Bayesian and Minimax Solutions to the Adaptation of a Plant to a Changing Environment

S. Merad*
Centre for Software Reliability, Department of Computing Science, University of Newcastle, Newcastle upon Tyne, UK
email: Salah.Merad@newcastle.ac.uk

Abstract

In this paper, we consider the decision problem that an automated control system faces when its plant has to adapt to a changing environment which can be in a finite number of states. The law of transition between states is unknown, and we hence adopt the worst case scenario. The controller gathers data at regular intervals and, periodically, it has to select the best control for the next period. The optimality criterion used in the selection process is the minimum expected loss in utility from not applying the best control. We construct a model and propose solution procedures for both stable and unstable environments. It is shown that when the environment is stable, the decision problem is equivalent to a classification problem. When it is unstable, we formulate the problem as a two-person zero-sum game. In both cases, the minimax principle plays a major role.

Keywords Bayesian change point, Classification, Minimax criterion, Two-person zero-sum game, Utility loss.

1 Introduction

In this paper, we consider a statistical decision problem in an adaptive control system composed of a controller and a plant, and operating in a changing environment. The controller makes regular readings from the plant through its sensors. The readings collected during a period give information about the state of the external environment which can be in a finite number of states. We assume that the transition law between the states is unknown. The readings made in the same environmental state are modelled using independent identically distributed random variables. We assume that the different probability distributions, which correspond to the different states,

*Supported by an EPSRC grant.

belong to the same parametric family, and hence they differ only in the values of their parameters.

The controller has a number of controls available to choose from. Every control is best suited to one of the states, but it may also be used in other states. The performance and suitability of the controls in the various environmental states is quantified using a utility function. Periodically, and on the basis of the information contained in the readings, the controller has to select the best control to apply in the next period. We hence have a decision problem with observations [5]. The optimality criterion used in the selection process is the minimum expected loss in utility from not applying the best control, regardless of the transition state.

We consider two extreme cases: the first case is when the environment is very stable so that changes occur rarely, and the second case is when the environment is very unstable so that it changes in almost every sampling period. In the former, the problem reduces to an identification or classification problem, whereas the latter is formulated as a two-person zero-sum game. In both cases, the solution is based on the minimax approach. We assume that the a priori distribution of the time of change is uniform, but the transition state when change occurs is chosen according with the worst case scenario.

An example of an adaptive control system is the automated suspension of a vehicle considered in [15]. A controller monitors the height of the suspension of a vehicle operating on many types of road, for example, motorway, gravel, and off-road. Regular readings of the fluctuations of the height of the chassis around the selected set point are made. These fluctuations, which result from the road surface, are assumed to follow a Normal probability distribution with mean 0 and a variance which is specific to the road type. And, depending on the available information, the controller must choose the appropriate set point for the chassis in the next period. This set point can be either low (best for a motorway), medium (best for a gravelled road) or high (best for off-road).

The problem of change-detection in a dynamic environment has been extensively studied in many areas of engineering, for example, in monitoring communication channels [7], failure detection [17], and search radar [6]. The goal of the change-detection in theses problems is to sequentially observe a sequence and raise an alarm as soon as a change in distribution has been detected, subject to some type of false alarm constraint. The main criteria used in the cited contributions are the Shyryayev's criterion [13] in which the objective is to minimise the expected delay under a constraint on the probability of stopping before change occurs (false alarm), and Lorden's criterion [9] in which the objective is to minimise an upper bound on the expected delay over all possible change times, and it is hence a minimax criterion. In [8], the Shyryayev's approach was extended to the case where an a priori distribution of change time is not available. Bayesian procedures have also been used in the generic problem of the detection of a multiple abrupt changes in the parameters of piecewise stationary univariate time series [11]. In all the above papers, the decision making is over a single period, whereas in our contribution it is over many periods.

In section 2, we present the model together with the simplifying assumptions we make. In section 3, we introduce the Bayesian change-point statistical test in a sampling period. In section 4, we consider the problem of determining the maximin prior of the transition state, and illustrate the method with an example where measurements are independent Normal. In section 5, we consider the optimal selection problem when the environment is stable and unstable, in turn. In section 6, we give some concluding remarks and suggest possible lines of exploration for future work.

2 The Model

Let θ denote the state of the external environment. It is assumed that the environment can be in M possible states denoted by $\theta_1, \dots, \theta_M$. Let Θ denote the set of the environments possible states. We also assume that the controller can choose from the finite set of controls $A = \{a_1, a_2, \dots, a_M\}$. For each state of the environment θ_j, $j = 1, 2, \dots, M$, we associate a utility U_{ij} to every control a_i, $i = 1, 2, \dots, M$. The utilities are assigned using the method of von Neumann [16]- see also [10] for a more detailed account. The controls are indexed such that when the environment is in state θ_j, a_j is the best control to apply. Let $x_1, x_2, \dots, x_N$ be N readings from the plant made at regular intervals during a period of length N. We assume that $x_n \in S$ for all $n = 1, 2, \dots, N$, where the set S can be either a real interval, or a set of integers, or the set $\{0, 1\}$, depending on the environmental variable being measured.

