Estimating the Out-of-Control Rate from Control Chart Data in the Presence of Multiple Causes and Process Improvement

Diane P. Bischak
Edward A. Silver

Operations Management Area
Haskayne School of Business
University of Calgary
2500 University Drive NW
Calgary, AB T2N 1N4
CANADA

Corresponding author:
diane.bischak@haskayne.ucalgary.ca
Phone 403-220-3368
Fax 403-210-3327

October, 2003
(Revised April, 2004)

ABSTRACT

We consider a process that is monitored with an $\bar{X}$ chart. The process may go out of control due to the occurrence of one of several independent assignable causes. After the process has gone out of control and the assignable cause has been determined, the process undergoes improvement that results in a reduction of the rate due to that cause. We develop a Bayesian estimator of the rate at which the process is going out of control as well as of the rates of the individual assignable causes. The estimation procedure makes use of the Markov chain Monte Carlo technique of data augmentation. We provide numerical illustrations that indicate how the posterior results depend upon the data and upon the choice of the parameters of the prior distributions.

Keywords: $\bar{X}$ control chart, multiple assignable causes, process improvement, parameter estimation, Bayesian estimation, Markov chain Monte Carlo, data augmentation.

1. INTRODUCTION

The problem addressed in this paper is the estimation of the rate at which a process is going out of control due to assignable causes. We impose several realistic complications within this basic problem. First, the information available for estimation is limited to that obtained while an $\bar{X}$ control chart is being used to monitor the process. Second, there are J independent assignable causes, so that the total out-of-control rate is the sum of the rates at which these individual causes occur; we would like to obtain estimates of these individual rates as well. Third, the process undergoes improvement every time we detect an out-of-control condition, in that the rate at which the identified assignable cause occurs is reduced to a fraction of its current value. Fourth, we wish to incorporate prior information that a manager may have concerning how frequently the process goes out of control due to particular causes.

Estimating the out-of-control rate is not as easy as it may first appear, even when a manager has had considerable experience with a process. Since in statistical process control a process is observed only through relatively infrequent samples of output, the exact time at which the process goes out of control cannot be determined, although a lower bound on that time is the last moment at which a false alarm occurred, i.e., an investigation ascertained that the process was still in control. Furthermore, it may require several samples to detect the out-of-control condition, as there is some probability that a sample from an out-of-control process may suggest that the process is still in control. Thus the rate is somewhat problematic to estimate.

An estimate of the out-of-control rate is necessary to design a control chart that is based on economic models (e.g., Duncan 1956; Lorenzen and Vance 1986) or economic-statistical models (e.g., Saniga 1989; Keats et al. 1997). These models take into consideration the relevant costs of a control charting program and, in the latter case, the statistical properties of the resulting chart, such as the Type I and Type II error probabilities. Several economic models for control charts in the literature go further and assume multiple assignable causes, requiring estimates of the associated individual rates (Duncan 1971; Chiu and Wetherill 1974; Tagaras and Lee 1988; Chung 1994).

Estimating these rates and explicitly taking into account the fact that the out-of-control rates will change due to process improvement will itself facilitate process improvement; as Woodall (1986) suggests, “by considering [the parameter for the time before a shift is expected to occur] to be fixed, the economic approach [to the design of control charts] can be a barrier to process improvement.” A model for the economic design of $\bar{X}$ control charts that recognize process improvement was presented in Silver and Rohleder (1999), who found that including the process improvement and using it to dynamically update the width of the control chart limits and the size of the sample and sampling interval could result in substantial cost savings.

However, little work has been done on estimation of the input parameters for these charts. Chiu (1976) presents a very simple approach for estimating the out-of-control rate and other parameters required for the economic design of $\bar{X}$ charts, but he does not examine the performance of this estimator. Bischak and Silver (2001) use a classical statistical approach to develop four estimators for the rate when assignable causes are not distinguished. A maximum likelihood estimator and another estimator that is a function of the expected time until the process goes out of control, given the time of the last sample known to be from the in-control process and the time of the sample at which detection occurred, were shown to be superior to a simple, relatively naïve, estimator. Lo (1994) addresses the problem of estimating parameters for economic control charts for attributes, including the rate at which assignable causes occur. He uses an iterative technique known as the Expectation-Maximization (EM) algorithm to develop a maximum likelihood estimate of the rate. The EM algorithm has some similarities to the approach in the present paper in that it depends upon an unobservable value – in Lo's case, the number of samples required to detect an out-of-control condition after the process has gone out of control – which, if known, would greatly simplify the form of the likelihood function.

We take a Bayesian approach to estimation (see, e.g., Carlin and Louis 2000). Bayesian methods are appropriate in cases where a manager would have some knowledge either of the process of interest or a similar process and where the amount of available data is limited. Prior distributions are placed on the parameters of interest, namely the rates of occurrence of the individual causes. The likelihood function, which is the probability of seeing the control chart data given particular values of these parameters, must be derived. As data are observed, the likelihood function is used to update the prior distributions on the rates, thus obtaining posterior distributions. We then use estimates of the expected values of these posterior distributions as estimates of the rates.

Bayesian updating procedures have previously been used in the quality control literature for process control, where the probability that the process has moved into an out-of-control state is updated dynamically and may be used to adaptively select the sampling interval, sample size, or control limits for a control chart (Tagaras 1996; Calabrese 1995; Porteus and Angelus 1997; Tagaras and Nikolaidis 2002) and to set control limits when the parameters of the process distribution are unknown (Menzefricke 2002). Bayesian estimation of the out-of-control rate from control chart data has been considered only by Silver and Bischak (2004), who developed a posterior estimator of the expected value of the rate in the case where there is no process improvement. The inclusion of process improvement on individual assignable causes in the current paper makes the estimation of the total rate significantly more challenging.

In the remainder of the paper, we first define our assumptions and notation. We then derive the likelihood function in the case that the exact out-of-control times are unknown and in the case that they are known. We make use of the simpler form of the likelihood with known out-of-control times, through a data augmentation procedure, to obtain estimates of the rates for each assignable cause and the overall rate. Finally, we provide numerical illustrations of the data augmentation procedure on a base case, with variations to show the effects of changes in the prior distributions, the number of assignable causes, and the data. We conclude with suggestions for future work.

2. ASSUMPTIONS AND NOTATION

We use the following set of assumptions, most of which are consistent with the literature on $\overline{X}$ charts:

i) Samples of size *n* are taken from the process of interest every *h* periods. (Without loss of generality, we assume throughout that $h=1$, so that *h* defines the time unit.)

ii) Individual observations have a standard deviation σ, assumed to be known.

iii) Sample averages are assumed to be normally distributed.

iv) The $\overline{X}$ chart has control limits set at a distance of $\pm k\,\sigma/\sqrt{n}$ units from the in-control process mean. (The more general case where *h*, *n*, and/or *k* varied from cycle to cycle could easily be handled.)

v) There are *J* possible assignable causes that occur independently of each other. The time until cause *j* is activated is exponentially distributed with initial rate parameter λ_j (j = 1, 2, ..., *J*). As soon as a cause activates, no other assignable causes can occur until the process is put back into control.

vi) When an assignable cause occurs (i.e., the process goes out of control), the mean shifts by a known constant distance, $+\delta\,\sigma$ or $-\delta\,\sigma$, independent of the specific cause.

vii) Once an out-of-control situation is detected, the underlying assignable cause (j) is found and corrected. The rate at which this cause is occurring is reduced to a fraction ϕ_j of its previous rate of occurrence. (We assume that $\phi_j > 0$, but it can be arbitrarily small.) The process is returned to its in-control state and the clock is reset to zero.

The following additional notation is used.

