on the sampling distribution, e.g., all common continuous distributions such as the normal, lognormal, and gamma distributions will satisfy the conditions. As an example, for the usual normal distribution assumption on $\{X_i\}$, with mean μ_j and standard deviation σ, we have the following:

$$\left(\frac{d\mathrm{ARL}_j}{d\mu_j}\right)_{\mathrm{LR}} = L \cdot \sum_{i=1}^{L} \frac{X_i - \mu_j}{\sigma^2}, \tag{15}$$

$$\left(\frac{d\mathrm{ARL}_j}{d\sigma}\right)_{\mathrm{LR}} = L \cdot \sum_{i=1}^{L} \left(\frac{(X_i - \mu_j)^2}{\sigma^3} - \frac{1}{\sigma}\right). \tag{16}$$

For the PA estimator, the effect of changing a parameter in F_j is to cause a shift in $\{X_i\}$, which is similar to shifting both of the control limits *simultaneously*. First, we recall the standard PA form for the derivative of a random variable with respect to a parameter θ in its known distribution:

$$\frac{dX}{d\theta} = -\frac{dF_j(X; \theta)/d\theta}{f_j(X; \theta)}. \tag{17}$$

Common cases are $dX/d\theta = 1$ when θ is a location parameter (as in the mean of a normal distribution or a uniform distribution); and $dX/d\theta = X/\theta$ when θ is a scale parameter (as in the mean of an exponential distribution).

Since the control chart has dependence on previous data in the general form of the chart given by (1), the relationship between Y_i and X_i is used to "propagate" perturbations. Differentiating, we have

$$\frac{dY_i}{d\theta} = \frac{d\psi}{dX_i}\frac{dX_i}{d\theta} + \frac{d\psi}{dY_{i-1}}\frac{dY_{i-1}}{d\theta}, \quad i > 1, \quad \frac{dY_1}{d\theta} = \frac{dX_1}{d\theta}. \tag{18}$$

In particular, for the EWMA control chart, we have

$$\frac{dY_i}{d\theta} = \alpha\frac{dX_i}{d\theta} + (1-\alpha)\frac{dY_{i-1}}{d\theta}, \quad i > 1, \quad \frac{dY_1}{d\theta} = \frac{dX_1}{d\theta}. \tag{19}$$

The general form of the estimator is then derived by considering the possible changes in X_i caused by a change in μ_j. If they are positively correlated, then the estimator is given by the negative of the sum of the two left-hand derivative estimators w.r.t. the control limits, whereas if they are negatively correlated, then the estimator is given by the sum of the two right-hand derivative estimators w.r.t. the control limits. Thus, the final estimator can be written as follows:

$$\begin{aligned}\left(\frac{d\mathrm{ARL}_j}{d\mu_j}\right)_{\mathrm{PA}} &= \left(\frac{dY_L}{d\mu_j}\right)^{-}\mathbf{1}\{Y_L > u\}\frac{f_j(\psi_L^{-1}(u, Y_{L-1}))}{1 - F_j(\psi_L^{-1}(u, Y_{L-1}))} \\ &\times \frac{d\psi_L^{-1}(u, Y_{L-1})}{du} J(\tau_j, \psi(X, u)) \\ &- \sum_{i<L}\frac{f_j(\psi_i^{-1}(l, Y_{i-1}))}{F_j(\psi_i^{-1}(u, Y_{i-1})) - F_j(\psi_i^{-1}(l, Y_{i-1}))} \\ &\times \frac{d\psi_i^{-1}(l, Y_{i-1})}{dl}\left(\frac{dY_i}{d\mu_j}\right)^{-} J(\tau_j, \psi(X, l)) \\ &+ \left(\frac{dY_L}{d\mu_j}\right)^{+}\mathbf{1}\{Y_L < l\}\frac{f_j(\psi_L^{-1}(l, Y_{L-1}))}{F_j(\psi_L^{-1}(l, Y_{L-1}))} \\ &\times \frac{d\psi_L^{-1}(l, Y_{L-1})}{dl} J(\tau_j, \psi(X, l)) \\ &- \sum_{i<L}\frac{f_j(\psi_i^{-1}(u, Y_{i-1}))}{F_j(\psi_i^{-1}(u, Y_{i-1})) - F_j(\psi_i^{-1}(l, Y_{i-1}))} \\ &\times \frac{d\psi_i^{-1}(u, Y_{i-1})}{du}\left(\frac{dY_i}{d\mu_j}\right)^{+} J(\tau_j, \psi(X, u)),\end{aligned} \tag{20}$$

where $x^+ = \max(x, 0)$ and $x^- = \max(-x, 0)$. Note that at most two of the four terms will be nonzero for a given case, since the sign of $d/d\mu_j$ is either positive or negative.

Example 1. Shewhart chart:

$$\begin{aligned}\left(\frac{d\mathrm{ARL}_j}{d\mu_j}\right)_{\mathrm{PA}} &= \left(\frac{f_j(u)}{1 - F_j(u)}\mathbf{1}\{X_L > u\}\cdot L \right.\\ &\left.- \frac{f_j(l)}{F_j(u) - F_j(l)}\frac{L(L-1)}{2}\right)\mathbf{1}\left\{\frac{dX}{d\mu_j} < 0\right\} \\ &+ \left(\frac{f_j(l)}{F_j(l)}\mathbf{1}\{X_L < l\}\cdot L \right.\\ &\left.- \frac{f_j(u)}{F_j(u) - F_j(l)}\frac{L(L-1)}{2}\right)\mathbf{1}\left\{\frac{dX}{d\mu_j} > 0\right\}.\end{aligned}$$

Example 2. EWMA chart:

$$\begin{aligned}\left(\frac{d\mathrm{ARL}_j}{d\mu_j}\right)_{\mathrm{PA}} &= \left(\frac{dY_L}{d\mu_j}\right)^{-}\mathbf{1}\{Y_L > u\}\frac{f_j(\psi_L^{-1}(u, Y_{L-1}))}{1 - F_j(\psi_L^{-1}(u, Y_{L-1}))}\alpha^{-\mathbf{1}\{L>1\}} \\ &\times J(\tau_j, \psi(X, u)) \\ &- \sum_{i<L}\frac{f_j(\psi_i^{-1}(l, Y_{i-1}))}{F_j(\psi_i^{-1}(u, Y_{i-1})) - F_j(\psi_i^{-1}(l, Y_{i-1}))}\alpha^{-\mathbf{1}\{i>1\}} \\ &\times \left(\frac{dY_i}{d\mu_j}\right)^{-} J(\tau_j, \psi(X, l)) + \left(\frac{dY_L}{d\mu_j}\right)^{+}\mathbf{1}\{Y_L < l\} \\ &\times \frac{f_j(\psi_L^{-1}(l, Y_{L-1}))}{F_j(\psi_L^{-1}(l, Y_{L-1}))}\alpha^{-\mathbf{1}\{L>1\}} J(\tau_j, \psi(X, l)) \\ &- \sum_{i<L}\frac{f_j(\psi_i^{-1}(u, Y_{i-1}))}{F_j(\psi_i^{-1}(u, Y_{i-1})) - F_j(\psi_i^{-1}(l, Y_{i-1}))}\alpha^{-\mathbf{1}\{i>1\}} \\ &\times \left(\frac{dY_i}{d\mu_j}\right)^{+} J(\tau_j, \psi(X, u)).\end{aligned}$$

Next, we note that we can also consider derivatives with respect to parameters in the relationship ψ itself, as well. For example, in the case of the EWMA control chart, one might be interested in sensitivities with respect to the smoothing constant α. Again, we simply differentiate the relationship to obtain a propagation rule:

$$\frac{dY_i}{d\alpha} = X_i - Y_{i-1}, \tag{21}$$

which is used in the estimator (20) above.