We make the following assumption concerning the collected data: The readings $x_1, x_2, \dots, x_N$ are the observed values of the independent random variables $X_1, X_2, \dots, X_N$. For simplicity of notation, we identify the state of the environment with the parameter of the distribution so that when the environment is in state θ_j, the readings have probability density function $g(.|\theta_j)$. The probability density functions are hence assumed to be from the same parametric family for all states.
Let $T_1, T_2, \dots, T_n, \dots$ be the times at which the computations necessary for the decision making are made. These times are such that

$$T_{n+1} - T_n = N, \quad \text{forall } n \geq 1.$$

The choice of sampling rate and sample size are parallel and interlinked problems [2]. If the system structure is known, then optimal measurement and sampling can be achieved with the use of Fisher's information matrix [12].

For simplicity, we assume that the time required to reach a decision at time T_n is small compared with the the length of a period, and hence it is neglected. If the controller finds that action a_i is the best, this action will be applied between times T_n and T_{n+1}.

Let γ be the probability of a change in environmental state during a period. For simplicity, we assume that there is at most one change during any period.

The controller has to use the information in the data to update its knowledge about the state of the environment at time T_{n+1}. But because the state of the environment can change during the sampling period, the updating process has to integrate this possibility of change, and hence we need to work in the framework of a change-point statistical problem.

Note that if all the readings are from the same environmental state, then the joint density function is given by

$$f_j(\mathbf{x}) = \prod_{n=1}^{N} g(x_n;\theta_j),$$

where $\mathbf{x} = (x_1, x_2, \cdots, x_N)$.

3 A change-point statistical problem

Consider the period $[T_n, T_{n+1}]$, and let $p_\theta^{(0)}(\theta_k)$ be the probability that the environment is in state θ_k at the beginning of this period. We want to find the posterior probability distribution of the state of the environment, given the prior $p_\theta^{(0)}(\theta_k)$, the collected data $\mathbf{x} = (x_1, x_2, \cdots, x_N)$, and the probability of change γ.

Let t_c, $1 \le t_c \le N$ be the time at which a change of environmental state occurs in a period. The case of $t_c = N$ corresponds to the situation of no change in the period.

The statistical change point problem has been studied extensively in the literature. See, for example, [4] for an account of both classical parametric and non-parametric methods with emphasis on limit theorems, and [14] for a Bayesian account. [18] gives a survey of both classical and Bayesian approaches to the change point problem.

In our problem, the decision making is periodic and data are collected throughout. Hence, to make full use of the information in the data, we will adopt the Bayesian approach. Let $\theta^{(1)}$ be the state of the environment in the time interval $[T_n, T_n + t_c)$ and $\theta^{(2)}$ be the state of the environment in the time interval $[T_n + t_c, T_{n+1}]$. Let $q^{(0)}(t_c)$ be a prior over the set of possible change points, where $q^{(0)}(N)$ is the probability of no change in the period.
We assume that the prior distribution of t_c is uniform over $\{1, 2, \ldots, N-1\}$ so that

$$q^{(0)}(t_c) = \gamma/(N-1), \quad \text{for} \quad t_c = 1, 2, \cdots, N-1,$$

and

$$q^{(0)}(N) = 1 - \gamma.$$

Let $p^{(0)}_{\theta^{(1)},\theta^{(2)}|t_c}(\theta_k, \theta_j|t_c)$ be the joint prior probability distribution of $\theta^{(1)}$, and $\theta^{(2)}$ given the change of state t_c time units after the beginning of the period. We then have

$$p^{(0)}_{\theta^{(1)},\theta^{(2)}|t_c}(\theta_j, \theta_j|t_c) = p^{(0)}_{\theta^{(2)}|\theta^{(1)},t_c}(\theta_j|\theta_k, t_c)\, p^{(0)}_{\theta^{(1)}}(\theta_k) \tag{1}$$

Note that $p^{(0)}_{\theta^{(1)}}(\theta_k) = p^{(0)}_\theta(\theta_k)$.
We have assumed that there is no information about the transitions between states when change occurs. The prior probability distribution of $\theta^{(2)}$, $p^{(0)}_{\theta^{(2)}|\theta^{(1)},t_c}(\theta_j|\theta_k, t_c)$, is hence chosen in such a way that it is the least favourable distribution with respect to some utility loss. We treat the problem of finding this distribution in the next

section.
Let $q^{(N)}(t_c)$ be the posterior distribution of the time of change t_c. Then

$$q^{(N)}(t_c) \propto f(x_1, x_2, \ldots, x_N|t_c)q^{(0)}(t_c), \tag{2}$$

where

$$f(x_1,x_2,\ldots,x_N|t_c) = \sum_{1\le k\le M}\sum_{\substack{1\le k\le M\\ k\ne j}} f(x_1,x_2,\ldots,x_N|t_c,\theta_k,\theta_j)p^{(0)}_{\theta^{(1)},\theta^{(2)}|t_n}(\theta_k,\theta_j|t_c). \tag{3}$$

Moreover, the joint posterior distribution of $\theta^{(1)}$ and $\theta^{(2)}$ is given by

$$p^{(N)}_{\theta^{(1)},\theta^{(2)}}(\theta_k,\theta_j) = \sum_{t_n=1}^{N-1} p^{(N)}_{\theta^{(1)},\theta^{(2)}|t_n}(\theta_k,\theta_j|t_c)q^{(N)}(t_c), \tag{4}$$

where

$$p^{(N)}_{\theta^{(1)},\theta^{(2)}|t_n}(\theta_k,\theta_j|t_c) \propto f(x_1,x_2,\ldots,x_N|t_c,\theta_k,\theta_j)p^{(0)}_{\theta^{(1)},\theta^{(2)}|t_n}(\theta_k,\theta_j|t_c). \tag{5}$$