$\Phi(\ )$: the cumulative standard normal distribution function.

$\alpha = 2\,\Phi(-k)$: the probability that a point (sample mean) will fall outside the control limits while the process is in control, i.e., the probability of a false alarm.

$p = \Phi\left(\delta\sqrt{n} - k\right) + \Phi\left(-\delta\sqrt{n} - k\right)$: the probability that a point will fall outside the control limits after the process has gone out of control.

$\lambda_j(i)$: the rate at which assignable cause *j* will occur in cycle *i*, after $i-1$ cycles of data have been seen and the process improvement at the end of those cycles has been completed.

$\lambda_j \equiv \lambda_j(1)$, the initial rate at which assignable cause *j* occurs.

$\underline{\lambda}$: vector representation of the initial rates of the *J* assignable causes.

$z(i) = \sum_{j=1}^{J} \lambda_j(i)$: the total rate at which underlying assignable causes will occur during cycle *i*.

$z \equiv z(1)$.

β and γ: parameters of the prior Gamma distribution on the initial total rate, *z*.

β_j and γ_j: parameters of the prior Gamma distribution on the initial rates λ_j, *j*=1, 2,…, *J*.

$\beta_j(i)$ and $\gamma_j(i)$: parameters of the posterior Gamma distribution on rate *j* after *i* cycles of data and the related process improvement, *j*=1, 2,…, *J*.

N: the total number of cycles of data.

r(*i*): the sample number of the last false alarm in cycle *i*.

m(*i*): the sample number at which the out-of-control condition is detected in cycle *i*.

c(*i*): the assignable cause detected in cycle *i*.

t(*i*): the time at which the assignable cause in cycle *i* occurred.

$n_j(i)$: the number of times cause *j* was seen in the first *i* cycles.

3. THE LIKELIHOOD FUNCTION

3.1. The likelihood when out-of-control times are unknown

To set up a Bayesian procedure to estimate the out-of-control rates for each assignable cause, we first repeat (from Bischak and Silver 2001) the likelihood function in the case in which there is a single cycle of data (i.e., one triplet *r*, *m*, *c*) for which the exact out-of-control time is unknown and there is no process improvement. Note that this likelihood function is a function of the initial rates, before any data are seen.

Since a sample mean that falls outside the control limits is determined to be a false alarm if we find that the process is still in control at that time, the latest time during the cycle that we are certain the process is in control is the time at which the last false alarm is seen. Suppressing the cycle index for now, suppose that the last false alarm we see occurs at time *r*, the process is detected as out of control at time *m*, and the cause is found to be cause *c*. The likelihood function is the probability of seeing *r*, *m*, and *c*, given that the rates of the various assignable causes occurring are $\underline{\lambda}$ (with a total rate of $z = \sum \lambda_j$). Letting i_0 be the interval in which the process goes out of control (where $r < i_0 \le m$), we have

$$P(i_0 \mid \underline{\lambda}) = e^{-z(i_0-1)} - e^{-z i_0}, \tag{1}$$

$$P\left(c \mid \underline{\lambda}, i_0\right) = P\left(c \mid \underline{\lambda}\right) = \lambda_c / z, \tag{2}$$

$$P(r \mid \underline{\lambda}, i_0, c) = P(r \mid i_0) = \begin{cases} (1-\alpha)^{i_0-1} & r = 0 \\ \alpha(1-\alpha)^{i_0-r-1} & r = 1, 2, \ldots, i_0 - 1, \end{cases} \tag{3}$$

and

$$P(m \mid r, \underline{\lambda}, i_0, c) = P(m \mid i_0) = p(1-p)^{m-i_0}, \quad m = i_0, i_0+1, \ldots . \tag{4}$$

Equations (1) and (2) follow from the assumption that the times between out-of-control event occurrences for each independent cause are exponentially distributed. Figure 1 depicts a cycle in which $r = 2$, $i_0 = 3$, and $m = 4$.

[Insert figure 1 about here]

From equations (1) through (4), we see that the likelihood function for this single cycle is

$$L \equiv P(r, m, c \mid \underline{\lambda}) = \sum_{i_0=r+1}^{m} P(i_0 \mid \underline{\lambda}) \cdot P(c \mid \underline{\lambda}, i_0) \cdot P(r \mid \underline{\lambda}, i_0, c) \cdot P(m \mid r, \underline{\lambda}, i_0, c)$$

$$= \begin{cases} \displaystyle\sum_{i_0=1}^{m} \left[e^{-z(i_0-1)} - e^{-zi_0} \right] (1-\alpha)^{i_0-1} p (1-p)^{m-i_0} \lambda_c / z & r = 0 \\ \displaystyle\sum_{i_0=r+1}^{m} \left[e^{-z(i_0-1)} - e^{-zi_0} \right] \alpha (1-\alpha)^{i_0-r-1} p (1-p)^{m-i_0} \lambda_c / z & r = 1, 2, \ldots, m-1 \end{cases}$$

$$= \begin{cases} (1-e^{-z}) p (1-p)^{m-1} \left[\dfrac{1-a^m}{1-a} \right] \lambda_c / z & r = 0 \\ (1-e^{-z}) \alpha (1-\alpha)^{-r} p (1-p)^{m-1} \left[\dfrac{a^r - a^m}{1-a} \right] \lambda_c / z & r = 1, 2, \ldots, m-1, \end{cases}$$

where

$$a = e^{-z}(1-\alpha)/(1-p).$$

We now make the additional assumption that process improvement in the form of a fractional reduction in the rate λ_j occurs each time cause *j* is seen. If $n_j(i)$ is the number of times cause *j* was seen in the first *i* cycles, then

$$\lambda_j(i) = \phi_j^{\,n_j(i-1)} \lambda_j$$

is the rate at which cause *j* will occur in cycle *i*. Let

$$z(i) = \sum_{j=1}^{J} \lambda_j(i)$$

and $a_i = e^{-z(i)}\ 1-\alpha\ /\ 1-p$. Since each cycle is independent of the others given the initial rates, the likelihood function after *N* cycles of data have been seen is $L = L_1 \cdot L_2 \cdots L_N$, where

$$L_i = K_i \left[\frac{1-e^{-z(i)}}{z(i)(1-a_i)} \right] (a_i^{\,r(i)} - a_i^{\,m(i)}) \lambda_{c(i)}(i)\,, \qquad (5)$$

K_i is a constant independent of the rates equal to $p(1-p)^{m(i)-1}$ when $r(i) = 0$ and $\alpha(1-\alpha)^{-r(i)}\, p(1-p)^{m(i)-1}$ when $r(i) > 0$, and $c(i)$ is the observed cause. Note, however, that *L* is not separable into expressions for the individual rates λ_j. Thus the denominator of the posterior distribution on *z*, which is the integral of the prior and the likelihood over the interval (0, ∞) for each cause, would be a multiple integral of dimension *J*. Numerical integration of this expression would be extremely difficult, other than for relatively small values of *J*.

3.2. The likelihood and posterior when out-of-control times are known

Suppose that somehow we knew the exact time $t(i)$ at which the process went out of control in cycle i. As an example, say that we have seen N=2 cycles of data, and there are only two assignable causes, with cause 1 observed in cycle 1 and cause 2 observed in cycle 2. In terms of the initial rates, the likelihood function for cycle 1 would be

$$L_1 = \frac{\lambda_1}{\lambda_1+\lambda_2} \cdot \;\; \lambda_1+\lambda_2 \;\; e^{-(\lambda_1+\lambda_2)t(1)} = \lambda_1 e^{-(\lambda_1+\lambda_2)t(1)},$$

the likelihood function for cycle 2 would be

$$L_2 = \frac{\lambda_2}{\phi_1\lambda_1+\lambda_2} \cdot (\phi_1\lambda_1+\lambda_2) e^{-(\phi_1\lambda_1+\lambda_2)t(2)} = \lambda_2 e^{-(\phi_1\lambda_1+\lambda_2)t(2)},$$

and the likelihood function for the two cycles together would be

$$L = L_1 \cdot L_2 = \lambda_1\lambda_2 e^{-[(\lambda_1+\lambda_2)t(1)+(\phi_1\lambda_1+\lambda_2)t(2)]}.$$

Again, note that this is the likelihood function with respect to the initial rates.

