UNIVERSITY OF SOUTHERN CALIFORNIA

SCHOOL OF MEDICINE

LABORATORY OF APPLIED PHARMACOKINETICS

Technical Report: 91-1

APPLICATION OF STOCHASTIC CONTROL THEORY TO OPTIMAL DESIGN OF DOSAGE REGIMENS*

by

Alan Schumitzky
Department of Mathematics
University of Southern California
Los Angeles, California 90089

*This work was supported in part by the National Institutes of Health under Grant RR01629.

To appear in: Advanced Methods in Pharmacokinetic and Pharmacodynamic Systems Analysis, edited by D. Z. D'Argenio, Plenum Press, New York, 1991.

APPLICATION OF STOCHASTIC CONTROL THEORY TO OPTIMAL DESIGN OF DOSAGE REGIMENS

Alan Schumitzky

Department of Mathematics
University of Southern California

ABSTRACT

Designing a dosage regimen for a pharmacokinetic/pharmacodynamic system involves defining: i) a patient-dependent model, which includes structure, parameter, and measurement uncertainties; ii) the choice of controls, which can include dose amounts, dose times and/or sampling times; and iii) an appropriate performance index to evaluate achievement of a clinically chosen therapeutic goal. The control problem then is to choose the dosage regimen that optimizes the expected value of the performance index. This problem fits within the framework of stochastic control theory. Examples are given to illustrate the variety of this class of problems, including: optimal dose regimens for target level and target window cost; and optimal sampling schedules for maximal information. By varying the class of admissible controls, different strategies are generated. Control strategies to be discussed include: open loop, open loop feedback, separation principle, and iteration in policy space. Monte Carlo simulation studies of a terminal cost type problem are presented.

INTRODUCTION

Designing a dosage regimen for a pharmacokinetic/pharmacodynamic (PK/PD) system involves defining:

- a patient-dependent PK/PD model, which includes structure, parameter, and measurement uncertainties;
- the choice of "controls", which can include dose amounts, dose times and/or sampling times; and
- an appropriate performance index to evaluate achievement of a clinically chosen therapeutic goal.

The control problem then is to choose the dosage regimen that optimizes the expected value of the performance index. This problem fits within the framework of "stochastic" control theory, i.e., control in the presence of uncertainty.

Examples are given to illustrate the variety of this class of problems in the PK/PD setting. Included will be: optimal dose regimens for target level and target window cost; and optimal sampling schedules for maximal information.

By varying the class of admissible controls, different strategies are generated. Control strategies to be discussed include: open loop, open loop feedback, separation principle and iteration in policy space. Monte Carlo simulation studies of a terminal cost type problem are presented.

STOCHASTIC CONTROL FORMALISM

In this section we define the various ingredients which make up a stochastic control problem.

State Equations

The time history of a drug concentration in a traditional PK model is typically described by a system of deterministic ordinary differential equations. When "noise" is added to this model the system of differential equations becomes "stochastic". Since stochastic differential equations are mathematically quite sophisticated, it is desirable to seek a simpler formulation. The possibility of such a simplification comes from the fact that dose inputs and infusion rates are changed only at discrete time points (as opposed to continuously). In the linear case, this leads to an equivalent linear system of "discrete time" stochastic equations (as illustrated below). In the nonlinear case the appropriate model is given by a discrete time "Markov process" [1]. For purposes of this exposition, we take the middle ground and assume that the PK/PD model can be described by a nonlinear system of discrete time stochastic equations as follows:

$$x_{n+1} = f_n(x_n, u_n, w_n, \phi), \qquad n = 0, 1, \ldots, N \tag{1}$$

where at "stage" n, x_n is the "state" vector of the system, u_n is the external "control" vector, w_n is the "process noise" vector, and f_n is a known vector function. Further, ϕ is an unknown time invariant parameter vector.

In the PK/PD setting, the state x_n typically corresponds to amounts and concentrations of drug in various compartments and/or to various drug effects. The control u_n typically corresponds to the drug amounts and/or infusion rates of one or more drugs into various compartments; additionally, the control u_n could also include various design entities such as dose times and sample times. The parameter vector ϕ typically corresponds to various PK entities such as rate constants, distribution volumes, clearances, and/or to various PD entities such as C_{50} and E_{max}. The process noise vector w_n typically corresponds to various errors made in dose amount, dose timing and model mispecification. Finally, the stage n typically corresponds to a dose time or infusion time or sample time.

Measurement Equations

The measurement model is also described by discrete time stochastic equations. These equations can be written as :

$$y_n = h_n(x_n, u_n, v_n, \phi), \qquad n = 1, 2, \ldots, N \tag{2}$$

where at stage n, y_n is the "measurement" vector ; v_n is the corresponding "measurement noise" vector, and h_n is a known vector function.

In the PK/PD setting, the measurement y_n typically corresponds to concentrations of drug in serum or amounts of drug in urine but could also be measured drug effects; the measurement noise v_n typically corresponds to assay or measurement device noise but could also correspond to the errors in recording the time of observations.