Finally, the posterior distribution of the environmental state at time T_{n+1} is given by

$$p^{(N)}_{\theta^{(2)}}(\theta_j) = \sum_{\substack{1\le k\le M\\ k\ne j}} p^{(N)}_{\theta^{(1)},\theta^{(2)}}(\theta_k,\theta_j). \tag{6}$$

4 Maximin prior

We now consider the problem of finding the prior probability distribution of the state of the environment after a change of state, that is, $p^{(0)}_{\theta^{(2)}|\theta^{(1)},t_n}(\theta_j|\theta_k,t_c)$. This prior is chosen in such a way that it is the least favourable with respect to identifying the true state. Suppose that $\theta^{(1)} = \theta_k$, $(1 \le k \le M)$. We then define the following loss function

$$L^{(-k)}(i|j) = \begin{cases} 0 & \text{for } i = j \\ 1 & \text{for } i \ne j \end{cases} \tag{7}$$

for $i, j \in \{1, 2, \ldots, M\} - \{k\}$.

Minimax classification Let $\mathbf{x}_{t_n} = (x_{t_n}, x_{t_n+1}, \ldots, x_N)$ represent the readings collected in the new state. We want to tell from which environmental state the sample comes from. Hence we want to classify the sample $\mathbf{x}_{t_n}$ into one of the environmental states in $\Theta^{(-k)} = \Theta - \{\theta_k\}$. Such classification or identification will be subject to errors. We therefore need a classification that is best according to some well specified criterion. The criterion we use is the minimum expected utility loss with respect to the function $L^{(-k)}(i|j)$. For this classification problem, we adapt the account in [1] using the notation which is appropriate to our problem.

We want to divide the space of observations given by

$$\Omega_{t_n} = \{\mathbf{x}_{t_n} = (x_{t_n}, \dots, x_N)\,, x_t \in S, t_n \le t \le N\}$$

into $M-1$ mutually exclusive and exhaustive regions $D_1^{(-k)}(t_n), D_2^{(-k)}(t_n), \dots, D_{k-1}^{(-k)}(t_n)$, $D_{k+1}^{(-k)}(t_n), \dots, D_M^{(-k)}(t_n)$. That is,

$$\Omega_{t_n} = \bigcup_{\substack{1 \le i \le M \\ i \ne k}} D_i^{(-k)}(t_n),$$

and

$$D_i^{(-k)}(t_n) \cup D_j^{(-k)}(t_n) = \emptyset \quad \text{for } i \ne j.$$

We denote this classification by $\mathfrak{D}_{t_n}^{(-k)}$. Then, if a sample falls into region $D_i^{(-k)}(t_n)$, we shall identify the state as θ_i.

Under classification rule $\mathfrak{D}_{t_n}^{(-k)}$, the probability of classifying sample $\mathbf{x}_{t_n}$ into environmental state θ_i when it in fact comes from environmental state θ_j is given by

$$P(i|j, \mathfrak{D}_{t_n}^{(-k)}) = \int_{D_i^{(-k)}(t_n)} f_j(\mathbf{x}_{t_n}) d\mathbf{x}_{t_n}, \tag{8}$$

where $d\mathbf{x}_{t_n} = dx_{t_n} dx_{t_n+1} \cdots dx_N$, and $i \ne j$.
Let $\Delta_{t_n}^{(-k)}$ be the set of all classification rules. The expected loss incurred, also called risk, under strategy $\mathfrak{D}_{t_n}^{(-k)} \in \Delta_{t_n}^{(-k)}$ when the sample comes from an environment in state θ_j is

$$R(j, \mathfrak{D}_{t_n}^{(-k)}) = \sum_{i \ne j} L(i|j) P(i|j, \mathfrak{D}_{t_n}^{(-k)}). \tag{9}$$

We next give the definitions of an admissible rule and a Bayes rule.

Definition 4.1 (Admissible rule). *A procedure $\mathfrak{D}$ is at least as good as procedure $\mathfrak{D}'$ if for all j,*

$$R(j, \mathfrak{D}) \le R(j, \mathfrak{D}'),$$

and is better than $\mathfrak{D}'$ if at least one of these inequalities is a strict inequality. A procedure $\mathfrak{D}$ is said to be admissible if there is no better procedure.

We assume that the state of the environment before the change occurs is θ_k. For $1 \le j \le M, j \ne k$, let

$$p_j^{(-k)} = p^{(0)}_{\theta^{(2)}|\theta^{(1)}, t_n}(\theta_j | \theta_k, t_n),$$

and

$$\mathbf{p}_{t_n}^{(-k)} = (p_1^{(-k)}, p_2^{(-k)}, \dots, p_{k-1}^{(-k)}, p_{k+1}^{(-k)}, \dots, p_M^{(-k)}).$$

Note that we simplified the notation by adopting $p_j^{(-k)}$ instead of $p_j^{(-k)}(t_n)$. We also denote the set of all priors over $\Theta^{(-k)}$ by $\mathcal{P}(\Theta^{(-k)})$.