4. Process Shift ARL

Now we consider the case where the process begins in control and goes out of control after a random time T. Thus, the distribution of each X_i depends on T, so that L also depends on T implicitly. The dependence of $\{X_i\}$ on T is not on the actual value of T, but just on which sampling interval it occurs. Therefore, we define the index random variable for the last in-control sample taken (refer to Figure 3)

$$\eta = \max\{j : hj < T\}$$

i.e.,

$$X_i \sim \begin{cases} F_0 & \text{for } i \le \eta, \\ F_1 & \text{for } i > \eta. \end{cases}$$

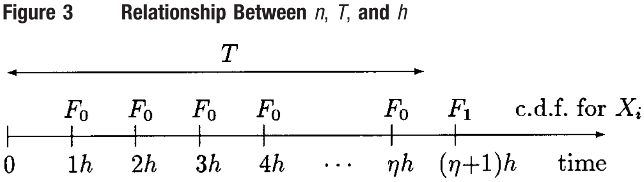
4.1. Control Limits

These estimators resemble very closely the corresponding in-control and out-of-control estimators, with the following exceptions:

- The initial condition on the additional run length involves the residual time of T at the time of the out-of-control signal;
- The probability rate term is replaced with a random distribution G_i, equal to either F_1 or F_0 depending on whether the out-of-control signal is a true or false alarm, respectively.

Otherwise proceeding in the same manner, we obtain the following estimators:

Figure 3 **Relationship Between** *n*, *T*, **and** *h*



$$\left(\frac{d\text{ARL}_j}{du}\right)_{\text{PA},R} = \frac{g_L(\psi_L^{-1}(u, Y_{L-1}))}{1 - G_L(\psi_L^{-1}(u, Y_{L-1}))} \frac{d\psi_L^{-1}(u, Y_{L-1})}{du} \times J(\tau_{\text{res}}, \psi(X, u))\mathbf{1}\{Y_L > u\}, \tag{22}$$

$$\left(\frac{d\text{ARL}_j}{du}\right)_{\text{PA},L} = \sum_{i<L} \frac{g_i(\psi_i^{-1}(u, Y_{i-1}))}{G_i(\psi_i^{-1}(u, Y_{i-1})) - G_i(\psi_i^{-1}(l, Y_{i-1}))} \times \frac{d\psi_i^{-1}(u, Y_{i-1})}{du} J(\tau_{\text{res}}, \psi(X, u)), \tag{23}$$

$$\left(\frac{d\text{ARL}_j}{dl}\right)_{\text{PA},R} = -\sum_{i<L} \frac{g_i(\psi_i^{-1}(l, Y_{i-1}))}{G_i(\psi_i^{-1}(u, Y_{i-1})) - G_i(\psi_i^{-1}(l, Y_{i-1}))} \times \frac{d\psi_i^{-1}(l, Y_{i-1})}{dl} J(\tau_{\text{res}}, \psi(X, l)), \tag{24}$$

$$\left(\frac{d\text{ARL}_j}{dl}\right)_{\text{PA},L} = -\frac{g_L(\psi_L^{-1}(l, Y_{L-1}))}{G_L(\psi_L^{-1}(l, Y_{L-1}))} \frac{d\psi_L^{-1}(l, Y_{L-1})}{dl} \times J(\tau_{\text{res}}, \psi(X, l))\mathbf{1}\{Y_L < l\}, \tag{25}$$

where

$$\tau_{\text{res}} = (T - hL)^+ \tag{26}$$

is the residual time until the system actually goes out of control from the epoch at which an out-of-control signal is declared. Thus, if $T \le hL$, the system is already out of control, and hence the residual time is zero.

Example 1. Shewhart chart:

$$\left(\frac{d\text{ARL}_j}{du}\right)_{\text{PA},R} = \frac{g_L(u)}{1 - G_L(u)} J(\tau_{\text{res}})\mathbf{1}\{X_L > u\}, \tag{27}$$

$$\left(\frac{d\text{ARL}_j}{du}\right)_{\text{PA},L} = \sum_{i<L} \frac{g_i(u)}{G_i(u) - G_i(l)} J(\tau_{\text{res}}), \tag{28}$$

$$\left(\frac{d\text{ARL}_j}{dl}\right)_{\text{PA},R} = -\sum_{i<L} \frac{g_i(l)}{G_i(u) - G_i(l)} J(\tau_{\text{res}}), \tag{29}$$

$$\left(\frac{d\,\mathrm{ARL}_j}{dl}\right)_{\mathrm{PA},L} = -\frac{g_L(l)}{G_L(l)} J(\tau_{\mathrm{res}})\mathbf{1}\{X_L < l\}. \quad (30)$$

EXAMPLE 2. EWMA chart: ψ_i^{-1} given by (10):

$$\left(\frac{d\,\mathrm{ARL}_j}{du}\right)_{\mathrm{PA},R} = \frac{g_L(\psi_L^{-1}(u, Y_{L-1}))}{1 - G_L(\psi_L^{-1}(u, Y_{L-1}))} \alpha^{-\mathbf{1}\{L>1\}} \times J(\tau_{\mathrm{res}}, \psi(X, u))\mathbf{1}\{Y_L > u\}, \quad (31)$$

$$\left(\frac{d\,\mathrm{ARL}_j}{du}\right)_{\mathrm{PA},L} = \sum_{i<L} \frac{g_i(\psi_i^{-1}(u, Y_{i-1}))}{G_i(\psi_i^{-1}(u, Y_{i-1})) - G_i(\psi_i^{-1}(l, Y_{i-1}))} \alpha^{-\mathbf{1}\{i>1\}} \times J(\tau_{\mathrm{res}}, \psi(X, u)), \quad (32)$$

$$\left(\frac{d\,\mathrm{ARL}_j}{dl}\right)_{\mathrm{PA},R} = -\sum_{i<L} \frac{g_i(\psi_i^{-1}(l, Y_{i-1}))}{G_i(\psi_i^{-1}(u, Y_{i-1})) - G_i(\psi_i^{-1}(l, Y_{i-1}))} \alpha^{-\mathbf{1}\{i>1\}} \times J(\tau_{\mathrm{res}}, \psi(X, l)), \quad (33)$$

$$\left(\frac{d\,\mathrm{ARL}_j}{dl}\right)_{\mathrm{PA},L} = -\frac{g_L(\psi_L^{-1}(l, Y_{L-1}))}{G_L(\psi_L^{-1}(l, Y_{L-1}))} \alpha^{-\mathbf{1}\{L>1\}} \times J(\tau_{\mathrm{res}}, \psi(X, l))\mathbf{1}\{Y_L < l\}. \quad (34)$$