Prior Distributions

In the model of Eqs. (1) and (2), the vectors x_0, w_n, v_n, and ϕ are the basic random variables; and the prior probability density functions:

$$p(x_0 \mid \phi), \qquad p(w_n \mid x_n, u_n, \phi), \qquad p(v_n \mid x_n, \phi), \qquad p(\phi) \tag{3}$$

are assumed known. All the other random vectors x_n, u_n, and z_n are functions of x_0, w_n, v_n, and ϕ.

The determination of these prior distributions is an important ingredient in the stochastic control formulation. For the measurement noise, the prior distribution of v_n (given x_n and ϕ) comes from "calibrating" the measurement devices. For the PK/PD parameters, the prior distribution of ϕ comes from analyzing previous studies. This latter problem is called "Population Analysis" and is the subject of much interest in PK/PD applications. (See [2, 3] for survey articles and the chapter by Mentre and Mallet in this volume.) The initial state vector x_0 is quite often known exactly (e.g., $x_0 = 0$). Otherwise the prior distribution of x_0 must be determined from "prior" events. On the other hand, the process noise term w_n is a relatively new addition to PK/PD problems. The determination of its prior distribution (given x_n, u_n and ϕ) is, for all practical purposes an essentially unexplored problem. (See [4, 5] and the chapter by D'Argenio in this volume.)

Admissible Controls

A realizable control u_n must be *nonanticipatory*. This means that u_n can only depend on present and past data. More precisely:

$$u_n \text{ is a function of the information } I_n = (y_1, \ldots, y_n; u_0, \ldots, u_{n-1}). \tag{4}$$

For our PK/PD applications, the components of u_n will be constrained (doses cannot be negative or arbitrarily large, dose times must be sequential, etc.). Further, it may also be necessary to constrain certain components of the state x_n (serum levels should not be too large, platelet counts should not be too small). Since x_n is random, these constraints will only be required to hold in a probabilistic sense. Such constraints also (implicitly) imply constraints on the controls $(u_0, u_1, \ldots, u_n)$. All these constraints will be collected under the assumption:

$$u_n \text{ belongs to some set } \mathcal{U}_n \text{ which may depend on } (u_0, u_1, \ldots, u_{n-1}). \tag{5}$$

A control *policy* $U \equiv (u_0, u_1, \ldots, u_N)$ is called *admissible*, if U and the resulting $X \equiv (x_0, x_1, \ldots, x_{N+1})$ satisfy Eqs. (1)–(5). The control problem then is to choose the admissible control policy U to maximize some "performance index" or minimize some "cost function".

For example it may be required to design a regimen to maximize the probability that certain drug levels and/or effects belong to some therapeutic window. Here an appropriate performance index would be of the form:

$$J(U) = \sum_{n=0}^{N} \alpha_n \text{ Prob}\{g_n(x_n) \in S_n\} \tag{6}$$

where $g_n(x)$ is a given function of x, and S_n is a given set (the therapeutic window). In this case U would be chosen to maximize $J(U)$.

Similarly, it may required to design a regimen to minimize the error between certain drug levels effects and some desired response. Here an appropriate cost function would be of the form:

$$J(U) = \text{E}\{\sum_{n=0}^{N} \alpha_n[\| x_{n+1} - L_{n+1} \|^2 + \gamma_n \| u_n \|^2]\} \tag{7}$$

where L_n is the desired response at stage n, $\| \cdot \|$ is some vector "norm", and where α_n and γ_n are given nonnegative constants reflecting the relative importance of the corresponding terms (including, for example $\gamma_n \equiv 0$). In this case U would be chosen to minimize $J(U)$.

Finally, to illustrate an example that is not normally considered in the context of stochastic control, it may be required to design a sampling schedule to maximize some index of "information". Here an appropriate performance index could be of the form:

$$J(U) = E\{\det[M(\phi, U)]\} \tag{8}$$

where

$$M(\phi, U) = E\left\{(\partial \log p(Y \mid \phi)/\partial\phi)(\partial \log p(Y \mid \phi)/\partial\phi)^T \mid \phi\right\} \tag{9}$$

is the Fisher information matrix and where $Y = (y_1, y_2, \ldots, y_N)$. In this case U would be chosen to maximize $J(U)$.

These three examples can be put into the following form: Some criterion function $C = C(U, X, \phi)$ is given and the expected value of C is to be optimized. To be specific, we will suppose that $E\{C\}$ is to be minimized. (For maximization problems, just replace C by $-C$.)

Further, for technical reasons, it will be assumed that $C(X, U, \phi)$ is of the form:

$$C(U, X, \phi) = \sum_{n=0}^{N} g_n(x_{n+1}, u_n, \phi) \tag{10}$$

where g_n is the "cost" at stage n. (It is clear that the criterion functions in Eqs. (6),(7) are in the form of Eq. (10). To put the criterion function in Eq. (9) into this form requires "additional" state variables. This is illustrated in Example 2.)