Definition 4.2 (Bayes rule). *A rule* $\mathcal{D}^{\mathbf{p}_{k_n}^{(-k)}}$ *is said to be a Bayes rule with respect to the prior* $\mathbf{p}_{k_n}^{(-k)} \in \mathcal{P}(\Theta^{(-k)})$ *if*

$$\sum_{\substack{1 \le j \le M \\ j \ne k}} p_j^{(-k)} R(j, \mathcal{D}^{\mathbf{p}_{k_n}^{(-k)}}) \le \sum_{\substack{1 \le j \le M \\ j \ne k}} p_j^{(-k)} R(j, \mathcal{D}) \quad \forall \mathcal{D} \in \Delta_{k_n}^{(-k)}.$$

Note that it can be shown that a Bayes rule is admissible [3]. Now, we want to find a rule $\mathcal{D}_{k_n}^{(-k)*}$ that is optimal regardless of the new state of the environment. It can be shown that this rule can be found by solving the two-person zero-sum game between the controller which wants to minimise its expected utility loss and the environment which adopts a distribution $\mathbf{p}_{k_n}^{(-k)*}$ so as to maximise the controller's loss. We have

$$\begin{aligned} \sum_{\substack{1 \le j \le M \\ j \ne k}} p_j^{(-k)*} R(j, \mathcal{D}_{k_n}^{(-k)*}) &= \inf_{\mathcal{D}} \sup_{\mathbf{p}_{k_n}^{(-k)}} \sum_{\substack{1 \le j \le M \\ j \ne k}} p_j^{(-k)} R(j, \mathcal{D}) \\ &= \sup_{\mathbf{p}_{k_n}^{(-k)}} \inf_{\mathcal{D}} \sum_{\substack{1 \le j \le M \\ j \ne k}} p_j^{(-k)} R(j, \mathcal{D}). \end{aligned} \tag{10}$$

This problem is called a statistical game, $\mathcal{D}_{k_n}^{(-k)*}$ is called a minimax rule, and distribution $\mathbf{p}_{k_n}^{(-k)*}$ is the least favourable a priori distribution or maximin prior. To find the minimax rule and the associated maximin prior, we use the following result.

Theorem 4.1. *If* $\mathcal{D}_{k_n}^{(-k)*} \in \Delta_{k_n}^{(-k)}$ *is Bayes with respect to* $\mathbf{p}_{k_n}^{(-k)*} \in \mathcal{P}(\Theta^{(-k)})$*, and*

$$R(j, \mathcal{D}_{k_n}^{(-k)*}) \le E_{\mathbf{p}_{k_n}^{(-k)*}}[R(i, \mathcal{D}_{k_n}^{(-k)*})]$$

for all $(1 \le j \le M, j \ne k)$*, then* $\mathcal{D}_{k_n}^{(-k)*}$ *is minimax and* $\mathbf{p}_{k_n}^{(-k)*}$ *is maximin.*

Proof. See [3], p 211. □

The following procedure, which is a corollary of Theorem 4.1, allows us to find the minimax regions and the maximin prior:

1. Find the form of the Bayes rule with respect to a prior $\mathbf{p}_{k_n}^{(-k)}$.
2. Equalise the risk functions; that is, set

$$\begin{aligned} R(\theta_1, \mathcal{D}_{k_n}^{(-k)*}) = R(\theta_2, \mathcal{D}_{k_n}^{(-k)*}) = \cdots = R(\theta_{k-1}, \mathcal{D}_{k_n}^{(-k)*}) \\ = R(\theta_{k+1}, \mathcal{D}_{k_n}^{(-k)*}) = \cdots = R(\theta_M, \mathcal{D}_{k_n}^{(-k)*}), \end{aligned} \tag{11}$$

   and solve to find the minimax regions.
3. If the Bayes regions in step 1 are given explicitly in terms of the prior $\mathbf{p}_{k_n}^{(-k)}$, then, using results from step 2, we can solve the appropriate equations to find $\mathbf{p}_{k_n}^{(-k)*}$.

Note that this procedure is carried out entirely off-line, and all the results stored. The above procedure has to be repeated for all possible values of the time of change, and for all state spaces $\Theta^{(-k)}, 1 \leq k \leq M$; that is, the procedure has to be repeated $(N-1)M$ times. We now illustrate the procedure for Normally distributed measurements.

Example Assume that when a system operates in an environment in state θ_j, the resulting measurements are normally distributed with mean 0 and variance σ_j^2. In the automated suspension example, the measurements are the difference between the height of the chassis and the selected set point. Assume that the environmental states are indexed so that

$$\sigma_1 < \sigma_2 < \cdots < \sigma_M.$$

Given the simple form of the loss function given by Equation 7, it is easy to verify that the Bayes rule $\mathcal{D}^{\mathbf{p}_{t_n}^{(-k)}}$ with respect to the prior $\mathbf{p}_{t_n}^{(-k)}$ is given by the sets $D_i^{\mathbf{p}_{t_n}^{(-k)}}$, $(1 \leq i \leq M, i \neq k)$, where

$$D_i^{\mathbf{p}_{t_n}^{(-k)}} = \left\{ \mathbf{x}_{t_n} : p_i^{(-k)} f(\mathbf{x}_{t_n}|\theta_i) > p_l^{(-k)} f(\mathbf{x}_{t_n}|\theta_l), (1 \leq l \leq M, l \neq i, l \neq k) \right\}.$$

Now, for $i > l$, $\sigma_i > \sigma_l$, and it is easy to show that

$$p_i^{(-k)} f(\mathbf{x}|\theta_i) > p_l^{(-k)} f(\mathbf{x}|\theta_l)$$

if and only if

$$\sum_{n=t_n}^{N} (x_n - \mu)^2 > \frac{2}{\frac{1}{\sigma_l^2} - \frac{1}{\sigma_i^2}} \ln \left(\left(\frac{\sigma_i}{\sigma_l} \right)^{N-t_n} \frac{p_l^{(-k)}}{p_i^{(-k)}} \right) = \eta_{il} \tag{12}$$