To facilitate Bayesian estimation, an appropriate choice of prior distribution on each initial rate λ_j would be a Gamma (β_j, γ_j), that is,

$$f(\lambda_j) = \frac{\lambda_j^{\beta_j-1} \gamma_j^{\beta_j} e^{-\gamma_j\lambda_j}}{\Gamma(\beta_j)}, \qquad j = 1, 2, \ldots, J\ .$$

This is a convenient choice because the form of the likelihood function makes this a conjugate prior distribution (see, e.g., DeGroot 1970); the prior can be easily updated to the posterior, which will also be a Gamma distribution. Continuing the above example, we calculate the posterior as

$$f(\lambda_1,\lambda_2)=\frac{\frac{\lambda_1^{\beta_1-1}\gamma_1^{\beta_1}e^{-\gamma_1\lambda_1}}{\Gamma(\beta_1)}\cdot\frac{\lambda_2^{\beta_2-1}\gamma_2^{\beta_2}e^{-\gamma_2\lambda_2}}{\Gamma(\beta_2)}\cdot\lambda_1\lambda_2 e^{-[(\lambda_1+\lambda_2)t(1)+(\phi_1\lambda_1+\lambda_2)t(2)]}}{\int_{\lambda_1}\int_{\lambda_2}\frac{\lambda_1^{\beta_1-1}\gamma_1^{\beta_1}e^{-\gamma_1\lambda_1}}{\Gamma(\beta_1)}\cdot\frac{\lambda_2^{\beta_2-1}\gamma_2^{\beta_2}e^{-\gamma_2\lambda_2}}{\Gamma(\beta_2)}\cdot\lambda_1\lambda_2 e^{-[(\lambda_1+\lambda_2)t(1)+(\phi_1\lambda_1+\lambda_2)t(2)]}d\lambda_2 d\lambda_1}$$

$$=\frac{\frac{\lambda_1^{\beta_1}\gamma_1^{\beta_1}e^{-(\gamma_1+t(1)+\phi_1 t(2))\lambda_1}}{\Gamma(\beta_1)}\cdot\frac{\lambda_2^{\beta_2}\gamma_2^{\beta_2}e^{-(\gamma_2+t(1)+t(2))\lambda_2}}{\Gamma(\beta_2)}}{\int_{\lambda_1}\int_{\lambda_2}\frac{\lambda_1^{\beta_1}\gamma_1^{\beta_1}e^{-(\gamma_1+t(1)+\phi_1 t(2))\lambda_1}}{\Gamma(\beta_1)}\cdot\frac{\lambda_2^{\beta_2}\gamma_2^{\beta_2}e^{-(\gamma_2+t(1)+t(2))\lambda_2}}{\Gamma(\beta_2)}d\lambda_2 d\lambda_1}$$

$$=\frac{\lambda_1^{\beta_1}(\gamma_1+t(1)+\phi_1 t(2))^{(\beta_1+1)}e^{-(\gamma_1+t(1)+\phi_1 t(2))\lambda_1}}{\Gamma(\beta_1+1)}\cdot\frac{\lambda_2^{\beta_2}(\gamma_2+t(1)+t(2))^{(\beta_2+1)}e^{-(\gamma_2+t(1)+t(2))\lambda_2}}{\Gamma(\beta_2+1)}.$$

This expression can be seen to be the product of two Gamma probability density functions, the one a density on λ_1 with parameters β_1+1 and $\gamma_1+t(1)+\phi_1 t(2)$, and the other a density on λ_2 with parameters β_2+1 and $\gamma_2+t(1)+t(2)$. Note that this is the posterior on the original λ_j values, not the rates $\lambda_j(3)=\phi_j^{n_j(2)}\lambda_j=\phi_j\lambda_j$, j=1, 2, that would be in effect after the two cycles. To obtain the posterior on the updated rates, we use the fact that if $Y=\phi X$ and X has a Gamma (β,γ) distribution, then Y ~ Gamma$(\beta,\gamma/\phi)$ (see, for example, Evans et al. 2000, p. 24-25). Thus the posterior distribution of the updated rates as the third cycle begins is the product of a Gamma distribution on $\lambda_1(3)$ with parameters $[\beta_1+1,\ (\gamma_1+t(1)+\phi_1 t(2))/\phi_1]$ and a Gamma distribution on $\lambda_2(3)$ with parameters $[\beta_2+1,\ (\gamma_2+t(1)+t(2))/\phi_2]$.

In the case that we know the $t(i)$s and use a Gamma (β_j, γ_j) prior on λ_j , we can derive a general expression for the parameters of the posterior distribution on $\lambda_j(N+1)$ after N cycles of data have been seen and process improvement has taken place. The posterior distribution on rate j as we enter cycle N + 1 will be a Gamma ($\beta_j(N), \gamma_j(N)$), where

$$\beta_j(N) = \phi_j^{n_j(N)}(\beta_j + n_j(N)) . \tag{6}$$

We can determine $\gamma_j(N)$ by letting $n_j(0) = 0$ for j = 1, 2,…, J and observing that

$$n_j(i) = n_j(i-1) \quad \text{for } j \neq c(i)$$

and

$$n_{c(i)}(i) = n_{c(i)}(i-1) + 1 .$$

Then if we let $\gamma_j(0) = \gamma_j$ for j = 1, 2,…, J, we have

$$\gamma_j(i) = \gamma_j(i-1) + \phi_j^{n_j(i-1)} t(i), \qquad i = 1, 2, ..., N. \tag{7}$$

4. THE DATA AUGMENTATION PROCEDURE

In the preceding two sections, we developed two versions of the likelihood function in the case that process improvement is present. The first version, which is the product of expressions L_i such as that given for a representative cycle i in equation (5), holds in the case that the exact times at which the process went out of control are unknown. However, it is difficult to use this likelihood to obtain estimates of posterior values, as the posterior distribution will include a multiple integral of dimension J. It turns out to be much simpler and more intuitive to use the second version of the likelihood function derived in the last section to obtain posterior estimates. At first glance it would appear that this version of the likelihood is not helpful in the problem that we address, since the $t(i)$s are unknown. However, another Markov chain Monte Carlo procedure known as data augmentation (Tanner, 1996) allows us to make use of this "easier" likelihood function to develop posterior estimates of the rates $\lambda_j(N+1)$. Data augmentation has been used by Nair et al. (2001) to determine posterior distributions for mixing parameters in quality and reliability models, and a related Markov chain Monte Carlo procedure, Gibbs sampling, has been implemented by Hamada et al. (2003) to obtain posterior values in calibration models.

The data augmentation algorithm is displayed in figure 2. The procedure, which is run after N cycles of data have been seen, consists of an inner loop and an outer loop. For the initial execution of the inner loop, the Gamma ($\beta_j(N), \gamma_j$) distribution serves as an approximation to the posterior distribution on $\lambda_j(N+1)$. The first parameter of this distribution is as given in equation (6). We would prefer to have $\gamma_j(N)$, derived from equation (7), as the second parameter of this distribution, but $\gamma_j(N)$ is a function of the $t(i)$s, unknown to us. (In subsequent executions of the inner loop we are able to avoid this difficulty through data augmentation, as described below.) From equation (7), it is clear that $\gamma_j(i) \geq \gamma_j$, so that the use of γ_j tends to underestimate the $\gamma_j(i)$'s.