For the general case where the control limits are also indexed by the sample number, $\{[l_i, u_i]\}$:

$$\left(\frac{d\,\mathrm{ARL}_j}{du_i}\right)_{\mathrm{PA},R} = \frac{g_i(\psi_i^{-1}(u_i, Y_{i-1}))}{1 - G_i(\psi_i^{-1}(u_i, Y_{i-1}))} \frac{d\psi_i^{-1}(u_i, Y_{i-1})}{du_i} \times J(\tau_{\mathrm{res}}, \psi(X, u_i))\mathbf{1}\{Y_L > u_L\}\mathbf{1}\{L = i\},$$

$$\left(\frac{d\,\mathrm{ARL}_j}{du_i}\right)_{\mathrm{PA},L} = \frac{g_i(\psi_i^{-1}(u_i, Y_{i-1}))}{G_i(\psi_i^{-1}(u_i, Y_{i-1})) - G_i(\psi_i^{-1}(l_i, Y_{i-1}))} \times \frac{d\psi_i^{-1}(u_i, Y_{i-1})}{du_i} J(\tau_{\mathrm{res}}, \psi(X, u_i))\mathbf{1}\{L > i\},$$

$$\left(\frac{d\,\mathrm{ARL}_j}{dl_i}\right)_{\mathrm{PA},R} = -\frac{g_i(\psi_i^{-1}(l_i, Y_{i-1}))}{G_i(\psi_i^{-1}(u_i, Y_{i-1})) - G_i(\psi_i^{-1}(l_i, Y_{i-1}))} \times \frac{d\psi_i^{-1}(l_i, Y_{i-1})}{dl_i} J(\tau_{\mathrm{res}}, \psi(X, l_i))\mathbf{1}\{L > i\},$$

$$\left(\frac{d\,\mathrm{ARL}_j}{dl_i}\right)_{\mathrm{PA},L} = -\frac{g_i(\psi_i^{-1}(l_i, Y_{i-1}))}{G_i(\psi_i^{-1}(l_i, Y_{i-1}))} \frac{d\psi_i^{-1}(l_i, Y_{i-1})}{dl_i} \times J(\tau_{\mathrm{res}}, \psi(X, l_i))\mathbf{1}\{Y_L < l_L\}\mathbf{1}\{L = i\}.$$

4.2. In-Control and Out-of-Control Process Parameters

Similar to the case of the in-control and out-of-control ARLs, we can easily derive estimators for parameters in the in-control and out-of-control sampling distributions. Basically, the only difference in the estimators is that there are nonzero contributions only when the sampling distribution being used for X_i is the one in which the parameter μ_j enters:

$$\frac{dX_i}{d\mu_0} = \frac{d \ln g_i(\cdot)}{d\mu_0} = 0 \quad \text{for } i \geq T/h \text{ (where } X_i \sim F_1), \quad (35)$$

and

$$\frac{dX_i}{d\mu_1} = \frac{d \ln g_i(\cdot)}{d\mu_1} = 0 \quad \text{for } i < T/h \text{ (where } X_i \sim F_0). \quad (36)$$

Recall that the random variable T is assumed independent of μ_0 and μ_1.

Again, straightforward LR can be applied to give the following estimator:

$$\left(\frac{d\,\mathrm{ARL}_T}{d\mu_j}\right)_{\mathrm{LR}} = L \cdot \sum_{i=1}^{L} \frac{d \ln g_i(X_i)}{d\mu_j}. \quad (37)$$

Similarly, we can derive the following PA estimator:

$$\left(\frac{d\mathrm{ARL}_T}{d\mu_j}\right)_{\mathrm{PA}} = \left(\frac{dY_L}{d\mu_j}\right)^{-} \mathbf{1}\{Y_L > u\} \frac{g_L(\psi_L^{-1}(u, Y_{L-1}))}{1 - G_L(\psi_L^{-1}(u, Y_{L-1}))} \times \frac{d\psi_L^{-1}(u, Y_{L-1})}{du} J(\tau_{\mathrm{res}}, \psi(X, u))$$
$$- \sum_{i<L} \frac{g_i(\psi_i^{-1}(l, Y_{i-1}))}{G_i(\psi_i^{-1}(u, Y_{i-1})) - G_i(\psi_i^{-1}(l, Y_{i-1}))} \times \frac{d\psi_i^{-1}(l, Y_{i-1})}{du} \left(\frac{dY_i}{d\mu_j}\right)^{-} J(\tau_{\mathrm{res}}, \psi(X, l))$$
$$+ \left(\frac{dY_L}{d\mu_j}\right)^{+} \mathbf{1}\{Y_L < l\} \frac{g_L(\psi_L^{-1}(l, Y_{L-1}))}{G_L(\psi_L^{-1}(l, Y_{L-1}))} \times \frac{d\psi_L^{-1}(l, Y_{L-1})}{du} J(\tau_{\mathrm{res}}, \psi(X, l))$$
$$- \sum_{i<L} \frac{g_i(\psi_i^{-1}(u, Y_{i-1}))}{G_i(\psi_i^{-1}(u, Y_{i-1})) - G_i(\psi_i^{-1}(l, Y_{i-1}))} \times \frac{d\psi_i^{-1}(u, Y_{i-1})}{du} \left(\frac{dY_i}{d\mu_j}\right)^{+} J(\tau_{\mathrm{res}}, \psi(X, u)). \quad (38)$$

EXAMPLE 1. Shewhart chart: τ_{res} given by (26),

$$\left(\frac{d\mathrm{ARL}_T}{d\mu_j}\right)_{\mathrm{PA}} = \left(\frac{dX_L}{d\mu_j}\right)^{-} \mathbf{1}\{X_L > u\} \frac{g_L(u)}{1 - G_L(u)} J(\tau_{\mathrm{res}}) - \sum_{i<L} \frac{g_i(l)}{G_i(u) - G_i(l)} \left(\frac{dX_i}{d\mu_j}\right)^{-} J(\tau_{\mathrm{res}})$$
$$+ \left(\frac{dX_L}{d\mu_j}\right)^{+} \mathbf{1}\{X_L < l\} \frac{g_L(l)}{G_L(l)} J(\tau_{\mathrm{res}}) - \sum_{i<L} \frac{g_i(u)}{G_i(u) - G_i(l)} \left(\frac{dX_i}{d\mu_j}\right)^{+} J(\tau_{\mathrm{res}}). \quad (39)$$

Note that if T is exponentially distributed, then by the memoryless property, $J(\tau_{\mathrm{res}})$ can be replaced by L, and no extra simulation is required.

EXAMPLE 2. EWMA chart: τ_{res} given by (26),

$$\left(\frac{d\mathrm{ARL}_T}{d\mu_j}\right)_{\mathrm{PA}} = \left(\frac{dY_L}{d\mu_j}\right)^{-} \mathbf{1}\{Y_L > u\} \frac{g_L(\psi_L^{-1}(u, Y_{L-1}))}{1 - G_L(\psi_L^{-1}(u, Y_{L-1}))} \alpha^{-\mathbf{1}\{L>1\}} \times J(\tau_{\mathrm{res}}, \psi(X, u))$$
$$- \sum_{i<L} \frac{g_i(\psi_i^{-1}(l, Y_{i-1}))}{G_i(\psi_i^{-1}(u, Y_{i-1})) - G_i(\psi_i^{-1}(l, Y_{i-1}))} \alpha^{-\mathbf{1}\{i>1\}} \times \left(\frac{dY_i}{d\mu_j}\right)^{-} J(\tau_{\mathrm{res}}, \psi(X, l)) + \left(\frac{dY_L}{d\mu_j}\right)^{+} \mathbf{1}\{Y_L < l\}$$
$$\times \frac{g_L(\psi_L^{-1}(l, Y_{L-1}))}{G_L(\psi_L^{-1}(l, Y_{L-1}))} \alpha^{-\mathbf{1}\{L>1\}} J(\tau_{\mathrm{res}}, \psi(X, l))$$
$$- \sum_{i<L} \frac{g_i(\psi_i^{-1}(u, Y_{i-1}))}{G_i(\psi_i^{-1}(u, Y_{i-1})) - G_i(\psi_i^{-1}(l, Y_{i-1}))} \alpha^{-\mathbf{1}\{i>1\}} \times \left(\frac{dY_i}{d\mu_j}\right)^{+} J(\tau_{\mathrm{res}}, \psi(X, u)). \quad (40)$$