The stochastic control problem can now be stated precisely as:

Find the admissible policy $U^* = (u_0^*, u_1^*, \ldots, u_N^*)$ which minimizes

$$J(U) = E\{C(U, X)\} \tag{11}$$

over all admissible U.

CONTROL POLICIES

The stochastic control problem as stated above is extremely general. It can be made to reflect many of the decisions made in a clinical situation. Unfortunately, the optimal control U^* cannot be implemented except for the most trivial cases. The reason for this is outside the scope of this paper, but it essentially goes under the name of "The Curse of Dimensionality". In this section we, therefore, discuss various types of suboptimal control policies which can be implemented. We restrict our attention to those policies which have appeared in PK/PD settings.

Open Loop

In the *open loop* policy, the controller ignores all measurement data and depends only on the prior "information" $I_0 = \{p(x_0 \mid \phi),\ p(w_n \mid x_0, \phi),\ p(\phi)\}$. The optimal open loop (OL) control $U^{OL} \equiv (u_0^{OL}, \ldots, u_N^{OL})$ minimizes the expression

$$E\left\{\sum_{i=0}^{N} g_i(x_{i+1}, u_i, \phi) \mid I_0\right\} \tag{12}$$

with respect to $(u_0, u_1, \ldots, u_N)$. (For notational simplicity, in Eq. (12) and throughout this section, we assume that $g_n(x_{n+1}, u_n, \phi) \equiv \infty$, if $u_n \notin \mathcal{U}_n$ so that the constraints of Eq. (5) can be suppressed.)

Open Loop Feedback

In the *open loop feedback* policy, the controller at stage n acknowledges that the information I_n is available, but assumes that no measurements will be taken in the future. The optimal open loop feedback (OLF) control $U^{OLF} \equiv (u_0^{OLF}, \ldots, u_N^{OLF})$ is such that, at each stage n, u_n^{OLF} minimizes the expression

$$E\{g_n(x_{n+1}, u_n, \phi) + \min_{u_{n+1} \cdots u_N} \left[\sum_{i=n+1}^{N} g_i(x_{i+1}, u_i, \phi)\right] \mid I_n\} \tag{13}$$

with respect to u_n.

In Eqs. (12) and (13), the expression $E\{C \mid I_n\}$ denotes the expectation of C conditioned on the information I_n. Note that the optimal OLF control is just the optimal OL control starting at stage n with "prior" information I_n. The fact that measurements can be used to advantage is reflected by the result [6]:

$$J(U^{OLF}) \leq J(U^{OL})$$

Separation Principle

In the *separation principle* policy, the controller applies at each stage the deterministic control that would be applied if all random terms were fixed at their expected values. The optimal separation principle (SP) control $U^{SP} \equiv (u_0^{SP}, \ldots, u_N^{SP})$ is such that, at each stage n, u_n^{SP} minimizes the deterministic expression:

$$g_n(x_{n+1}, u_n, \hat{\phi}) + \min_{u_{n+1} \cdots u_N} \left[\sum_{i=n+1}^{N} g_i(x_{i+1}, u_i, \hat{\phi})\right]$$

with respect to u_n, where

$$x_{i+1} = f_i(x_i, u_i, 0, \hat{\phi}), \qquad x_i = \hat{x}_n, \quad i = n, \ldots, N \tag{14}$$

In Eq. (14) it is assumed that $E\{w_i\} = 0$. Further, $\hat{x}_n$ and $\hat{\phi}$ are some estimates of x_n and ϕ based on the information I_n.

The separation principle is so named as the resulting controller "separates" the problem of estimation and control. The popularity of the SP controller comes from its ease of computation and from the fact that the optimal SP controller is actually optimal with respect to all admissible controls in the linear, quadratic, Gaussian case (LQG). In the LQG case: the state and measurement equations are linear; the cost function is quadratic; all noise terms are independent Gaussian; there is no unknown parameter vector ϕ; and $\hat{x}_n = E\{x_n \mid I_n\}$. Unfortunately these conditions rarely hold in PK/PD applications. And in general:

$$J(U^{SP}) \not\leq J(U^{OL})$$

Iteration in Policy Space

The optimal control policy $U^* = (u_0^*, u_1^*, \ldots, u_N^*)$ defined in Eq. (11) satisfies an important recursion relationship. At each stage n, the control u_n^* minimizes the expression

$$E\{g_n(x_{n+1}, u_n, \phi) + \sum_{i=n+1}^{N} g_i(x_{i+1}, u_i^*(I_i), \phi) \mid I_n\} \tag{15}$$

with respect to u_n. In Eq. (15), the dependence of u_i^* on I_i is explicitly shown to indicate the way that u_i^* depends on u_n, for $i = n+1, \ldots, N$. Equation (15) leads to Bellman's method of Stochastic Dynamic Programming. (This was one of the earliest and most important results in stochastic control theory [6, 7].)