[Insert figure 2 about here]

The inner loop is then executed *l* times, as follows. Candidate values for $\lambda_j(N+1)$, denoted by $\tilde{\lambda}_j(N+1)$, j = 1, …, *J*, are drawn randomly from the current approximations to the posterior distributions on the $\lambda_j(N+1)$ s. Using the generated candidate values, we first enlarge these rates back to what would have been their starting values before any process improvement was performed, i.e., back to $\tilde{\lambda}_j = \tilde{\lambda}_j(N+1)/\phi_j^{n_j(N)}$. We then augment the observed data in the *N* cycles by using these $\tilde{\lambda}_j$ s to generate a set of possible out-of-control times, one for each cycle, according to the following procedure. The full rate $\tilde{z}(i)$ at which out-of-control conditions are posited to have occurred during a particular cycle *i* is the sum of the $\tilde{\lambda}_j(i)$ s, i.e.,

$$\tilde{z}(i) = \sum_{j=1}^{J} \tilde{\lambda}_j(i) \cdot \phi_j^{n_j(i-1)} .$$

For each interval $s(i) \in \{1, 2, ..., m(i) - r(i)\}$ after the last false alarm and before detection, let $G_{s(i)}$ be the event that the process went out of control during that interval. Then (Bischak and Silver 2001)

$$P(G_{s(i)} \mid r(i), m(i)) = \frac{(e^{-\tilde{z}(i)(s(i)-1)} - e^{-\tilde{z}(i)s(i)})(1-p)^{-s(i)}(1-b)}{b(e^{\tilde{z}(i)} - 1)(1 - b^{(m(i)-r(i))})} = \frac{b^{s(i)-1}(1-b)}{1 - b^{m(i)-r(i)}},$$

where $b = e^{-\tilde{z}(i)}/(1-p)$.

We randomly select, according to the probabilities $P(G_{s(i)} \mid r(i), m(i))$, an interval $\tilde{s}(i)$ in which the process could have gone out of control. We then randomly generate an imputed time $\tilde{t}(i)$ at which the process could have gone out of control in that interval.

The generation of $\tilde{t}(i)$ is via the inverse transform method using the assumption of exponentially distributed times until the process goes out of control, as follows:

$$\tilde{t}(i) = -(1/\tilde{z}(i))\ln(u \cdot e^{-\tilde{z}(i)\tilde{s}(i)} + (1-u) \cdot e^{-\tilde{z}(i)(\tilde{s}(i)-1)}),$$

where u is an observation from the Uniform (0,1) distribution. This concludes one execution of the inner loop.

When the inner loop has been repeated l times, we have l sets of imputed $\tilde{t}(i)$ values. At this point, a reasonable estimate of the actual posterior density of $\lambda_j(N+1)$ is the mixture of l Gamma ($\beta_j(N), \tilde{\gamma}_j(N)$) densities, where the $\tilde{\gamma}_j(N)$ parameter of each density is as given in equation (7) but with the $t(i)$ s replaced by the corresponding set of $\tilde{t}(i)$ s. In the outer loop, we then replace the current approximation to the posterior distribution by this mixture, so that on each of the following l iterations of the inner loop we draw new candidate values of $\tilde{\lambda}_j(N+1)$ from this mixture, i.e., on each inner loop iteration we select one of these l densities and use it as the source of new, randomly drawn $\tilde{\lambda}_j(N+1)$ values. The posterior distribution of $\lambda_j(N+1)$ that results from execution of the data augmentation algorithm will converge to the true posterior distribution under mild regularity conditions (Tanner and Wong 1987) that are met in our context.

5. NUMERICAL ILLUSTRATIONS

We demonstrate the data augmentation procedure with a series of numerical illustrations. After presenting a base case of priors and data, we show the effects of a single change in each of a number of aspects of the priors or the data. For all numerical illustrations we set the values of the control chart parameters to $\delta = 1$, n = 5, and k = 2, and for simplicity we set $\phi_j = \phi$ for all j. Ten replications of the entire algorithm were used. In each replication, the outer loop was executed L = 30 times; the inner loop was executed l = 2 times for the first 20 executions of the outer loop, after which l was doubled in every successive pass, so that in the last pass l = 2048. The choices of l are aligned with Tanner and Wong (1987), who suggest an initially small value of l that is increased with successive iterations. The choice of L was made on the basis of plots of the convergence of the estimates. A typical convergence pattern is shown in figure 3, where estimates of E[z] for the base case with ϕ = 1.0 are plotted. Each plotted point is the average of the l values of $\beta_j(N)/\tilde{\gamma}_j(N)$ obtained after a pass through the outer loop. It is clear from figure 3 that each replication converges to a very similar estimate. The estimated posterior mean reported in the experiments is the average across all ten replications of the averages obtained in the final outer loop iteration.

[Insert figure 3 about here]

5.1. BASE CASE

In the base case numerical illustration, we assume that there are ten assignable causes, the first three of which have previously been seen, as shown in table 1. Using a Gamma prior on the total rate z with $\beta = 9$ and $\gamma = 180$, the mean is 0.05 and the coefficient of variation (CV) is 1/3. The parameters of the prior Gamma distribution on λ_j are $\beta_j = q_j \beta$, where q_j is the proportion of out-of-control conditions attributed to cause *j*, and $\gamma_j = 180$ (see, for example, Evans et al. 2000, p. 99-100). There is considerable discussion in Silver and Bischak (2004) regarding how to determine the parameters of the prior distributions.

[Insert table 1 about here]

As shown in table 2, four cycles of data are seen before the prior is updated, so that $N = 4$. In cycle 2 an originally unknown cause is observed and labeled as cause 4. The average of the four observed m values is 19.5. This implies an average time until the process went out of control of somewhat less than 19.5, since the process could have been out of control for some time before detection occurred. For the case of no process improvement (ϕ =1), the prior mean of 0.05 suggests an average of 20 periods until an out-of-control condition occurs, so the posterior distribution on z, which includes the information contained in the data, is shifted somewhat to the right in this case. When ϕ =1, moments of the posterior can be obtained with a single-dimension numerical integration (see Silver and Bischak 2004) which in this case gives a posterior mean for z of 0.0514 > 0.05. (The estimate of the total rate z when ϕ =1 and the value obtained through numerical integration match for all the experiments).

[Insert table 2 about here]

Table 3 shows for each cause *j* the posterior parameters $\beta_j(5)$ and estimated expected values $\lambda_j(5)$ along with the standard error of the estimate across replications for ϕ=1.0, 0.8, and 0.5. The prior and estimated posterior Gamma distributions at ϕ=1.0 and 0.5 for each cause are also shown in figure 4. For these graphs we use $\beta_j(5)$ divided by the estimated posterior mean as an ad hoc estimate of $\gamma_j(5)$, since it is a function of the unknown out-of-control times.

[Insert table 3 about here]

[Insert figure 4 about here]

With no process improvement, the posterior means for the three observed causes (1, 3, and 4) are larger than the prior means. The increase is especially great for the previously unobserved cause 4, which had been assigned a low prior mean. At ϕ=0.8, however, cause 1's posterior mean has been reduced to less than its prior mean due to process improvement. For cause 3, the posterior is about the same as the prior mean – the increase due to its appearance in the data and the decrease due to subsequent process improvement just balance out. When ϕ=0.5, all posterior means are lower except for cause 4. Note that for the causes that were not seen in the four cycles, the posterior means will always be smaller than the prior means, since in this Bayesian analysis the fact that we have not seen these causes lowers our initial estimate of the rate at which they occur. The estimates of the total rate after process improvement when ϕ=0.8 and 0.5 drop to 0.0426 and 0.0311, respectively.