Similarly, for $i < l$, $\sigma_i < \sigma_l$, and we can verify that

$$p_i^{(-k)} f(\mathbf{x}|\theta_i) > p_l^{(-k)} f(\mathbf{x}|\theta_l)$$

if and only if

$$\sum_{n=t_n}^{N} (x_n - \mu)^2 < \frac{2}{\frac{1}{\sigma_i^2} - \frac{1}{\sigma_l^2}} \ln \left(\left(\frac{\sigma_l}{\sigma_i} \right)^{N-t_n} \frac{p_i^{(-k)}}{p_l^{(-k)}} \right) = \eta_{li} \tag{13}$$

It can be easily shown that the inequalities in Equation (12) and Equation 13 imply that there exists constants $\lambda_0, \lambda_1, \cdots, \lambda_{k-1}, \lambda_{k+1}, \cdots, \lambda_M$ indexed so that

$$0 = \lambda_0 \leq \lambda_1 \leq \ldots \leq \lambda_{k-1} \leq \lambda_{k+1} \leq \ldots \leq \lambda_M = \infty,$$

such that

$$D_i^{\mathbf{p}_{t_n}^{(-k)}} = \left\{ \mathbf{x}_{t_n} : \lambda_{i-1} \leq \sum_{n=t_n}^{N} (x_n - \mu)^2 \leq \lambda_i \right\}$$

for $(1 \leq i \leq k-1)$ and $(k+1 < i \leq M-1)$, and

$$D_{k+1}^{\mathbf{D}_{t_c}^{(-n)}} = \left\{ \mathbf{x}_{t_c} : \lambda_{k-1} \leq \sum_{n=t_c}^{N} (x_n - \mu)^2 \leq \lambda_{k+1} \right\}.$$

Now, the risk function is given by

$$\begin{aligned} R(\theta_j, \mathbf{D}_{t_c}^{(-n)}) &= \sum_{\substack{1 \leq i \leq M-1 \\ i \neq k}} L^{(-k)}(i|j) P(a_i|\theta_j, \mathbf{D}_{t_c}^{(-n)}) \\ &= 1 - P(a_j|\theta_j, \mathbf{D}_{t_c}^{(-n)}) \\ &= 1 - P\left(\lambda_{j-1} \leq \sum_{n=t_c}^{N} (X_n - \mu)^2 \leq \lambda_j | X_n \sim N(\mu, \sigma_j^2) \right) \\ &= 1 - P\left(\frac{\lambda_{j-1}}{\sigma_j^2} \leq Y \leq \frac{\lambda_j}{\sigma_j^2} \right), \end{aligned}$$

where Y has a χ^2 distribution with $N - t_c + 1$ degrees of freedom. To find the constants, we equalise the risks and obtain

$$\begin{aligned} & P\left(\frac{\lambda_0}{\sigma_1^2} \leq Y \leq \frac{\lambda_1}{\sigma_1^2} \right) = P\left(\frac{\lambda_1}{\sigma_2^2} \leq Y \leq \frac{\lambda_2}{\sigma_2^2} \right) = \cdots \\ &= P\left(\frac{\lambda_{k-2}}{\sigma_{k-1}^2} \leq Y \leq \frac{\lambda_{k-1}}{\sigma_{k-1}^2} \right) = P\left(\frac{\lambda_{k-1}}{\sigma_{k+1}^2} \leq Y \leq \frac{\lambda_{k+1}}{\sigma_{k+1}^2} \right) \\ &= P\left(\frac{\lambda_{k+1}}{\sigma_{k+2}^2} \leq Y \leq \frac{\lambda_{k+2}}{\sigma_{k+2}^2} \right) = \cdots = P\left(\frac{\lambda_{M-1}}{\sigma_M^2} \leq Y \leq \frac{\lambda_M}{\sigma_M^2} \right) = c. \end{aligned} \tag{14}$$

We can see that, for $1 \leq j \leq k-1$ and $k+2 \leq j \leq M-2$, once c and λ_{j-1} are known, we can determine λ_j easily trough the equation

$$P\left(Y \leq \frac{\lambda_j}{\sigma_j^2} \right) = P\left(Y \leq \frac{\lambda_{j-1}}{\sigma_j^2} \right) + c, \tag{15}$$

and once λ_{k-1} is known, λ_{k+1} can be determined through the equation

$$P\left(Y \leq \frac{\lambda_{k+1}}{\sigma_{k+1}^2} \right) = P\left(Y \leq \frac{\lambda_{k-1}}{\sigma_{k+1}^2} \right) + c. \tag{16}$$

On the other hand, λ_1 and λ_{M-1} verify respectively

$$P\left(Y \leq \frac{\lambda_1}{\sigma_1^2} \right) = c \tag{17}$$

and

$$P\left(Y \leq \frac{\lambda_{M-1}}{\sigma_M^2} \right) = 1 - c. \tag{18}$$

We then have $M-1$ equations with $M-1$ unknowns, $\{\lambda_i, 1 \leq i \leq M-1, i \neq k\}$ and c. However, these equations are not linear and cannot be solved directly. We propose the following iterative method:

1. Set $n = 1$ and $c = c_0(n) = 0.5$.

2. Find λ_1 and λ_{M-1} by solving Equations 17 and 18, respectively, and λ_j, $2 \leq j \leq M-2, j \neq k$ using equations 15 and 16.