The major drawback of using Eq. (15) for computational purposes is that it must be solved "backwards in time", since otherwise u_n^* would depend on the future controls $(u_{n+1}^*, \ldots, u_N^*)$. This problem led Bayard [8, 9] to consider analogous methods which could be solved "forwards in time". A brief description of Bayard's approach is as follows: Let $U^0 \equiv (\mu_0, \ldots, \mu_N)$ be any admissible control policy (called the *nominal policy*). Then an *iteration in policy space* (IPS), with respect to U^0, is the control policy $U^{IPS} \equiv (u_0^{IPS}, \ldots, u_N^{IPS})$ such that, at each stage n, u_n^{IPS} minimizes the expression

$$E\{g_n(x_{n+1}, u_n, \phi) + \sum_{i=n+1} g_i(x_{i+1}, \mu_i(I_i), \phi) \mid I_n\} \tag{16}$$

with respect to u_n. It can be shown that U^{IPS} improves on U^0, i.e.,:

$$J(U^{IPS}) \leq J(U^0)$$

(Note the similarity between Eq. (15) and Eq. (16). Also note that calculating U^{IPS} from Eq. (16) is, in fact, simpler than calculating U^{OLF} from Eq. (13).)

This iteration process can be continued. The control policy U^{IPS} is admissible and can, therefore, be substituted for the original nominal policy. The resulting iteration on the policy U^{IPS}, now denoted by U^{2-IPS}, will, therefore, satisfy

$$J(U^{2-IPS}) \leq J(U^{IPS}).$$

A remarkable result of Bayard [8] is that if this process is continued for (at most) N iterations, then the optimal control U^* of Eq. (11) is obtained, i.e.,

$$U^{N-IPS} = U^*$$

However even one iteration in policy space can dramatically improve control performance. This is illustrated in the simulation study presented below (Example 1).

Active vs. Passive Learning

An important feature of the optimal control policy U^* is how it "learns" about the unknown parameters and states. In contrast to the OL, OLF, and SP controllers which learn only "passively", the optimal policy learns "actively" by probing the system. Probing comes from the anticipation that future measurements will be made, so that "mistakes" can be corrected [10]. In this sense, the optimal policy does experimental design "on line".

It is somewhat surprising that even one iteration in policy space on a "passive" nominal can generate a "active" controller. For example, it is shown in [11] that if the nominal policy is taken as U^{OLF} then the resulting U^{IPS} has this behavior. It is observed in the simulation study presented below (Example 1), that the same is true if the nominal policy is U^{SP}.

Previous PK/PD Applications

There have been only a few PK/PD applications of stochastic control theory to dynamical systems with process noise. In [12], an SP controller was simulated for a PK/PD problem in anticoagulant therapy; in [13] an OL controller was simulated for a PK problem in theophylline therapy; and in [14], an IPS type controller was simulated for a PK "terminal cost" problem.

Deterministic State Equations

In traditional PK/PD problems, the state Eq. (1) is deterministic (for given ϕ). That is, the process noise is assumed to be zero ($w_n \equiv 0$) and the initial state x_0 is known exactly, e.g., $x_0 = 0$. Now the cost function, Eq. (11), is of the form:

$$J(U) = \int C(U, X, \phi) p(\phi) d\phi$$

In this case the special policies OL, OLF, SP, and IPS above are all considerably simpler. Most previous PK/PD applications of stochastic control have been in this setting.

One of the earliest applications of "sophisticated" stochastic control theory was due to Gaillot, Steimer, Mallet, Thebault, and Bieder in 1979 where an optimal OL controller was utilized for lithium therapy [15] . In Richter and Reinhardt [16], a similar OL controller was utilized for theophylline therapy. More recently, Mallet, *et al.* [17] combined sophisticated population analysis with an optimal OL controller in designing dosage regimens for gentamicin therapy.

Simulated PK applications of optimal OLF control appeared in Katz and D'Argenio [18] and D'Argenio and Katz [19]. Applications of SP type controllers are more numerous. One of the earliest was the MAP Bayesian controller of Sheiner [20]. Similar controllers are found in the USC PC Pack of Jelliffe, *et al.* [21, 22]. Surveys and tutorials in this subject are found in Vozeh and Steimer [23], and Schumitzky [24, 25].

EXAMPLE 1: OPTIMAL INFUSION REGIMEN

In the next two sections we illustrate by example some of the ingredients of the stochastic control formalism. For this purpose it is sufficient to consider the simplest PK settings.

In this section we consider a one compartment model with iv infusion. In general one of the problems with the "textbook" version of the stochastic control formalism is that it does not easily conform to the standard setting in PK/PD applications. The solution to this problem is a version of the stochastic control formalism which includes continuous time state equations and discrete time measurement equations. However this is not the place for such a digression. It will be apparent in the example below that certain awkward constructions could be avoided by a "continuous-discrete" formalism.

State Equation

In the deterministic case, the time history of the drug concentration satisfies the differential equation:

$$\frac{dC(t)}{dt} = -kC(t) + \frac{r(t)}{V}, \quad t \geq 0, \qquad C(0) = 0 \tag{17}$$

where at time t, $C(t)$ is the concentration of drug, $C(0) = 0$; and $r(t)$ is the iv infusion rate which is assumed to be piecewise constant:

$$r(t) = r_n, \; t \in [t_n, t_{n+1}), \quad n = 0, 1, \ldots, N,$$

where $\{t_n\}$ are the times at which the infusion rates can change. Further, k is the elimination rate constant and V is the volume of distribution.