5.2. Changing the prior distributions

Two changes to the base case prior distribution on *z*, and hence to the prior distributions on the λ_j s, are illustrated in tables 4 and 5, respectively:

i) the prior mean is doubled, but the CV is the same ($\beta = 9$, $\gamma = 90$),

ii) the prior mean is the same, but the CV is reduced ($\beta = 25$, $\gamma = 500$).

The results for the first case (table 4) reveal the importance of the choice of the prior mean when there is so little data. The prior means are sufficiently large that all the posterior means (with the continued exception of cause 4) are actually smaller than their prior values for every ϕ.

[Insert table 4 about here]

In the second case (table 5), we see that the tighter prior distributions reduce the effect of the data on the posterior mean. When ϕ =1, the results for the observed causes are similar to those of the base case: the means adjust in the same direction as in the base case but not as far. When there is process improvement, this dampening of the movement away from the prior mean results in a post-improvement posterior mean that can be greater than or less than the prior mean but is always less than the base case posterior mean.

[Insert table 5 about here]

5.3. Changing the number of assignable causes

We next increase the number of assignable causes (*J*) from 10 to 17, allocating the percentage of out-of-control conditions due to unknown causes across 14 rather than seven causes. As displayed in table 6, the posterior result for a cause remains the same as in the base case if the prior mean remains the same, but lowering the prior mean for the previously unknown causes (as there are more of them) reduces the posterior mean for each of these causes. The overall rate remains essentially the same.

[Insert table 6 about here]

5.4. Changing the durations of the cycles

We next examine the effects of doubling the length of each cycle, $m(i)$, to 34, 20, 40, and 62, respectively. Longer cycles suggest a lower overall rate of occurrence of causes, and indeed the posterior means drop below the prior means (for all but cause 4), as displayed in table 7.

[Insert table 7 about here]

5.5. Increasing the times of the latest observed false alarms

Lastly, we increase the time of the latest observed false alarm in each cycle by five samples. Because it is quite likely that an out-of-control condition is detected quickly ($p = 0.59$ for the values of δ, n, and k we use here) and the base case r's and m's are not very close together, this has virtually no effect on the posterior means, as the results in table 8 indicate.

[Insert table 8 about here]

6. SUMMARY AND POSSIBLE EXTENSIONS

In this paper we have developed a Bayesian estimator of the total rate at which a process that is being monitored by an $\overline{X}$ chart is going out of control, as well as estimates of the rates at which individual assignable causes occur. Process improvement was modeled as being a fractional reduction in the future rate of occurrence of the cause identified in a cycle. The Markov chain Monte Carlo technique of data augmentation allowed us to make use of a simplified likelihood function despite not knowing the exact times at which the process went out of control.

One possible extension of this work would be to include the number of assignable causes as a parameter to be estimated, along with the individual rates. Another possibility would be to estimate the fraction reduction in rate *j*, ϕ_j, as a parameter, rather than assuming its value. Both of these extensions would add considerably to the complexity of the model.

Future work in this area would include the development of procedures to estimate other parameters that are of interest in control chart design, such as the size of the process shift when an out-of-control condition occurs. We are currently assessing the value of data augmentation and other Markov chain Monte Carlo techniques in addressing this estimation problem.

ACKNOWLEDGEMENTS

The research leading to this paper was supported by the Natural Sciences and Engineering Research Council of Canada under Grant Nos. A1485 and 239153-01 and by the Carma Chair at the University of Calgary. The authors wish to thank Charles Romeo for suggesting the use of the technique of data augmentation. The authors also thank Marc Maes and Haim Shore for their suggestions during the research.

REFERENCES

BISCHAK, D.P., and SILVER, E.A., 2001, Estimating the rate at which a process goes out of control in a statistical process control context. *International Journal of Production Research*, 39, 2957-2971.

CALABRESE, J. M., 1995, Bayesian process control for attributes. *Management Science*, 41, 637-645.

CARLIN, B.P., and LOUIS, T.A., 2000, *Bayes and Empirical Bayes Methods for Data Analysis* (Boca Raton, Florida: Chapman & Hall/CRC), 2nd edition.

CHIU, W.K., 1976, On the estimation of data parameters for economic optimum $\overline{X}$ charts. *Metrika*, 23, 135-147.

CHIU, W.K., and WETHERILL, G.B., 1974, A simplified scheme for the economic design of $\overline{X}$ charts. *Journal of Quality Technology*, 24, 63-69.

CHUNG, K.-J., 1994, An algorithm for computing the economically optimal $\overline{X}$ -control chart for a process with multiple assignable causes. *European Journal of Operational Research*, 72, 350-363.

DEGROOT, M.H., *Optimal Statistical Decisions*. McGraw-Hill Book Co., New York, NY, 1970.

DUNCAN, A.J., 1956, The economic design of $\overline{X}$ charts used to maintain current control of a process. *American Statistical Association Journal*, 51, 228-242.

DUNCAN, A.J., 1971, The economic design of $\overline{X}$ charts when there is a multiplicity of assignable causes. *Journal of the American Statistical Association*, 66, 107-121.

EVANS, M., HASTINGS, N., and PEACOCK, B., 2000, *Statistical Distributions*, 3rd edition (New York, NY: Wiley Interscience).

HAMADA, M., POHL, A., SPIEGELMAN, C., and WENDELBERGER, J., 2003, A Bayesian approach to calibration intervals and properly calibrated tolerance intervals. *Journal of Quality Technology*, 35, 194-205.

KEATS, J.B., DEL CASTILLO, E., VON COLLANI, E., and SANIGA, E.M., 1997, Economic modeling for statistical process control. *Journal of Quality Technology*, 29, 144-147.

LO, V.S.Y., 1994, On estimation methods of input parameters for economic attribute control charts. *The Statistician*, 43, 439-451.

LORENZEN, T.J., and VANCE, L.C., 1986, The economic design of control charts: a unified approach. *Technometrics*, 28, 3-10.

MENZEFRICKE, U., 2002, On the evaluation of control chart limits based on predictive distributions. *Communications in Statistics: Theory and Methods*, 31, 1423-1440.

NAIR, V.N., TANG, B., and XU, L.-A., 2001, Bayesian inference for some mixture problems in quality and reliability. *Journal of Quality Technology*, 33, 16-28.

PORTEUS, E.L., and ANGELUS, A., 1997, Opportunities for improved statistical process control. *Management Science*, 43, 1214-1228.

SANIGA, E.M., 1989, Economic statistical design of control charts with an application to *X*-bar and *R* charts. *Technometrics*, 31, 313-320.

SILVER, E.A., and BISCHAK, D.P., 2004, Bayesian estimation of the rate at which a process, monitored by an $\bar{X}$ chart, goes out of control. *International Journal of Production Research*, 42, 1227-1242.

SILVER, E.A., and ROHLEDER, T.R., 1999, The economic design of an $\bar{X}$ control chart recognizing process improvement. *International Journal of Production Research*, 37, 393-412.

TAGARAS, G., 1996, Dynamic control charts for finite production runs. *European Journal of Operational Research*, 91, 38-55.

TAGARAS, G., and LEE, H.L., 1988, Economic design of control charts with different control limits for different assignable causes. *Management Science*, 34, 1347-1366.