4.3. Sampling Frequency

We now consider the case where the parameter is h, the sampling interval. Since this is a structural parameter, the straightforward LRM is not applicable, so we consider only the PA estimator. The key characteristic in the analysis is to note that a change in the sampling interval may cause a change in the sample number in which the out-of-control signal is declared. Consider the right-hand estimator, $\Delta h > 0$. Writing $\eta(h)$ to denote the dependence of η on the sampling interval, we observe increasing the sampling interval may cause the interval in which the process actually goes out of control to decrease, specifically $\eta(h + \Delta h) = \eta(h) - 1$, which would imply that for $h + \Delta h$, we would have

$$X_i \sim \begin{cases} F_0 & \text{for } i \le \eta - 1, \\ F_1 & \text{for } i > \eta - 1, \end{cases}$$

so that the distribution of X_η would change. In other words, T has shifted from falling in the interval $[\eta h, (\eta + 1)h)$ to falling in the preceding interval (see

Figure 4). Based on this critical event change, one can derive the following estimator (see the Appendix):

$$\left(\frac{d\mathrm{ARL}_T}{dh}\right)_{\mathrm{PA},R} = \frac{\eta \cdot f(\eta h)}{F((\eta+1)h) - F(\eta h)}\,[J([\eta h]_-) - L],$$
$$[x]_- = (x - \epsilon),\ \epsilon \to 0^+. \quad (41)$$

Similarly, one can derive the following left-hand derivative estimator:

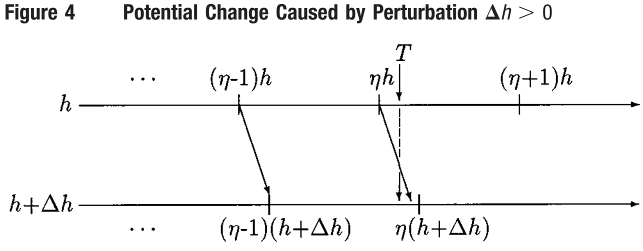
$$\left(\frac{d\mathrm{ARL}_T}{dh}\right)_{\mathrm{PA},L} = \frac{(\eta+1)\cdot f((\eta+1)h)}{F((\eta+1)h) - F(\eta h)} \times [L - J([(\eta+1)h]_+)],$$
$$[x]_+ = (x + \epsilon),\ \epsilon \to 0^+. \quad (42)$$

If T is a continuous random variable, it can be easily shown

$$E[J([\eta h]_-)] = E[J(T-h)],$$
$$E[J([(\eta+1)h]_+)] = E[J(T+h)].$$

In general, the estimators require one additional quantity to be estimated by simulation. However, in the case of the Shewhart chart, two particular coupling constructions (see Fu and Hu 1997b for details) can be used to derive efficient estimators for the expected values of $J([\eta h]_-) - L$ and $L - J([(\eta+1)h]_+)$, which in turn lead to the following estimators that require only the customary in-control and out-of-control ARLs:

Figure 4 Potential Change Caused by Perturbation $\Delta h > 0$



$$\left(\frac{d\mathrm{ARL}_T}{dh}\right)_{\mathrm{PA},R1} = \frac{\eta \cdot f(\eta h)}{F((\eta+1)h) - F(\eta h)}\,[(\eta - L)\mathbf{1}\{\eta < L, \times F_1^{-1}(F_0(X_\eta)) \notin [\mathrm{LCL},\mathrm{UCL}]\} + \mathrm{ARL}_1 \cdot \mathbf{1}\{\eta = L, \times F_1^{-1}(F_0(X_\eta)) \in [\mathrm{LCL},\mathrm{UCL}]\}],$$

$$\left(\frac{d\mathrm{ARL}_T}{dh}\right)_{\mathrm{PA},R2} = \frac{\eta \cdot f(\eta h)}{F((\eta+1)h) - F(\eta h)}\,[(-1)\mathbf{1}\{\eta < L\} + \mathrm{ARL}_1 \cdot \mathbf{1}\{\eta = L,\ X_{\eta+1} \in [\mathrm{LCL},\mathrm{UCL}]\}],$$

$$\left(\frac{d\mathrm{ARL}_T}{dh}\right)_{\mathrm{PA},L1} = -\frac{(\eta+1)\cdot f((\eta+1)h)}{F((\eta+1)h) - F(\eta h)}\,[((\eta+1) - L) \times \mathbf{1}\{\eta + 1 < L,\ F_0^{-1}(F_1(X_\eta)) \notin [\mathrm{LCL},\mathrm{UCL}]\} + \mathrm{ARL}_1 \cdot \mathbf{1}\{(\eta+1)h = L, \times F_0^{-1}(F_1(X_\eta)) \in [\mathrm{LCL},\mathrm{UCL}]\}],$$

$$\left(\frac{d\mathrm{ARL}_T}{dh}\right)_{\mathrm{PA},L2} = -\frac{(\eta+1)\cdot f((\eta+1)h)}{F((\eta+1)h) - F(\eta h)}\,[((\eta+1) - L) \times \mathbf{1}\{(\eta+1) < L,\ \tilde{X} \notin [\mathrm{LCL},\mathrm{UCL}]\} + \mathbf{1}\{\eta + 1 < L,\ \tilde{X} \in [\mathrm{LCL},\mathrm{UCL}]\} + \mathrm{ARL}_1 \cdot \mathbf{1}\{\eta + 1 = L,\ \tilde{X} \in [\mathrm{LCL},\mathrm{UCL}]\}].$$

4.4. Process "Drift" Parameters

Next, we consider sensitivities to parameters that enter the dynamics of the process going from in control to out of control. In particular, we consider the derivative with respect to γ, a parameter in the distribution of T.

Since γ is a distributional parameter, straightforward LR can be applied to give

$$\left(\frac{d\,\mathrm{ARL}_T}{d\gamma}\right)_{\mathrm{LR}} = L \cdot \frac{d \ln f(T)}{d\gamma}. \tag{43}$$

Note that in contrast to the previous LR estimators given by (14) and (37), this one just has a single multiplicative term and not a sum of a possibly large number of terms, the latter of which can cause an undesirable linear growth in the variance of the estimator. Thus, one would expect this estimator (43) to have relatively low variance.

For the PA estimator, a little thought reveals that the resulting estimator is similar to the previous one, since the change occurs only in the timing of the samplings. The only differences are the following: $\Delta T < 0$ corresponds to $\Delta h > 0$, $\Delta T > 0$ corresponds to $\Delta h < 0$, and the probability rate term is slightly different, as computed in the Appendix. Thus, we have the following estimators:

$$\left(\frac{d\,\mathrm{ARL}_T}{d\gamma}\right)_{\mathrm{PA},R} = \frac{-f(\eta h)}{F((\eta+1)h) - F(\eta h)} \left.\frac{dT}{d\gamma}\right|_{T=\eta h} \times [J([\eta h]_-) - L], \tag{44}$$

$$\left(\frac{d\,\mathrm{ARL}_T}{d\gamma}\right)_{\mathrm{PA},L} = \frac{-f((\eta+1)h)}{F((\eta+1)h) - F(\eta h)} \left.\frac{dT}{d\gamma}\right|_{T=(\eta+1)h} \times [L - J([(\eta+1)h]_+)]. \tag{45}$$

As in the previous section, for the Shewhart control chart an efficient estimator can be derived by using either of the two coupling constructions discussed in the Appendix for estimating the expected values of $J([\eta h]_-) - L$ and $L - J([\eta + 1)h]_-)$.

5. Implementation of the Estimators

In this section, we summarize the practical implementation of the various estimators. The LR estimators never require any additional simulation wherever applicable (the same is true for the IPA estimator, which is not applicable in this problem setting at all) and are always straightforward to implement. PA estimators based on SPA often require additional simulation to estimate quantities that are not available explicitly in the original simulation. In Fu and Hu (1992), it is noted that this could lead to an impractical amount of extra computation. However, for the problem considered here, the amount of extra burden is quite minimal, being on the order of one additional simulation or less. The quantities to be estimated in the PA estimators are just ARLs that can be easily estimated by simulations with different starting conditions on T and Y_1. Hence, all of the estimators are at least as efficient as finite differences in terms of amount of simulation required to generate an estimate. However, the precision of the estimators is just as important and will be tested for some examples in the next section.