3. Compute

$$c_1(n) = P\left(Y \leq \frac{\lambda_{M-1}}{\sigma^2_{M-1}}\right) - P\left(Y \leq \frac{\lambda_{M-2}}{\sigma^2_{M-1}}\right).$$

4. If $|c_0(n) - c_1(n)| \leq \delta$, where δ is a small positive real number, say 10^{-2}, stop. Otherwise, execute iteration $n+1$ by setting

$$c = c_0(n+1) = \frac{c_0(n) + c_1(n)}{2},$$

and moving to step 2.

To illustrate the above algorithm, we carry out the computations with $M = 4, \sigma_1 = 0.02, \sigma_2 = 0.10, \sigma_3 = 0.15, \sigma_4 = 0.20$, and $\delta = 0.01$. On the other hand, N and l_c are such that $N - l_c + 1 = 10$, so that $Y \sim \chi^2_{10}$. We assume that $\theta^{(1)} = \theta_4$. In this example, the minimax regions are found after three iterations.

Iteration 1.
$c_0(1) = 0.5$; $\lambda_1 = 15.20, \lambda_2 = 41.20$; $c_1(1) = 0.63$. $|c_1(1) - c_0(1)| = 0.13 > \delta$.
Set $c_0(2) = \frac{0.50+0.63}{2} = 0.57$.

Iteration 2.
$c_0(2) = 0.57$; $\lambda_1 = 14.57, \lambda_2 = 37.36$; $c_1(2) = 0.65$. $|c_1(2) - c_0(2)| = 0.07 > \delta$.
Set $c_0(3) = \frac{0.57+0.65}{2} = 0.61$.

Iteration 3.
$c_0(3) = 0.61$; $\lambda_1 = 15.26, \lambda_2 = 36.13$; $c_1(3) = 0.60$. $|c_1(3) - c_0(3)| = 0.01 = \delta$.

Once the minimax regions have been determined, the least favourable prior $\{p_j^{(4)*}, 1 \leq j \leq 3\}$ can be found by solving the system of equations:

$$\lambda_j = \min_{l \neq j} \eta_{lj} \quad (1 \leq j \leq 3), j \neq 4 \tag{19}$$

and

$$\sum_{\substack{j=1 \\ j \neq k}}^{M} p_j^{(4)*} = 1 \tag{20}$$

Department of Computing Science,
University of Newcastle upon Tyne

For our illustrative example, we have $\lambda_1 = \min\{\eta_{21}, \eta_{31}\}$ and $\lambda_2 = \eta_{32}$. We can verify that the latter equation can be written as

$$p_2^{\{4\}} = 1.93 p_3^{\{4\}}. \tag{21}$$

Now, assume that $\eta_{21} < \eta_{31}$, so that

$$\lambda_1 = \eta_{21}.$$

We can verify that it yields

$$p_1^{\{4\}} = 0.75 p_2^{\{4\}}. \tag{22}$$

Equation 20, Equation 21 and Equation 22 yield $p_1^{\{4\}} = 0.44$, $p_2^{\{4\}} = 0.33$, and $p_3^{\{4\}} = 0.23$. And we can verify that $\eta_{31} = 34.04 > \lambda_1$, which justifies our assumption that $\eta_{21} < \eta_{31}$ a posteriori.

5 Optimal selection of a control

Let $C(i|j)$ be the loss in utility for applying control a_i when the environment is in state θ_j, $j = 1, 2, \ldots, M$. Then

$$C(i|j) = U_{jj} - U_{ij}.$$

$C(i|j)$ is called the regret function. This is a measure of the loss in utility for not making the best choice in an environment in state θ_j.
Note that $C(i|j) = 0$ if $i = j$.

To simplify the analysis, we will consider the following two extreme cases. When the probability of change is very small, in which case we say that the environment is stable, and when the probability of change is close to 1, in which case we say that the environment is unstable.

5.1 Stable environment

In this case $\gamma \approx 0$ and hence it is reasonable to assume that the state of the environment will not change during the next period. The problem then reduces to the identification of the current state of the environment using the accumulated information. It can be shown - see Theorem 6.7.1 in [1]- that the optimal action a_{i^*} with respect to the utility loss from misidentification is the one that minimises the posterior expected cost at point $\mathbf{x}$; that is

$$\sum_{j=1}^{M} C(i^*|j)\, p_{\theta(x)}^{\{N\}}(\theta_j) = \min_{1 \le i \le M} \sum_{j=1}^{M} C(i|j)\, p_{\theta(x)}^{\{N\}}(\theta_j), \tag{23}$$

where $p_{\theta(x)}^{\{N\}}(\theta_j)$ is the posterior probability distribution of the state of the environment at the end of the period, and it is given by Equation 6.

So, in practice, the controller computes the posterior utility loss at $\mathbf{x}$ under all actions $a_i, 1 \le i \le M$, and the action that achieves the minimum posterior loss is selected for the next period.

5.2 Unstable environment

In this case, the environment is likely to change during a period, and because the law of state transition is unknown, we adopt the minimax approach.

To evaluate the various available actions, at the end of the period lasting from T_n to T_{n+1}, we compute the expected total loss in utility during the period from T_{n+1} to T_{n+2}.