TAGARAS, G., and NIKOLAIDIS, Y., 2002, Comparing the effectiveness of various Bayesian $\overline{X}$ control charts. *Operations Research*, 50, 878-888.

TANNER, M.A., 1996, *Tools for statistical inference: methods for the exploration of posterior distributions and likelihood functions* (Springer-Verlag), 3rd edition.

TANNER, M.A., and WONG, W.H., 1987, The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82, 528-540.

WOODALL, W.H., 1986, Weaknesses of the economic design of control charts. *Technometrics,* 28, 408-410.

Table 1 – The base prior on the λ_j s

j	1	2	3	4	5	6	…	10	Total
q_j	0.3	0.2	0.15	0.05	0.05	0.05	…	0.05	1
β_j	2.7	1.8	1.35	0.45	0.45	0.45	…	0.45	9
γ_j	180	180	180	180	180	180	…	180	
E[λ_j]	0.015	0.01	0.0075	0.0025	0.0025	0.0025	…	0.0025	0.05

Table 2 – The base case observed data

i	$r(i)$	$m(i)$	$c(i)$
1	0	17	1
2	3	10	4
3	10	20	3
4	0	31	1

Table 3 – The posterior parameters $\beta_j(5)$ and estimated $E[\lambda_j(5)]$ for the base case

j	1	2	3	4	5	6	…	10	Total
Prior E[λ_j]	0.015	0.01	0.0075	0.0025	0.0025	0.0025	…	0.0025	0.05
ϕ =1:									
$\beta_j(5)$	4.7	1.8	2.35	1.45	0.45	0.45	…	0.45	13
Estimated E[$\lambda_j(5)$]	0.0186	0.0071	0.0093	0.0057	0.0018	0.0018	…	0.0018	0.0514
SE	0.00108	0.00016	0.00027	0.0001	0.00001	0.00001	…	0.00001	
ϕ =0.8:									
$\beta_j(5)$	3.008	1.8	1.88	1.16	0.45	0.45	…	0.45	10.548
Estimated E[$\lambda_j(5)$]	0.0125	0.0071	0.0076	0.0048	0.0018	0.0018	…	0.0018	0.0426
SE	0.00049	0.00016	0.00018	0.00007	0.00001	0.00001	…	0.00001	
ϕ =0.5:									
$\beta_j(5)$	1.175	1.8	1.175	0.725	0.45	0.45	…	0.45	7.575
Estimated E[$\lambda_j(5)$]	0.0052	0.0071	0.0049	0.0032	0.0018	0.0018	…	0.0018	0.0311
SE	0.00009	0.00016	0.00008	0.00003	0.00001	0.00001	…	0.00001	

Table 4 – The posterior parameters $\beta_j(5)$ and estimated $E[\lambda_j(5)]$ when the prior means are doubled

j	1	2	3	4	5	6	…	10	Total
Prior E[λ_j]	0.03	0.02	0.015	0.005	0.005	0.005	…	0.005	0.10
ϕ =1:									
$\beta_j(5)$	4.7	1.8	2.35	1.45	0.45	0.45	…	0.45	13
Estimated E[$\lambda_j(5)$]	0.0289	0.0111	0.0144	0.0089	0.0028	0.0028	…	0.0028	0.0798
SE	0.00261	0.00038	0.00065	0.00025	0.00002	0.00002	…	0.00002	
ϕ =0.8:									
$\beta_j(5)$	3.008	1.8	1.88	1.16	0.45	0.45	…	0.45	10.548
Estimated E[$\lambda_j(5)$]	0.0199	0.0111	0.0120	0.0076	0.0028	0.0028	…	0.0028	0.0670
SE	0.00124	0.00038	0.00045	0.00018	0.00002	0.00002	…	0.00002	
ϕ =0.5:									
$\beta_j(5)$	1.175	1.8	1.175	0.725	0.45	0.45	…	0.45	7.575
Estimated E[$\lambda_j(5)$]	0.0088	0.0111	0.0079	0.0052	0.0028	0.0028	…	0.0028	0.0495
SE	0.00024	0.00038	0.0002	0.00009	0.00002	0.00002	…	0.00002	

Table 5 – The posterior parameters $\beta_j(5)$ and estimated $E[\lambda_j(5)]$ when the prior CV is reduced

j	1	2	3	4	5	6	…	10	Total
Prior E[λ_j]	0.015	0.01	0.0075	0.0025	0.0025	0.0025	…	0.0025	0.05
ϕ =1:									
$\beta_j(5)$	9.5	5	4.75	2.25	1.25	1.25	…	1.25	29
Estimated E[$\lambda_j(5)$]	0.0166	0.0087	0.0083	0.0039	0.0022	0.0022	…	0.0022	0.0506
SE	0.00086	0.00024	0.00022	0.00005	0.00001	0.00001	…	0.00001	
ϕ =0.8:									
$\beta_j(5)$	6.08	5	3.8	1.8	1.25	1.25	…	1.25	24.18
Estimated E[$\lambda_j(5)$]	0.0108	0.0087	0.0067	0.0032	0.0022	0.0022	…	0.0022	0.0425
SE	0.00037	0.00024	0.00014	0.00003	0.00001	0.00001	…	0.00001	
ϕ =0.5:									
$\beta_j(5)$	2.375	5	2.375	1.125	1.25	1.25	…	1.25	18.375
Estimated E[$\lambda_j(5)$]	0.0044	0.0087	0.0043	0.0021	0.0022	0.0022	…	0.0022	0.0325
SE	0.00006	0.00024	0.00006	0.00001	0.00001	0.00001	…	0.00001	

Table 6 – The posterior parameters $\beta_j(5)$ and estimated $E[\lambda_j(5)]$ when the number of causes is increased

j	1	2	3	4	5	6	…	17	Total
Prior E[λ_j]	0.015	0.01	0.0075	0.00125	0.00125	0.00125	…	0.00125	0.05
ϕ =1:									
$\beta_j(5)$	4.7	1.8	2.35	1.225	0.225	0.225	…	0.225	13
Estimated E[$\lambda_j(5)$]	0.0186	0.0071	0.0093	0.0048	0.0009	0.0009	…	0.0009	0.0514
SE	0.00108	0.00016	0.00027	0.00007	2.5×10^{-6}	2.5×10^{-6}	…	2.5×10^{-6}	
ϕ =0.8:									
$\beta_j(5)$	3.008	1.8	1.88	0.98	0.225	0.225	…	0.225	10.593
Estimated E[$\lambda_j(5)$]	0.0125	0.0071	0.0076	0.0040	0.0009	0.0009	…	0.0009	0.0428
SE	0.00049	0.00016	0.00018	0.00005	2.5×10^{-6}	2.5×10^{-6}	…	2.5×10^{-6}	
ϕ =0.5:									
$\beta_j(5)$	1.175	1.8	1.175	0.6125	0.225	0.225	…	0.225	7.6875
Estimated E[$\lambda_j(5)$]	0.0052	0.0071	0.0049	0.0027	0.0009	0.0009	…	0.0009	0.0315
SE	0.00009	0.00016	0.00008	0.00002	2.5×10^{-6}	2.5×10^{-6}	…	2.5×10^{-6}	