6. Numerical Results

Numerical properties of the estimators were investigated via simulation examples for the Shewhart and EWMA charts. We purposely selected most of the examples to be cases where analytical results are available in order to easily judge the accuracy of the estimates. The applicability of the estimators, however, is far wider than the examples considered here. For sensitivities with respect to control limits, we compared the PA estimates with brute-force finite difference (FD) estimates, particularly with symmetric difference (SD) estimates using common random numbers. For sensitivities with respect to parameters in the sampling distributions and the process shift timing distribution, we compared all three of the PA, LR, and SD estimates.

The various cases are based on the following distributions and parameter settings:

- normal sampling distribution F_j with $\sigma/\sqrt{n} = 1$, $\mu_0 = 0$ and $\mu_1 = 0.1, 1.0, 3.0$, corresponding to small, moderate, and large shifts, respectively;
- 95% and 99.5% control limits corresponding to ± 1.96 and ± 2.81, respectively (and hence, in-control ARLs of 20 and 200, respectively, for the Shewhart chart);
- exponential distribution F with mean $\gamma = 20$ for the process shift time T.

Without loss of generality, we took the sampling frequency to be $h = 1$. As shown in the tables displaying the results, not every combination was simulated. For the EWMA chart, the control limits were set at ± 1.94 for μ_1

Table 1 Shewhart Chart Derivative Estimators: Mean (Standard Error)

		ARL		d/du				d/dl				$d/d\mu$			
μ	Control Limits	sim	ana	PA, L	PA, R	SD	ana	PA, L	PA, R	SD	ana	PA	LR	SD	ana
0.0	±1.96	19.9	20.0	23.3	23.4	23.4	23.4	−23.1	−23.3	−21.9	−23.4	−0.2	2.4	−1.5	0.0
		(0.2)		(0.5)	(0.4)	(0.7)		(0.4)	(0.5)	(0.7)		(0.5)	(2.0)	(0.8)	
0.0	±2.81	200.6	200.0	306	306	302	314	−317	−306	−300	−314	11.2	25.3	−6.5	0.0
		(2.0)		(6.8)	(5.2)	(7.7)		(5.4)	(6.8)	(7.8)		(5.9)	(64.9)	(9.2)	
0.1	±1.96	19.4	19.5	26.6	27.0	26.9	27.0	−18.0	−18.0	−18.2	−18.3	−8.7	−6.4	−9.0	−8.8
		(0.2)		(0.6)	(0.4)	(0.7)		(0.4)	(0.4)	(0.6)		(0.6)	(1.9)	(0.8)	
0.1	±2.81	191.9	193.4	373	373	377	379	−218	−212	−206	−216	−155	−124	−159	−163
		(1.9)		(8.6)	(5.4)	(8.4)		(4.6)	(4.9)	(6.4)		(8.1)	(61.2)	(9.0)	
1.0	±1.96	5.8	5.9	8.4	8.6	8.4	8.7	−0.18	−0.17	−0.15	−0.17	−8.3	−7.5	−8.2	−8.5
		(0.1)		(.18)	(.08)	(.21)		(.02)	(.001)	(.03)		(.18)	(.35)	(.21)	
1.0	±2.81	28.1	28.4	59.5	61.8	60.9	62.5	−0.28	−0.22	−0.25	−0.23	−59.3	−58.9	−60.5	−62.3
		(0.3)		(1.3)	(0.6)	(1.3)		(.08)	(.004)	(.08)		(1.2)	(3.9)	(1.3)	
3.0	±1.96	1.2	1.2	0.32	0.32	0.35	0.32	0.00	0.00	0.00	0.00	−0.32	−0.31	−0.35	−0.32
		(.00)		(.01)	(.001)	(.02)		(.000)	(.000)	(.000)		(.01)	(.01)	(.02)	
3.0	±2.81	1.8	1.74	1.20	1.19	1.14	1.18	0.00	0.00	0.00	0.00	−1.20	−1.20	−1.14	−1.18
		(.01)		(.03)	(.01)	(.04)		(.000)	(.000)	(.000)		(.03)	(.05)	(.04)	

= 0.25, 1.0, 3.0, yielding respective theoretical ARLs of 61.1, 10.6, 1.4, respectively, and an in-control ARL of 82.5 (Crowder 1987). In considering the process shift dynamics, only the Shewhart chart with wider control limits was considered. We report the results for derivatives with respect to the control limits, the process means, and the process shift time mean. Omitted is the derivative with respect to sampling frequency, since it yields basically an identical PA estimator to that for the process shift mean.

Each case was run for 10,000 replications, and Tables 1 through 3 report the results in the form of the sample mean with standard error in parentheses; "ana" indicates analytical results where available. For the PA estimators that require extra simulation, we performed only one additional run, so the computational burden would not exceed that of the SD estimator. Finally, the variance of the SD estimator of course decreases as the difference size Δ is increased, but at the cost of bias. Different sizes for Δ were tried; the results reported here are for $\Delta = 0.1$, which gave reasonable variance without introducing too much bias. (For example, biases of well over 10% of the mean were obtained for the control limit derivatives when $\Delta = 0.5$ was used.)

The simulation results indicate that the PA estimates are generally superior to SD estimates, but some estimators are more promising than others. Table 4 takes a closer look at the variance estimates in the most promising setting, the RH PA estimator with respect to the control limits. In order to obtain approximately valid confidence intervals using standard statistical estimation (i.e., via the chi-squared distribution), we used a batched estimator with batches of size 5, giving a total of 20,000 samples. For further comparison, the variances of SD estimators without common random numbers (CRN)—i.e., using independent streams—are also reported. The results indicate that the variance reduction of the PA estimator over SD is statistically significant and can be quite substantial, exceeding that in going from independent SD to CRN SD, and achieving two orders of magnitude reduction in one case (implying the need for over 100 times more replications for SD than for PA to obtain the same precision). Greater variance reduction is achieved for moderate to large shifts (one standard deviation or more) in the process.

The LR estimates are worse than SD estimates for all cases, except the case with respect to a parameter in the process shift distribution, when it also beat the PA

Table 2 **Shewhart Charts with Process Dynamics: Mean (Standard Error)** LCL/UCL $= \pm 2.81$, $\sigma/\sqrt{n} = 1$, $\mu_0 = 0$, $T \sim$ exp(20)

	ARL		d/du				d/dl			
μ_1	sim	ana	PA, L	PA, R	SD	ana	PA, L	PA, R	SD	ana
1.0	192.3	194.12	365.1	366.8	365.8	372.4	−225.1	−219.4	−215.9	−223.7
	(1.9)		(8.5)	(5.3)	(8.3)		(4.8)	(5.0)	(6.6)	
1.0	43.2	43.7	60.3	62.2	60.9	63.0	−6.0	−6.1	−5.7	−6.2
	(0.3)		(1.2)	(0.6)	(1.3)		(0.4)	(0.1)	(0.4)	
3.0	19.1	19.4	3.64	3.91	3.40	3.73	−2.65	−2.57	−2.38	−2.65
	(0.2)		(.06)	(.18)	(.21)		(.18)	(.05)	(.20)	

	$d/d\mu_0$				$d/d\mu_1$				$d/dE[T]$				
μ_1	PA	LR	SD	ana	PA	LR	SD	ana	PA, L	PA, R	LR	SD	ana
0.1	−2.5	12.2	1.5	0.00	−137.5	−129.0	−140.8	−148.7	−0.04	0.17	0.03	0.24	0.03
	(1.7)	(11.7)	(2.8)		(9.4)	(61)	(8.6)		(.08)	(.14)	(.14)	(.26)	
1.0	−0.04	4.1	−0.4	0.00	−54.3	−53.9	−54.7	−56.8	0.67	0.78	0.68	0.60	0.71
	(.38)	(2.9)	(.62)		(1.2)	(4.5)	(1.2)		(.08)	(.09)	(.04)	(.16)	
3.0	0.08	−0.3	0.04	0.0	−1.1	−1.1	−1.1	−1.1	0.79	0.79	0.79	0.76	0.82
	(.18)	(1.9)	(.28)		(.03)	(0.3)	(.04)		(.02)	(.02)	(.03)	(.03)	

estimates. This is because there is only a single random variable involved in the estimator, so the usual increasing (with length of simulation run) variance problem is absent.