If a Gaussian "white" noise process $W(t)$ with mean 0 and "variance" Q is added to Eq. (17), then the differential equation becomes stochastic and is written mathematically as:

$$dC(t) = -\left\{kC(t) + \frac{r(t)}{V}\right\} dt + d\beta(t), \qquad t \geq 0 \tag{18}$$

where $\beta(t)$ is a so-called Brownian motion such that $W(t) = d\beta(t)/dt$. (The noise term $W(t)$ can be considered as model mispecification.)

The solution to Eq. (18) is given by the stochastic integral

$$\begin{aligned} C(t) &= \exp\{-k(t - t_n)\}C(t_n) + [1 - \exp\{-k(t - t_n)\}]/(kV)\, r_n \\ &+ \int_{t_n}^{t} \exp\{-k(t - s)\} d\beta(s), \qquad t \in [t_n, t_{n+1}) \end{aligned} \tag{19}$$

To define a discrete time state equation corresponding to Eq. (1), set: $x_n = C(t_n)$, $\phi = (k, V)$ and

$$A(t, \tau, \phi) = exp\{-k(t-\tau)\} \tag{20}$$

$$B(t, \tau, \phi) = [1 - A(t, \tau, \phi)] / (kV) \tag{21}$$

$$W(t, \tau) = \int_{\tau}^{t} \exp\{-k(t-s)\} d\beta(s) \tag{22}$$

It follows from Eqs. (20)-(22):

$$x_{n+1} = A_n(\phi) x_n + B_n(\phi) r_n + W_n, \qquad n = 0, 1, \ldots, N \tag{23}$$

where $A_n(\phi) = A(t_{n+1}, t_n, \phi)$, $B_n(\phi) = B(t_{n+1}, t_n, \phi)$ and $W_n = W(t_{n+1}, t_n)$. It can be shown [1] that the sequence $\{W_n\}$ is independent Gaussian with mean 0 and variance

$$Q_n = Q \frac{\left\{1 - [B_n(\phi)]^2\right\}}{2k}$$

Additionally, if the infusion rate applied at time t_n is not exactly r_n but is equal to $r_n + \delta r_n$, where the error sequence $\{\delta r_n\}$ is independent Gaussian with mean 0 and variance R_n, then Eq. (23) becomes

$$x_{n+1} = A_n(\phi) x_n + B_n(\phi) r_n + w_n, \qquad n = 0, 1, \ldots, N \tag{24}$$

where the process noise

$$w_n = B_n(\phi) r_n + W_n$$

is independent Gaussian with mean 0 and variance $[B_n(\phi)]^2 R_n + Q_n$. Equation (24) is then the state equation corresponding to Eq. (1).

Measurement Equation

Assume noisy measurements $y(s_m)$ of the concentrations $C(s_m)$ are taken at the sampling times $\{s_m\}$ such that:

$$y(s_m) = C(s_m) + \epsilon_m, \quad m = 1, 2, \ldots, M$$

where ϵ_m is the assay noise of the m^{th} measurement. The random variables $\{\epsilon_m\}$ are typically assumed to be independent Gaussian with mean 0 and standard deviation $\sigma_m = g(C(s_m))$. The function $g(C)$ comes from calibrating the assay device at various test concentrations C.

To define a discrete time measurement equation corresponding to Eq. (2), it is necessary to do some bookkeeping. Essentially y_n is the collection of all measurements $\{y(s_m) : s_m \in [t_n, t_{n+1})\}$. The relationship between y_n and x_n is determined as follows. Let $S = \{s_1, \ldots, s_M\}$ be the set of sample times. Then for each $n = 0, 1, \ldots, N-1$, either $S \cap [t_n, t_{n+1})$ is empty, or there are indices $m(n)$ and $m(n+1)$ such that

$$S \cap [t_n, t_{n+1}) = \{s_{m(n)+1}, \ldots, s_{m(n+1)}\}. \tag{25}$$

Therefore, define:

$$\begin{aligned} y_n &= 0, && \text{if } S \cap [t_n, t_{n+1}) \text{ is empty.} \\ y_n &= (y(s_{m(n)+1}, \ldots, y(s_{m(n+1)})), && \text{if Eq. (25) holds.} \end{aligned} \tag{26}$$

(For example, if $N = 4$, $M = 3$, $s_1, s_2 \in [t_0, t_1)$ and $s_3 \in [t_3, t_4)$, then $m(0) = 0$, $m(1) = 2$, $m(3) = 2$, $m(4) = 3$ and $y_1 = (y(s_1), y(s_2))^T$, $y_2 = y_3 = 0, y_4 = y(s_3)$.)