Table 7 – The posterior parameters $\beta_j(5)$ and estimated $E[\lambda_j(5)]$ when the cycle lengths are doubled

j	1	2	3	4	5	6	…	10	Total
Prior E[λ_j]	0.015	0.01	0.0075	0.0025	0.0025	0.0025	…	0.0025	0.05
ϕ =1:									
$\beta_j(5)$	4.7	1.8	2.35	1.45	0.45	0.45	…	0.45	13
Estimated E[$\lambda_j(5)$]	0.0142	0.0054	0.0071	0.0044	0.0014	0.0014	…	0.0014	0.0393
SE	0.00063	0.00009	0.00016	0.00006	5.8×10^{-6}	5.8×10^{-6}	…	5.8×10^{-6}	
ϕ =0.8:									
$\beta_j(5)$	3.008	1.8	1.88	1.16	0.45	0.45	…	0.45	10.548
Estimated E[$\lambda_j(5)$]	0.0098	0.0054	0.0059	0.0037	0.0014	0.0014	…	0.0014	0.0330
SE	0.0003	0.00009	0.00011	0.00004	5.8×10^{-6}	5.8×10^{-6}	…	5.8×10^{-6}	
ϕ =0.5:									
$\beta_j(5)$	1.175	1.8	1.175	0.725	0.45	0.45	…	0.45	7.575
Estimated E[$\lambda_j(5)$]	0.0043	0.0054	0.0039	0.0026	0.0014	0.0014	…	0.0014	0.0244
SE	0.00006	0.00009	0.00005	0.00002	5.8×10^{-6}	5.8×10^{-6}	…	5.8×10^{-6}	

Table 8 – The posterior parameters $\beta_j(5)$ and estimated $E[\lambda_j(5)]$ when the times until the last false alarm are increased

j	1	2	3	4	5	6	…	10	Total
Prior $E[\lambda_j]$	0.015	0.01	0.0075	0.0025	0.0025	0.0025	…	0.0025	0.05
$\phi=1$:									
$\beta_j(5)$	4.7	1.8	2.35	1.45	0.45	0.45	…	0.45	13
Estimated $E[\lambda_j(5)]$	0.0185	0.0071	0.0093	0.0057	0.0018	0.0018	…	0.0018	0.0513
SE	0.00108	0.00016	0.00027	0.0001	0.00001	0.00001	…	0.00001	
$\phi=0.8$:									
$\beta_j(5)$	3.008	1.8	1.88	1.16	0.45	0.45	…	0.45	10.548
Estimated $E[\lambda_j(5)]$	0.0124	0.0071	0.0076	0.0048	0.0018	0.0018	…	0.0018	0.0425
SE	0.00048	0.00016	0.00018	0.00007	0.00001	0.00001	…	0.00001	
$\phi=0.5$:									
$\beta_j(5)$	1.175	1.8	1.175	0.725	0.45	0.45	…	0.45	7.575
Estimated $E[\lambda_j(5)]$	0.0052	0.0071	0.0049	0.0032	0.0018	0.0018	…	0.0018	0.0310
SE	0.00009	0.00016	0.00008	0.00003	0.00001	0.00001	…	0.00001	