Let $(k, j; t)$ represent an environment that is in state θ_k at time T_{n+1} and then switches to state θ_j, $j \neq k$, at time $T_{n+1} + t$, where $1 \leq\ < N$.
Let $L_i(k, j; t)$ be the total expected loss in utility under action a_i in the interval $[T_{n+1}, T_{n+2}]$ when the environment is given by $(k, j; t)$. Then

$$L_i(k, j; t) = \begin{cases} tC(i|k) + (N - t)C(i|j) & if \quad j \neq i, k \neq i \\ (N - t)C(i|j) & if \quad k = i, j \neq i \\ tC(i|k) & if \quad k \neq i, j = i. \end{cases}$$

Let $q_{(k,j;t)}$ be the probability that an environment that is in state θ_k at time T_{n+1} switches to state θ_j, at time $T_{n+1} + t$. By the law of total probability, we have

$$\sum_{k=1}^{M} \sum_{\substack{j=1 \\ j \neq k}}^{M} \sum_{t=1}^{N-1} p^{(N)}_{\theta(3)}(\theta_j) q_{(k,j;t)} = \gamma. \tag{24}$$

Now, it is easy to see that the expected total loss in utility under action a_i in the interval $[T_{n+1}, T_{n+2}]$ is given by

$$E_{p^{(N)}_{\hat{\theta}(3)}}(\mathcal{L}_i) = (1 - \gamma) N \sum_{\substack{j=1 \\ j \neq i}}^{M} C(i|j) p^{(N)}_{\theta(3)}(\theta_j) + \gamma \sum_{k=1}^{M} \sum_{\substack{j=1 \\ j \neq k}}^{M} \sum_{t=1}^{N-1} q_{(k,j;t)} L_i(k, j; t) p^{(N)}_{\theta(3)}(\theta_j). \tag{25}$$

When change is very likely, $\gamma \sim 1$, and hence the first term in the r.h.s. of Equation 25 can be neglected. The probabilities $\beta_{(k,j;t)}$ are unknown, so we adopt the worst case scenario and seek to minimise the expected future utility loss regardless of the change which occurs in the next period (minimax approach).

We can think of this problem as a two-person zero-sum game, where player 1 is the controller and player 2 is the environment. The set of actions for the controller is the set A, and for the environment, the set of possible actions is

$$\mathcal{N} = \{\nu = (k, j; t), (1 \leq k \leq M), (1 \leq j \leq M, j \neq k), (1 \leq t < N)\}.$$

Let $\mathcal{P}(\mathcal{N})$ be the set of environmental strategies $\{\beta_\nu, \nu \in \mathcal{N}\}$ which satisfy the constraint given by Equation 24. Let

$$\hat{L}_{i\nu} = L_i(k, j; t) p^{(N)}_{\theta(3)}(\theta_j),$$

and

$$e(\delta, \beta) = \sum_{i=1}^{M} \sum_{\nu \in \mathcal{N}} \delta_i \beta_\nu \hat{L}_{i\nu}, \tag{26}$$

be the expected utility loss when the controller adopts strategy $\delta = (\delta_1(\mathbf{x}), \cdots, \delta_M(\mathbf{x}))$ and the environment adopts strategy $\beta = \{\beta_\nu, \nu \in \mathcal{N}\}$.

The problem from the environment's point of view can then be written as

$$\sup_\beta \inf_\delta e(\delta, \beta), \tag{27}$$

whereas the problem from the controller's point of view is

$$\inf_\delta \sup_\beta e(\delta, \beta). \tag{28}$$

The Nash equilibrium of the game (δ^*, β^*) is such that

$$e(\delta, \beta^*) \geq e(\delta^*, \beta^*) = v \geq e(\delta^*, \beta), \tag{29}$$

where v is the value of the game.

Now, when adopting a pure strategy a_i, $1 \leq i \leq M$, the first inequality in Equation 29 is clearly also satisfied, and this can be written as

$$\sum_{\nu \in \mathcal{N}} \beta_\nu \hat{L}_{i\nu} \geq v. \tag{30}$$

On the other hand, the environment's objective is to make v as high as possible. Thus, the problem can be written as the linear optimisation problem

$$
\begin{aligned}
LP1: \quad & \textit{maximise} \quad v \\
& \text{subject to} \\
& \qquad \sum_{\nu \in \mathcal{N}} \hat{L}_{i\nu} \beta_\nu \geq v, \quad 1 \leq i \leq M, \\
& \qquad \beta \in \mathcal{P}(\mathcal{N}).
\end{aligned}
$$

Now, let

$$\hat{\beta}_\nu = \frac{\beta_\nu}{v}, \tag{31}$$

so that Equation 24 becomes

$$\sum_{\nu \in \mathcal{N}} p^{(N)}_{\theta(i)}(\theta_j) \hat{\beta}_\nu = \frac{\gamma}{v}. \tag{32}$$

Thus, maximising v is equivalent to minimising γ/v or $\sum_{\nu \in \mathcal{N}} p^{(N)}_{\theta(i)}(\theta_j) \hat{\beta}_\nu$. Problem $LP1$ is then equivalent to problem

$$
\begin{aligned}
LP2: \quad & \textit{minimise} \quad \sum_{\nu \in \mathcal{N}} p^{(N)}_{\theta(i)}(\theta_j) \hat{\beta}_\nu \\
& \text{subject to} \\
& \qquad \sum_{\nu \in \mathcal{N}} \hat{L}_{i\nu} \hat{\beta}_\nu \geq 1, \quad 1 \leq i \leq M, \\
& \qquad \hat{\beta}_\nu \geq 0, \quad \nu \in \mathcal{N}.
\end{aligned}
$$