Further, for $s_m \in [t_n, t_{n+1})$, it follows from Eqs. (19)-(22):

$$C(s_m) = A(s_m, t_n, \phi)x_n + B(s_m, t_n, \phi)r_n + W(s_m, t_n)$$

Therefore when Eq. (25) holds:

$$y_n = D_n(\phi)x_n + E_n(\phi)r_n + v_n \tag{27}$$

where

$$\begin{aligned} D_n(\phi) &= (A(s_{m(n)+1}, t_n, \phi), \ldots, A(s_{m(n+1)}, t_n, \phi))^T \\ E_n(\phi) &= (B(s_{m(n)+1}, t_n, \phi)), \ldots, B(s_{m(n+1)}, t_n, \phi)^T \\ v_n &= (W(s_{m(n)+1}, t_n), \ldots, W(s_{m(n+1)}, t_n)^T + (\epsilon_{m(n)+1}, \ldots, \epsilon_{m(n+1)})^T \end{aligned}$$

As before it can be shown that the sequence of random vectors $\{v_n\}$ is independent Gaussian with mean 0 and covariance depending on x_n and ϕ. Equations (26) and (27) are then the measurement equations corresponding to Eq. (2).

Optimal design of infusion regimens

In this case, the infusion times and sample times are fixed and the infusion rates are to be optimized. The controls are therefore given by $u_n = r_n$.
The controls will be explicitly constrained to be nonnegative and implicitly constrained by the conditions:

$$E\{x_n\} \leq L_{max}, \quad n = 1, 2, \ldots, N+1 \tag{28}$$

where L_{max} is some maximum allowable level.

Cost Function: Target Level Control-Quadratic Cost

Here we want $x_n \approx L_n$, where L_n is some desired target level at stage n. A suitable cost function corresponding to Eq. (7) is given by the quadratic cost function:

$$J(U) = E\{\sum_{n=0}^{N} \alpha_n(x_{n+1} - L_{n+1})^2 + \gamma_n(u_n)^2\} \tag{29}$$

where α_n and γ_n are given nonnegative constants reflecting the relative importance of the corresponding terms (including, for example $\gamma_n \equiv 0$). The optimization problem is to choose the admissible control policy to minimize Eq. (29).

Simulation Study: *IPS* vs. *SP*

In this simulation study we compare a separation principle policy with an iteration in policy space. The results presented here are taken from [13]. In the stochastic control formulation given by Eqs. (24), (27)–(29), the following specifications are made: $t_0 = 0$, $t_n = s_n = n$, $n = 1, \ldots, 5$. In this case Eqs. (24) and (27) become

$$\begin{aligned} x_{n+1} &= ax_n + bu_n + w_n, \qquad x_0 = 0, \quad n = 0, \ldots, 5 \\ y_n &= x_n + v_n, \qquad\qquad\qquad\quad n = 1, \ldots, 5 \end{aligned}$$

where $a = \exp(-k)$ and $b = (1 - a)/(kV)$. A "natural" parameter vector for this model is then $\phi = (a, b)$.

For the cost function of Eq. (29) the targets and weights are chosen to be:

- $L_1 = 1.0$, $L_2 = 1.6$, $L_3 = 2.2$, $L_4 = 2.6$ and $L_5 = 2.8$.
- $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0.1$; $\alpha_5 = 1000$;
- $\gamma_0 = \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0.0001$.

This choice of weighting indicates that we are essentially interested in attaining the desired level only at the fifth stage of the regimen. This is essentially a "terminal cost" problem and should reveal the importance of "active" learning. (Early mistakes are not severely penalized.) The controls u_n are nonnegative and the state constraint Eq. (28) is given by $E\{x_n\} \leq 3$, $n = 1, 2, \ldots, 5$.

The prior distributions are chosen as follows: a, b, w_n and v_n are independent random variables such that:

$$\begin{aligned} w_n &\sim N(0, 0.01), \\ v_n &\sim N(0, 0.2), \\ a &\sim N(0.135, 0.0046) \quad \text{(C.V. = 34 \%)}, \\ b &\sim N(0.278, 0.0834) \quad \text{(C.V. = 30 \%)} \end{aligned}$$

where for any random variable v the notation $v \sim N(m, \sigma)$ means v is Gaussian with mean m and standard deviation σ.

Monte Carlo simulations of control performance were conducted. The simulations assessed comparative performance of an optimal separation principle (SP) type controller and an iteration in policy space (IPS) type controller. The SP controller employed an extended Kalman filter to calculate the conditional expectations of a, b, and x_n given the data I_n. The IPS controller was essentially one iteration in policy space using the previously calculated SP controller as the admissible nominal policy. (This particular controller is due to Bar-Shalom and Tse [26] and actually predates the development of the IPS formalism. See [8, 13] for more details.)

Figure 1 shows the control cost $J(U)$ in 100 simulations of the SP and IPS controllers. Not only is the mean cost for the IPS controller significantly lower than SP, but the standard deviation of the cost is also much lower. The existence of fewer outliers for IPS is readily apparent from the figure. The fact that the IPS performance is better than SP is expected from the definition of an IPS policy. The magnitude of this improvement was unexpected and it is argued that it comes from active learning.