Figure 1: An example of a cycle in which $r = 2$, $i_0 = 3$, and $m = 4$

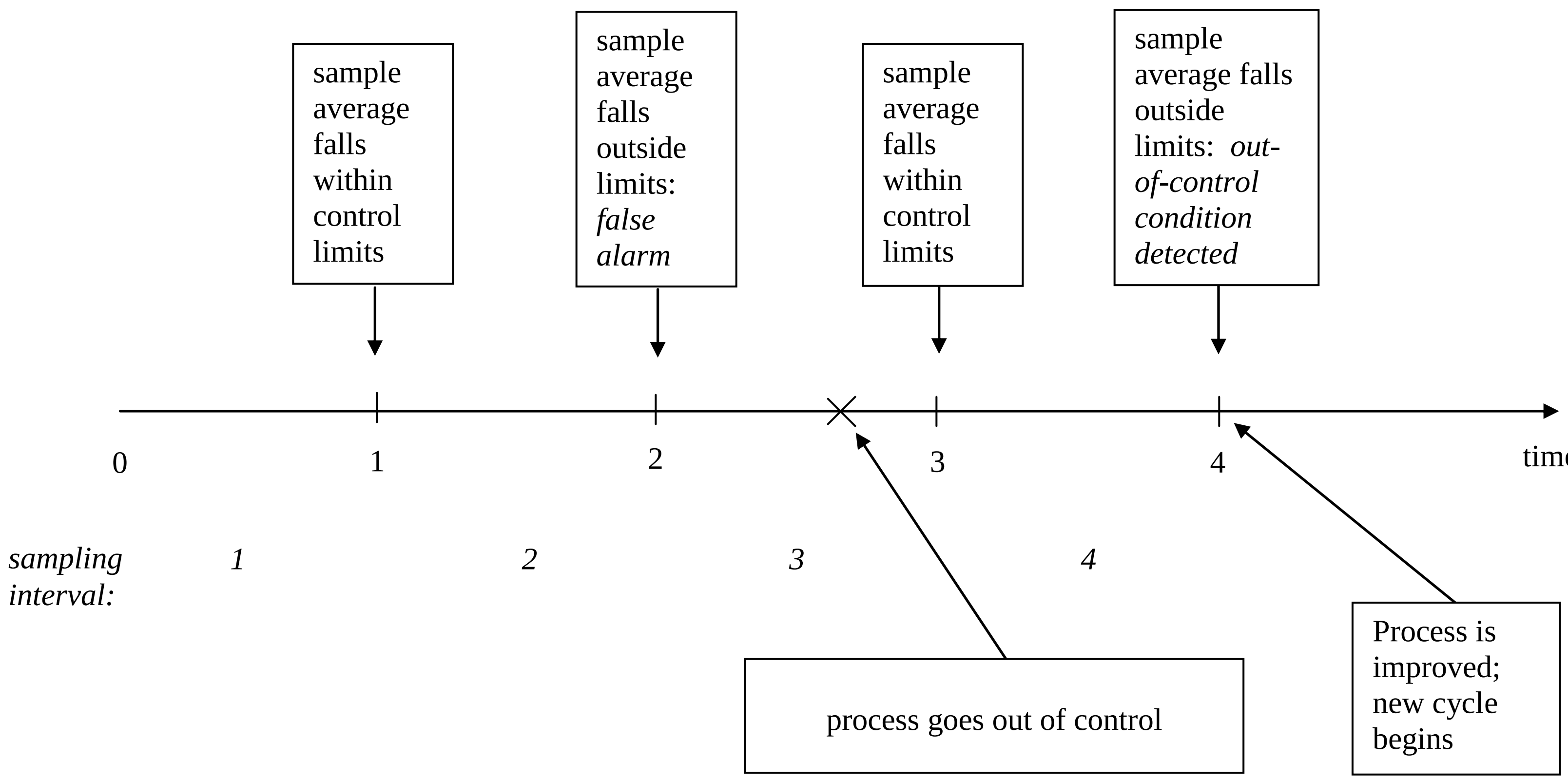


Figure 2: The data augmentation approach

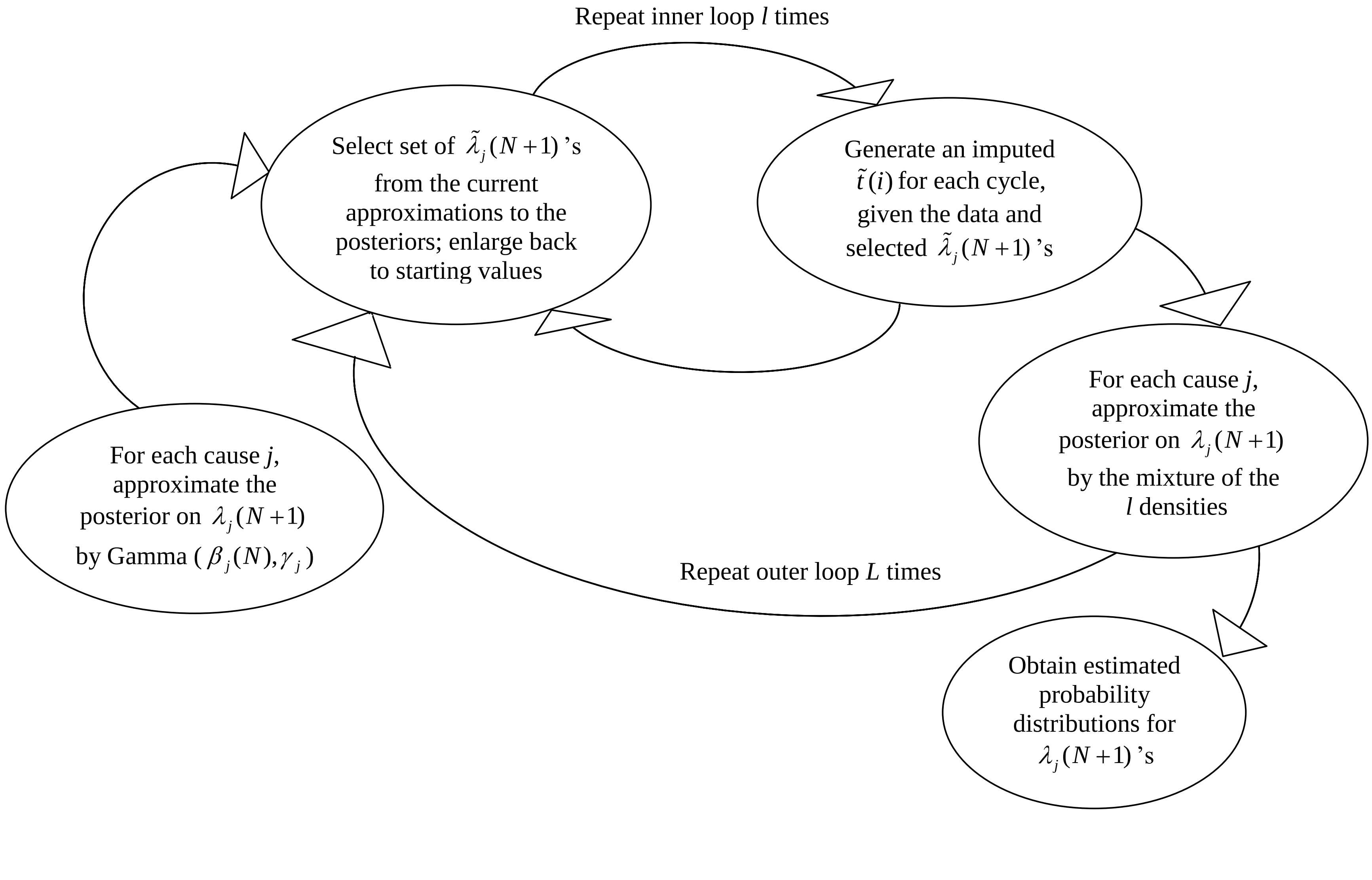


Figure 3: estimates of E[z] for the base case, $\phi = 1.0$, using ten replications with 30 outer loop iterations each

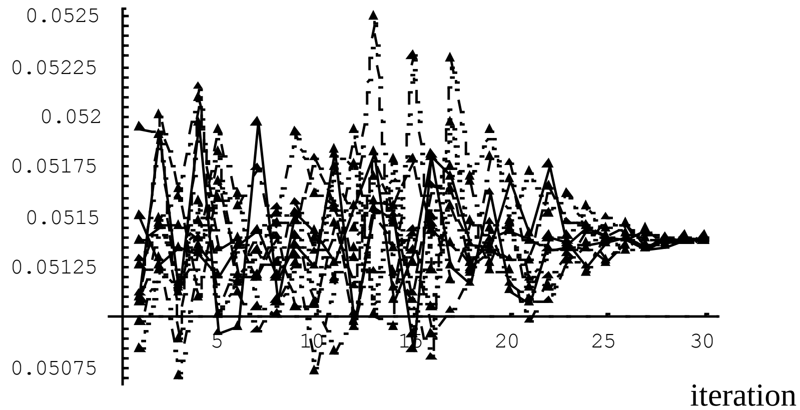
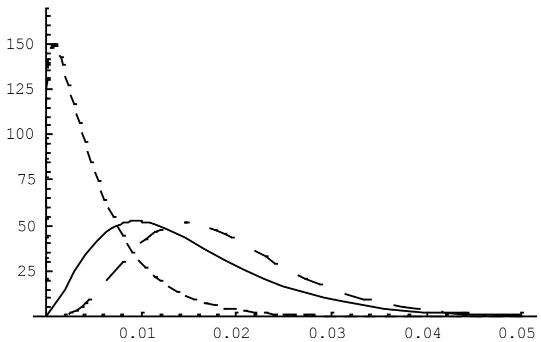
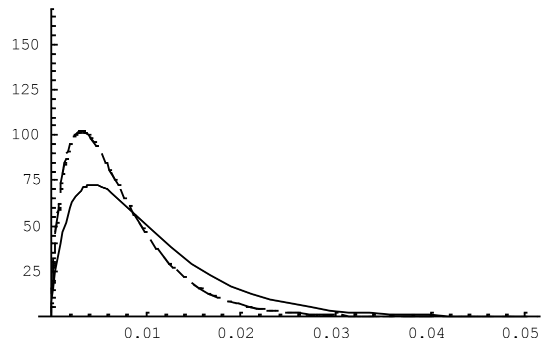


Figure 4: Prior and posterior distributions on individual causes, base case

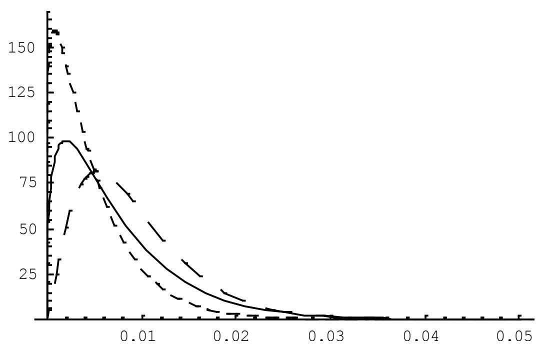
Solid line: prior distribution; long dashed line: posterior, ϕ=1.0; short dashed line: posterior, ϕ=0.5



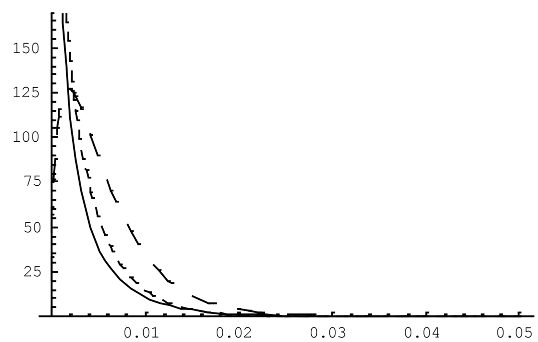
Cause 1



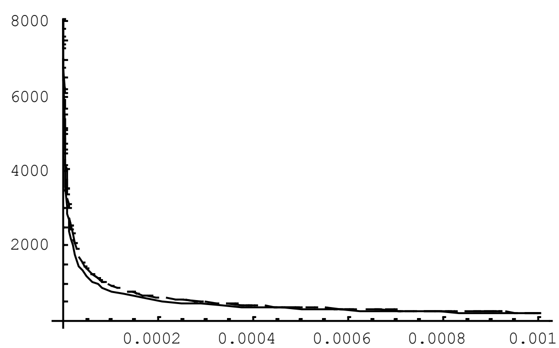
Cause 2



Cause 3



Cause 4



Causes 5 through 10