To find the controller's optimal strategy δ^*, we need to solve the dual of problem $LP2$ given by

$$LP3: \quad maximise \quad \sum_{1 \le i \le M} \hat{\delta}_i$$
$$\text{subject to}$$
$$\sum_{1 \le i \le M} \hat{L}_{i\nu} \hat{\delta}_i \le p_{\theta|\theta_j}^{(N)}(\theta_j), \quad \nu \in \mathcal{N},$$
$$\hat{\delta}_i \ge 0, \quad 1 \le i \le M.$$

Let $(\hat{\delta}_1^*, \hat{\delta}_2^*, \dots, \hat{\delta}_M^*)$ be the solution of problem $LP3$. Then

$$\delta_i^* = \frac{\hat{\delta}_i^*}{\sum_{1 \le i \le M} \hat{\delta}_i^*}, \quad \text{for all } 1 \le i \le M$$

Note that the future expected loss can be reduced by allowing the selection of up to two controls in a period, but at the expense of increasing the size of problem $LP3$.

When M and/or N are large, the time required for the solution of problem $LP3$ may not be negligible. In this case, we need to incorporate the delay in the implementation of the controls in the analysis.

6 Concluding remarks

The solution of the adaptation of a plant to a changing environment is composed of two steps. In the first step, we find the maximin prior of the state at the time of change for all possible states at the beginning of a period and times of change, and then compute the posterior probability distribution of the state of the environment at the end of the period. In the second step, in which the selection of the control is made, the problem is formulated as either a classification problem or a two-person zero-sum game, depending on whether change in a period is rare or almost certain.

When the probability of change γ is such that $0 \ll \gamma \ll 1$, it may be safer to treat the environment as unstable and adopt the resulting optimal strategy.

In the selection process, the criterion for evaluation of the available controls is the minimum expected loss in utility instead of the maximum expected utility, because the latter can lead to bad results for some values of the utilities; see [3] for a detailed discussion.

In this paper, we assumed that the data in a sample are the realisation of independent random variables. In practice, most samples will be the realisation of time series with some dependency. There are change point tests for some cases of dependency, and these problems are still the object of intensive research in statistics. For dependent data, the determination of the classification regions may be complex, and will need special attention. We also implicitly assumed that the utilities $\{U_{ij}, (1 \le i \le M), (1 \le j \le M)\}$ are constant. However, in many applications, these utilities can be time dependent; this is the case, for example, in systems whose functioning is affected when some controls are used for too long. The modelling and analysis of situations with time-dependent utilities is the object of future work.

Acknowledgements
The author would like to thank Rogério de Lemos, Tom Anderson and Emiliano Tramontana for the fruitful discussions, and to acknowledge the financial support of EPSRC/UK SafeGames project.

References

[1] T. W. Anderson, A Bibliography of multivariate statistical analysis. Oliver and Boyd, Edinburgh, 1954.

[2] R. J. Bennet, Spatial time series: Analysis-forecasting-control. Pion Limited, London, 1979.

[3] J. O. Berger, Statistical decision theory: Foundations, concepts and methods. Springer-Verlag, New York, 1980.

[4] M. Csorgo and L. Horvath, Limit theorems in change-point analysis. Wiley and Sons Ltd, Chichester, 1997.

[5] M. H. DeGroot, Optimal statistical decisions. McGraw-Hill, New York, 1970.

[6] R. M. Gagliardi and I. S. Reed, "On the sequential detection of emerging targets", *IEEE Trans. Inform. Theory*, vol. IT-6, 1965, pp. 260-263.

[7] Y. Liu, Quickest detection of a a random sequence with application to adaptive identification in fading channels. Ph.D. dissertation, Dept. of Elect. Eng., Queen's Univ., Kingston, Ont., Canada, 1993.

[8] Y. Liu and S. D. Blostein, "Quickest detection of an abrupt change in a random sequence with finite change-time", *IEEE Trans. Inform. Theory*, vol. 40, 1994, pp. 1985-1993.

[9] G. Lorden, "Procedures for reacting to a change in distribution", *Ann. Mathemat. Statist.*, vol. 42, 1971, pp. 1897-1908.

[10] R. D. Luce and H. Raiffa, Games and decisions. Wiley, New York, 1957.

[11] S. Macdougall and A. K. Nandi, "Hybrid bayesian procedures for automatic detection of change-points", *J. Franklin Inst.*, vol. 334B, 1997, pp. 575-597.

[12] R. K. Mehra, "Optimisation of measurement schedules and sensor designs for linear systems", *IEEE Transactions on Automatic Control*, AC-21, 1976, pp. 55-64

[13] A. N. Shyryayev, Optimal stopping rules. Springer-Verlag, New York, 1978.

[14] A. F. M. Smith, "Bayesian approach to inference about a change point in a sequence of random variables", *Biometrica*, vol. 62, 1975, pp. 407–416.

[15] T. Stauner, O. Mller and M. Fuchs, "Using HYTECH to verify an automative control system", in *Hybrid and Real-Time Systems, Lecture Notes in Computer Science*, vol. 1201, O. Maler (ed), Springer-Verlag, Berlin, Germany, 1997, pp. 139-153.

[16] J. von Neumann and O. Morgenstern, Theory of games and economic behavior. Wiley, 2nd edition, 1947.

[17] A. S. Willsky, "A survey of design methods for failure detection in dynamic systems", *Automatica*, vol. 12, 1976, pp. 601-611.

[18] S. Zacks, "Survey of classical and bayesian approaches to the change point problem", in *Recent Advances in Statistics*, Herman Chernoff Festschrift, Academic Press, New York, 1983, pp. 245-269.