7. Summary and Avenues for Further Research

We have derived PA and LR sensitivity estimators for control charts that can be efficiently implemented when Monte Carlo simulation is used for performance evaluation. Such estimators are useful for sensitivity analysis and optimization in the design of control charts. We considered the average run length performance measure and a fairly general control chart that includes both the Shewhart chart and the EWMA chart as special cases. Simulations performed on a few examples indicate that the RH PA estimators w.r.t. the control limits and the LR estimators w.r.t. a parameter

Table 3 **EWMA Charts: Mean (Standard Error)** $\alpha = 0.75$, LCL/UCL $= \pm 1.94$, $\sigma/\sqrt{n} = 1$

		d/du			d/dl			$d/d\mu$		
μ	ARL	PA, L	PA, R	SD	PA, L	PA, R	SD	PA	LR	SD
0.0	79.6	145.9	150.4	147.2	−145.3	−147.8	−145.2	−0.5	0.1	−3.8
	(0.8)	(2.5)	(2.6)	(3.4)	(2.6)	(2.6)	(3.5)	(3.0)	(16.6)	(4.0)
0.25	58.4	162.0	163.7	169.8	−34.5	−36.1	−34.3	−127.5	−124.6	−125.2
	(0.6)	(2.9)	(1.9)	(3.3)	(1.3)	(0.6)	(1.4)	(3.0)	(11.7)	(3.2)
1.0	9.3	19.8	19.8	19.9	−0.09	−0.10	−0.09	−19.7	−17.4	−19.8
	(0.1)	(.37)	(.20)	(.43)	(.02)	(.001)	(.03)	(.37)	(.87)	(.43)
2.0	2.1	2.01	2.02	1.98	0.00	0.00	0.00	−2.01	−1.93	−1.98
	(.02)	(.04)	(.02)	(.06)	(.00)	(.00)	(.00)	(.04)	(.08)	(.06)
3.0	1.2	0.34	0.34	0.37	0.00	0.00	0.00	−0.34	−0.33	−0.37
	(.00)	(.01)	(.002)	(.02)	(.00)	(.00)	(.00)	(.01)	(.02)	(.02)

Table 4 Variance Estimates for Shewhart Chart: Mean (95% c.i.) Based on 2000 Samples

		Variance of $dARL/d\mu$ Batched Estimator		
μ_1	Control Limits	Right-hand (RH) PA	SD CRN	SD non-CRN
0.0	±1.96	315	976	2634
		(296 335)	(919 1040)	(2478 2805)
0.0	±2.81	55428	138974	293651
		(52146 59031)	(130745 148009)	(276263 312742)
0.1	±1.96	326	1220	2865
		(306 347)	(1148 1300)	(2696 3052)
0.1	±2.81	57238	199348	361046
		(53848 60959)	(187545 212308)	(339667 384518)
1.0	±1.96	12.3	115	234
		(11.6 13.2)	(109 123)	(221 250)
1.0	±2.81	705	4858	7958
		(663 750)	(4571 5174)	(7487 8475)
2.0	±1.96	0.204	4.91	12.78
		(0.192 0.217)	(4.62 5.23)	(12.02 13.61)
2.0	±2.81	6.65	69.4	144.3
		(6.25 7.08)	(65.2 73.9)	(135.8 153.7)
3.0	±1.96	0.003	0.52	1.33
		(0.003 0.003)	(0.49 0.55)	(1.25 1.42)
3.0	±2.81	0.12	3.09	8.60
		(0.11 0.13)	(2.90 3.29)	(8.09 9.16)

in the process shift dynamics offer the potential for significant variance reduction over standard CRN SD estimators.

Although ARL performance measures are the most commonly used, in cases where the sample size and/or sampling interval are variable, other appropriate performance measures such as average time to signal and average number of observations to signal can also be handled. Also, the techniques are not restricted to just the mean of the process.

Application of the estimators to the economic optimal design of control charts is one natural avenue of research to pursue. In this approach, costs and profits are attached to various actions such as sampling and testing, investigation and correction, good production and nonconformance. Montgomery (1996) devotes a chapter to the problem of economic design; see also Ho and Case (1994) for a literature review. The focus has been on deriving analytical expressions for the time-average cost; in general, only the Shewhart case has yielded analytically tractable solutions, and, even in these cases, one must usually employ numerical analysis techniques to search for the optimum. A very general formulation of the economic design problem proposed by Lele (1997) assumes a control chart of the form analyzed in this paper. For the Shewhart X-bar charts, he was able to find the optimum design analytically, but for the EWMA charts and a Bayes' procedure, he had to employ Monte Carlo simulation. Derivatives were required in his solution procedure, obtained using finite differences; this would be a natural application for our approach, and is pursued in Fu et al. (1999).[1]

[1] M. C. Fu was supported in part by the National Science Foundation under Grants NSF EEC-9402384 and DMI-9713720 and by a 1996 Summer Research Grant from the Maryland Business School. J. Q. Hu was supported in part by the National Science Foundation under Grants EEC-9527422 and DDM-9212368. The authors thank the associate and departmental editors and two referees for useful comments and suggestions that led to a much improved paper; the inclusion of likelihood ratio estimators was suggested by the associate editor.

Appendix: Detailed Derivations and Theorems on Unbiasedness

Details of the derivations are provided in these appendices, with the exception of the coupling construction used in §4.3. Technical conditions and a full proof of unbiasedness are provided for the first case only. Due to space limitations here, refer to Fu and Hu (1997b) for the remainder of the detailed derivations and technical unbiasedness theorems.

Section 3.1. Control Limits

We derive in detail the estimator for the upper control limit u, and sketch the derivation for the lower control limit l. We consider both the left-hand and right-hand derivatives. For $\Delta u > 0$, it is obvious that $L(u + \Delta u) \neq L(u)$ if and only if $u < Y_{L(u)} \leq u + \Delta u$, i.e.,

$$E[L(u+\Delta u) - L(u)] = E[(L(u+\Delta u) - L(u))\mathbf{1}(u < Y_L \leq u + \Delta u)].$$

In this case, the perturbation Δu causes the test statistic to no longer signal out of control, and hence the run length is extended as a result (see Figure 1). In particular, we can think of the process as starting over with a new initial condition $\psi(X, Y_L)$, i.e., the additional length is equal to $J(\tau_j, \psi(X, Y_L))$. On the other hand, letting $z = (L, X_1, \ldots, X_{L-1})$, we have

$$\begin{aligned} E[(L(u+\Delta u) - L(u))\mathbf{1}(u < Y_L \leq u+\Delta u)] &= E[E[(L(u+\Delta u) - L(u))\mathbf{1}(u < Y_L \leq u+\Delta u)|z]] \\ &= E[E[(L(u+\Delta u) - L(u))|z, u < Y_L \leq u + \Delta u] \\ &\quad \times \mathbf{1}(Y_L > u) P(u < Y_L \leq u+\Delta u | z, Y_L > u)]. \end{aligned}$$

Note that in the last equation, we introduce the condition $Y_L > u$ so that our estimator will be simpler; otherwise, the implicit condition $Y_L \notin [l, u]$ would have to be used in calculating $P(u < Y_L \leq u + \Delta|z)$, resulting in a more complicated estimator.