Figure 2 shows the mean and one-sigma envelopes of the time histories of the serum concentrations for these same simulations. Note how the mean serum concentration of the IPS controller trajectory hugs the constraint boundary (exhibiting

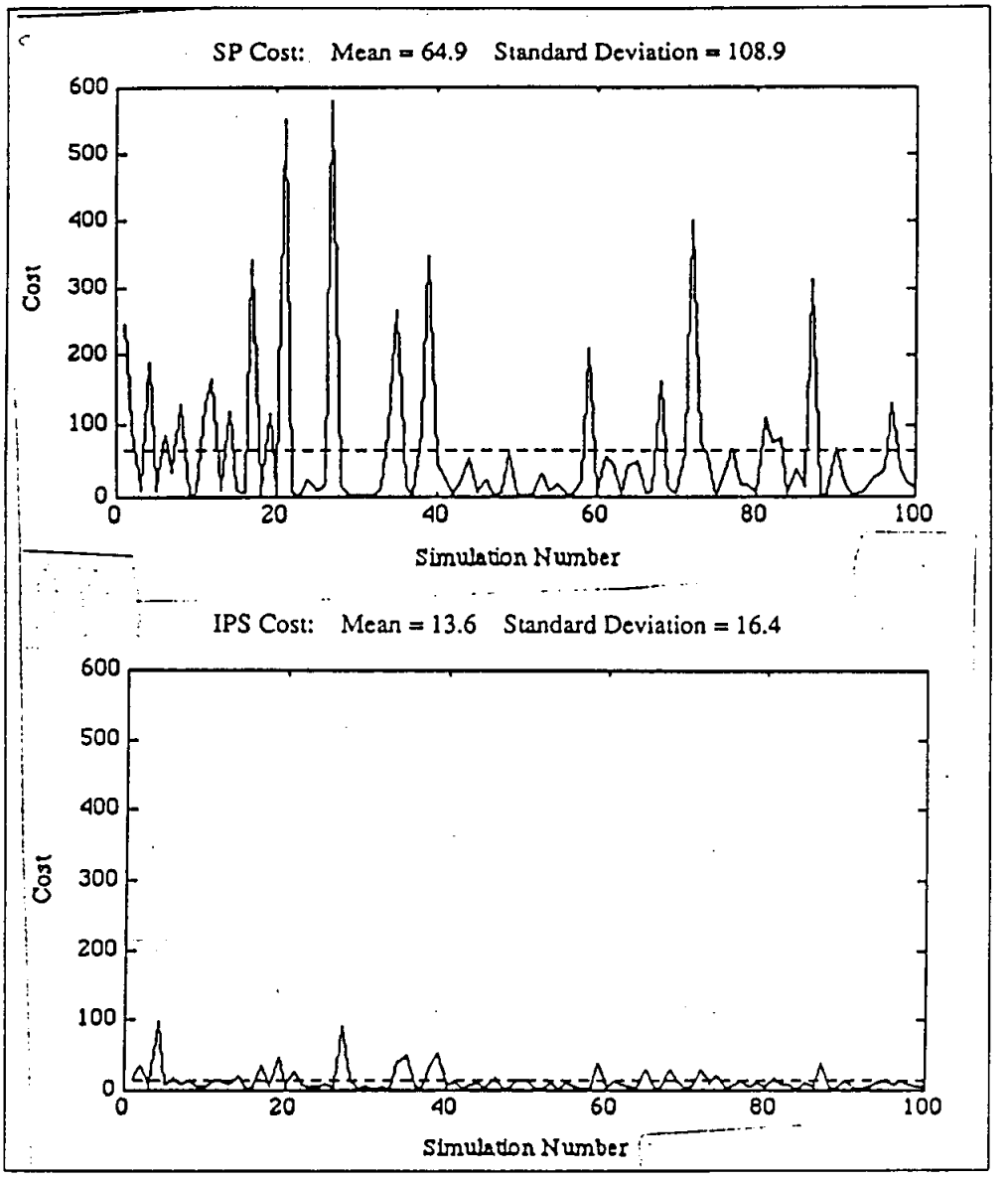

Fig. 1. *SP* and *IPS* costs.

an active learning behavior). Also note that both controllers hit the target exactly "on average".

EXAMPLE 2: OPTIMAL SAMPLING SCHEDULE DESIGN

In this section we consider the classical PK/PD problem of parameter estimation and sampling schedule design. Assume noisy measurements $y(s_n)$ are taken at the sampling times $\{s_n\}$ from a PK/PD system modeled by a "nonlinear regression" equation of the form:

$$y(s_n) = h_n(s_n, \phi) + \epsilon_n, \quad n = 1, 2, \ldots, N \tag{30}$$

where $h_n(s, \phi)$ is a known function and ϵ_n is the error of the n^{th} measurement. In Eq. (30) the following assumptions are made: The random variables $\{\epsilon_n\}$ are independent Gaussian with mean 0 and standard deviation σ_n, which is independent of ϕ but may depend on s_n. The function $h_n(s, \phi)$ has continuous partial derivatives