Therefore, we have the probability rate term for the change is calculated via

$$\lim_{\Delta u \to 0^+} \frac{P(u < Y_L \leq u + \Delta u | z, Y_L > u)}{\Delta u} = \begin{cases} \dfrac{f_j(\psi_L^{-1}(u, Y_{L-1}))}{1 - F_j(\psi_L^{-1}(u, Y_{L-1}))} \dfrac{d\psi_L^{-1}(u, Y_{L-1})}{du} & L > 1; \\ \dfrac{f_j(u)}{1 - F_j(u)} & L = 1; \end{cases}$$

where $\psi_i^{-1}(\cdot, \cdot)$ denotes the inverse with respect to the first argument. Also,

$$\begin{aligned} \lim_{\Delta u \to 0^+} E[(L(u+\Delta u) - L(u))|z, u < Y_L \leq u + \Delta u] &= \lim_{\Delta u \to 0^+} E[L|T = \tau_j, \tilde{Y}_1 = \psi(X_1, Y_L), u < Y_L \leq u + \Delta u] \\ &= E[L|T = \tau_j, \tilde{Y}_1 = \psi(X_1, u)]. \end{aligned}$$

Putting this all together yields the final right-hand estimator for $d\mathrm{ARL}_j/du$, $j = 0, 1$, given by (8). For unbiasedness, we need

$$\begin{aligned} \lim_{\Delta u \to 0} \frac{1}{\Delta u} E[L(u+\Delta u) - L(u)] &= E\Big[\lim_{\Delta u \to 0} E[(L(u+\Delta u) - L(u))|z, u < Y_L \leq u + \Delta u] \\ &\quad \times \mathbf{1}(Y_L > u) \frac{1}{\Delta u} P(u < Y_L \leq u + \Delta u | z, Y_L > u)\Big]. \end{aligned}$$

As usual, we use the dominated convergence theorem to establish this. Basically, to apply the dominated convergence theorem the key condition required is the bound

$$E\Big[\sup_{0 \leq \Delta u \leq \epsilon} E[L|T = \tau_j, \tilde{Y}_1 = \psi(X_1, u + \Delta u)] \times \mathbf{1}(Y_L > u) \frac{1}{\Delta u} P(u < Y_L \leq u + \Delta u | z, Y_L > u)\Big] < \infty,$$

for any $\epsilon > 0$, for which the following conditions suffice to establish.

CONDITION (A1). *$\psi(\cdot, \cdot)$ is continuously differentiable and strictly increasing w.r.t. its first argument.*

CONDITION (A2). *$\psi^{-1}(\cdot, \cdot)$ is a decreasing function w.r.t. its second argument,*

$$\left| \frac{d\psi^{-1}(x, \cdot)}{dx} \right| < K$$

for all x, where $K > 0$ is a constant, and $F_j(\psi^{-1}(u, l)) < 1$, for $j = 0, 1$.

CONDITION (A3). *$|f_j(x)| < K$ for all x, $j = 0, 1$.*

CONDITION (A4). *$E[L|T = \tau_j, \tilde{Y}_1 = \psi(X_1, u + \Delta u)] < K$, for $0 \leq \Delta u \leq \epsilon$, $j = 0, 1$.*

The conditions on ψ in (A1) and (A2) are mild. For example, it can easily be shown that the EWMA control chart satisfies them. Condition (A3) is easily verifiable and holds for most of the well-known distributions. On the other hand, (A4) is a technical condition not as straightforward to verify as the other conditions, but viewed as a bound on average run length, it is reasonable to assume it holds for systems of practical interest.

THEOREM 1. *Under (A1)–(A4), (8) is an unbiased estimator for $d\mathrm{ARL}_j/du$, $j = 0, 1$.*

PROOF. In the estimator (8), the existence of $\psi^{-1}(\cdot, \cdot)$ and its differentiability are guaranteed by (A1). We note that due to (A4), the actual quantity on which we need to establish a bound is

$$\left| \frac{1}{\Delta u} P(u < Y_L \leq u + \Delta u | z, Y_L > u) \right| .$$

Since $\psi^{-1}(\cdot, \cdot)$ is decreasing w.r.t. its second argument due to (A2) and $l \leq Y_{L-1} \leq u$,

$$\left|\frac{1}{\Delta u} P(u < Y_L \le u + \Delta u | z, Y_L > u)\right|$$

$$= \left|\frac{1}{\Delta u} \frac{F_j(\psi^{-1}(u + \Delta u, Y_{L-1})) - F_j(\psi^{-1}(u, Y_{L-1}))}{1 - F_j(\psi^{-1}(u, Y_{L-1}))}\right|$$

(using (A2) and (A3))

$$\le \frac{1}{\Delta u} \frac{K|\psi^{-1}(u + \Delta u, Y_{L-1}) - \psi^{-1}(u, Y_{L-1})|}{1 - F_j(\psi^{-1}(u, l))},$$

$$\le \frac{K^2}{1 - F_j(\psi^{-1}(u, l))} \quad \text{via (A2).} \qquad \square$$

For the left-hand derivative $\Delta u < 0$, we know $L(u + \Delta u) \ne L(u)$ if $u + \Delta u < Y_i \le u$ for some $i < L$. In this case, we have a larger set of possible changes, in that any in-control signal prior to the first out-of-control may be altered to out of control, thus shortening the run length to that point (see Figure 2). If such a change occurs for sample i, the run length is reduced to i, and we have

$$\lim_{\Delta u \to 0^-} E[(L(u + \Delta u) - L(u))|z, u + \Delta u < Y_i \le u]$$

$$= i - E[L|Y_i = u_-, L > i, T = \tau_j]$$

$$= -E[L|T = \tau_j, \tilde{Y}_1 = \psi(X_1, u)].$$

For each term i, we condition on the set of all sample information except X_i itself, $z_i = \{L, X_1, \ldots, X_L\} \backslash \{X_i\}$, so that the probability rate term for the change is calculated via

$$\lim_{\Delta u \to 0^-} \frac{P(Y_i > u + \Delta u | l \le Y_i \le u)}{\Delta u}$$

$$= \begin{cases} -\dfrac{f_j(\psi_i^{-1}(u, Y_{i-1}))}{F_j(\psi_i^{-1}(u, Y_{i-1})) - F_j(\psi_i^{-1}(l, Y_{i-1}))} \dfrac{d\psi_i^{-1}(u, Y_{i-1})}{du} & L > 1; \\ \dfrac{f_j(u)}{F_j(u) - F_j(l)} \quad L = 1; \end{cases}$$

and the final left-hand estimator for $d\text{ARL}_j/du$ is given by (9).

Section 4.3. Sampling Frequency

We condition on $\eta(h) = i$. By definition of $\eta(h)$, we have $hi < T < h(i + 1)$. Based on the possible values of $\eta(h + \Delta h)$, we consider four cases:

$\eta(h + \Delta h) > i \qquad \mathbf{1}\{T > (h + \Delta h)(i + 1)\} = 0 \quad$ since $\Delta h > 0$;

$\eta(h + \Delta h) < i - 1 \qquad \mathbf{1}\{T \le (h + \Delta h)(i - 1)\} = 0$ for $\Delta h < h/(i - 1)$;

$\eta(h + \Delta h) = i \qquad L(h + \Delta h) = L(h)$ if $(h + \Delta h)i < T \le (h + \Delta h)(i + 1)$;

$\eta(h + \Delta h) = i - 1 \qquad \Rightarrow (h + \Delta h)(i - 1) \le T < (h + \Delta h)i.$

Thus, we need only consider further the case $(h + \Delta h)(i - 1) \le T < (h + \Delta h)i$. The probability rate term in this case is calculated in the usual way by

$$P(\eta(h + \Delta h) = i - 1|\eta(h) = i)$$

$$= P((i - 1)(h + \Delta h) < T \le i(h + \Delta h)|ih < T \le (i + 1)h)$$

$$= P(ih < T \le i(h + \Delta h)|ih < T \le (i + 1)h), \quad 0 < \Delta h < h/i$$

$$= \frac{P(ih < T \le i(h + \Delta h))}{P(ih < T \le (i + 1)h)} = \frac{F(i(h + \Delta h)) - F(ih)}{F((i + 1)h) - F(ih)}$$

$$\Rightarrow \lim_{\Delta h \to 0} \frac{P(\eta(h + \Delta h) = i - 1|\eta(h) = i)}{\Delta h} = \frac{i \cdot f(ih)}{F((i + 1)h) - F(ih)},$$

and the final estimator is given by the following:

$$\frac{\eta \cdot f(\eta h)}{F((\eta + 1)h) - F(\eta h)} [L^{PP} - L^{DNP}], \tag{46}$$

where the three paths are defined as follows:

NP, the nominal (original) sample path: $\eta h < T \le (\eta + 1)h$;
DNP, the degenerated nominal path: $T = [\eta h]_+ \Rightarrow \eta^{DNP} = \eta$;
PP, the perturbed path: $T = [\eta h]_- \Rightarrow \eta^{PP} = \eta - 1$.

Since the run-length performance measure L does not depend on the actual value of T but just on η, L is the same on *DNP* as on *NP*, so $L^{DNP} = L$ and (46) yields (41).

We now turn to the left-hand derivative $\Delta h < 0$, to consider $P(\eta(h + \Delta h) = i + 1|\eta(h) = i)$:

$$P(\eta(h + \Delta h) = i + 1|\eta(h) = i)$$

$$= P((i + 1)(h + \Delta h) < T \le (i + 2)(h + \Delta h)|ih < T \le (i + 1)h)$$

$$= P((i + 1)(h + \Delta h) < T \le (i + 1)h|ih < T \le (i + 1)h), \quad -h(i + 2) < \Delta h < 0$$

$$= \frac{P((i + 1)(h + \Delta h) < T \le (i + 1)h)}{P(ih < T \le (i + 1)h)}$$

$$= \frac{F((i + 1)h) - F((i + 1)(h + \Delta h))}{F((i + 1)h) - F(ih)}$$

$$\Rightarrow \lim_{\Delta h \to 0} \frac{P(\eta(h + \Delta h) = i + 1|\eta(h) = i)}{\Delta h} = -\frac{(i + 1) \cdot f((i + 1)h)}{F((i + 1)h) - F(ih)},$$

and the final estimator is given by the following:

$$\frac{(\eta + 1) \cdot f((\eta + 1)h)}{F((\eta + 1)h) - F(\eta h)} [L^{DNP} - L^{PP}], \tag{47}$$

where the three paths are defined as follows:

NP, the nominal (original) sample path: $\eta h < T \le (\eta + 1)h$;
DNP, the degenerated nominal path: $T = [(\eta + 1)h]_- \Rightarrow \eta^{DNP} = \eta$;
PP, the perturbed path: $T = [(\eta + 1)h]_+ \Rightarrow \eta^{PP} = \eta + 1$.

Again, since L is the same on *DNP* as on *NP*, we have $L^{DNP} = L$, and the estimator (47) corresponds to (42).

Section 4.4. Process "Drift" Parameters
To obtain the left-hand derivative given by (44), consider the probability rate for $\Delta T < 0$:

$$P(\eta(T+\Delta T)=i-1|\eta(T)=i)$$
$$=P((i-1)h<T+\Delta T\le ih|ih<T\le(i+1)h)$$
$$=P(ih<T\le ih-\Delta T|ih<T\le(i+1)h)$$
$$=\frac{F(ih-\Delta T)-F(ih)}{F((i+1)h)-F(ih)}.$$

To obtain the right-hand derivative given by (45), consider the probability rate for $\Delta T > 0$:

$$P(\eta(T+\Delta T)=i+1|\eta(T)=i)$$
$$=P((i+1)h<T+\Delta T\le(i+2)h|ih<T\le(i+1)h)$$
$$=P(ih-\Delta T<T\le(i+1)h|ih<T\le(i+1)h)$$
$$=\frac{F((i+1)h)-F((i+1)h-\Delta T)}{F((i+1)h)-F(ih)}.$$

References

Albin, S. A., L. Kang, G. Shea. 1997. An X and EWMA chart for individual observations. *J. Quality Tech.* **29** (1) 41–48.

Barish, N. N., N. Hauser. 1963. Economic design for control decisions. *J. Industrial Eng.* **14** 125–134.

Baxley, R. V. 1995. An application of variable sampling interval control charts. *J. Quality Tech.* **27** (4) 275–282.

Cao, X. R. 1994. *Realization Probabilities: The Dynamics of Queueing Systems*. Springer Lecture Notes in Control and Optimization. **194** Springer-Verlag, New York.

Crowder, S. V. 1987. A simple method for studying run-length distributions of exponentially weighted moving average charts. *Technometrics* **29** (4) 401–407.

Fu, M. C. 1994. Optimization via simulation: A review. *Ann. Oper. Res.* **53** 199–248.

——, J. Q. Hu. 1992. Extensions and generalizations of smoothed perturbation analysis in a generalized semi-Markov process framework. *IEEE Trans. Automat. Control* **37** 1483–1500.

——, ——. 1997a. *Conditional Monte Carlo: Gradient Estimation and Optimization Applications*. Kluwer Academic Publishers, Boston, MA.

——, ——. 1997b. Application of perturbation analysis to the design and analysis of control charts. ISR Technical Report TR 97-91, University of Maryland, College Park, MD. Available in Postscript or PDF at `www.isr.umd.edu/TechReports/ISR/1997`.

——, S. Lele, T. W. Vossen. 1999. Economic design of control charts using Monte Carlo simulation. Working paper, University of Maryland, MD.

Glasserman, P. 1991. *Gradient Estimation Via Perturbation Analysis*. Kluwer Academic Publishers, Boston, MA.

Gong, W. B., Y. C. Ho. 1987. Smoothed perturbation analysis of discrete-event dynamic systems. *IEEE Trans. Automat. Control* **32** 858–867.

Grimshaw, S. D., F. B. Alt. 1997. Control charts for quantile function values. *J. Quality Tech.* **29** (1) 1–7.

Ho, Y. C., X. R. Cao. 1991. *Discrete Event Dynamic Systems and Perturbation Analysis*. Kluwer Academic Publishers, Boston, MA.

Ho, C., K. E. Case. 1994. Economic design of control charts: A literature review for 1981–1991. *J. Quality Tech.* **26** (1) 39–53.

Lele, Shreevardhan. 1996. Steady state analysis of three process monitoring procedures in quality control. Ph.D. dissertation, University Microfilms Inc., University of Michigan, Ann Arbor, MI.

——. 1997. Economic design of control charts using steady state distributions. Working paper.

Montgomery, D. C. 1996. *Introduction to Statistical Quality Control*, 3rd ed. John Wiley & Sons, New York.

Rubinstein, R. Y., Shapiro, A. 1993. *Discrete Event Systems: Sensitivity Analysis and Stochastic Optimization by the Score Function Method*. John Wiley & Sons, New York.

Accepted by Pierre L' Ecuyer; received April 9, 1997. This paper has been with the authors 6 months for 4 revisions.