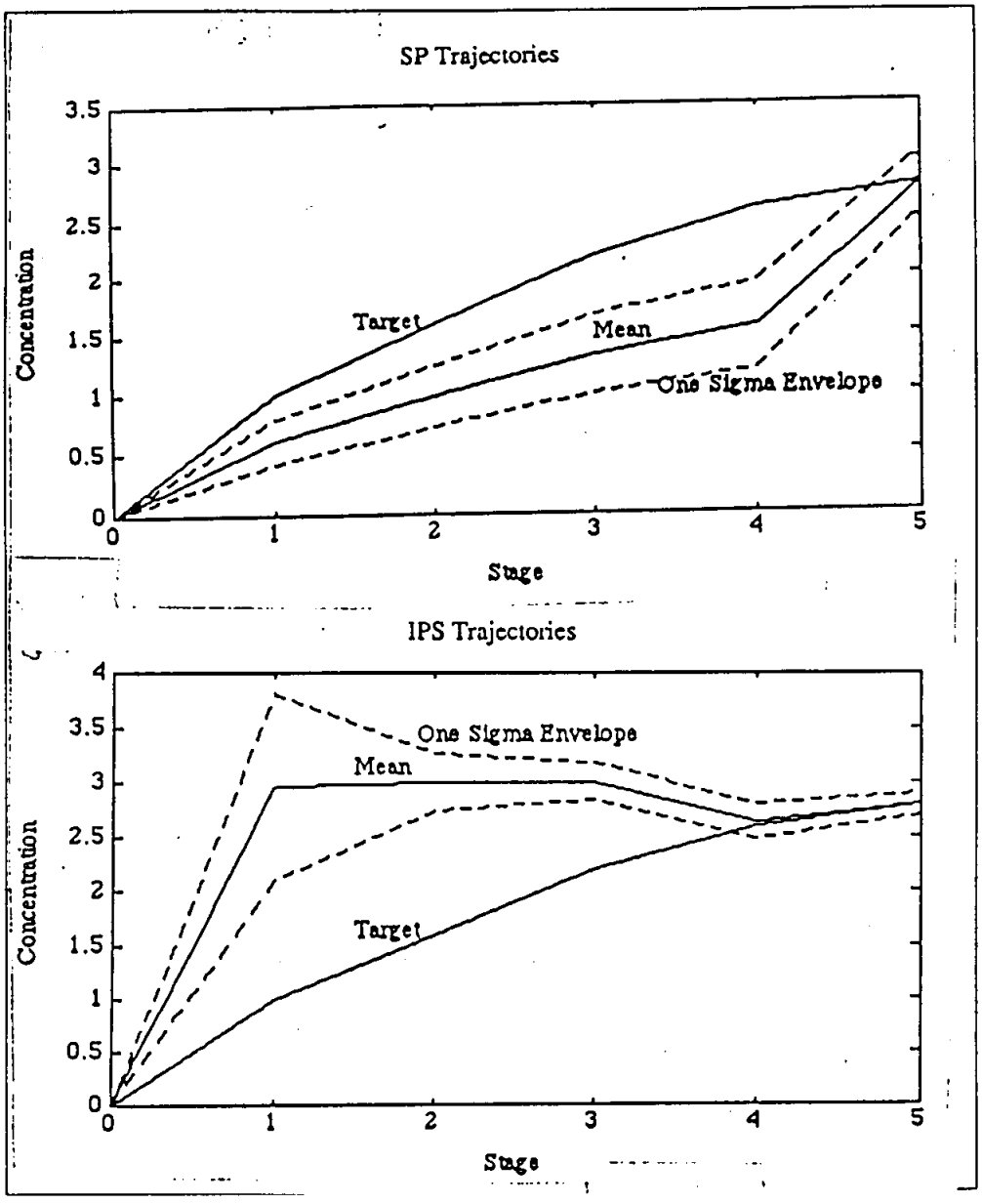

Fig. 2. *SP* and *IPS* trajectories.

with respect to the components of ϕ. The PK/PD vector ϕ is random with known probability density $p(\phi)$.

D-optimality and variants

The problem here is to design a sampling schedule so that the resulting estimate of ϕ provides as much "information" as possible. The most common measure of this information is the so-called "D-optimality" criterion; that is the criterion given in Eq. (8). Similar but not identical criteria occur when the $\det(M)$ in Eq. (9) is replaced by $\log \det(M)$. For the model of Eq. (30), it can be shown that the Fisher information matrix is:

$$M(\phi, U) = \sum_{n=0}^{N} \eta_n(\phi, s_n)^T \eta_n(\phi, s_n)/\sigma_n^2, \tag{31}$$

where $\eta_n(\phi, s) = \partial h_n(s, \phi)/\partial \phi = (\partial h_n(s, \phi)/\partial \phi_1, \ldots, \partial h_n(s, \phi)/\partial \phi_p)$ and $\phi = (\phi_1, \ldots, \phi_p)$.

In general, the objective is to maximize $E\{\Phi[M(\phi, U)]\}$ for some suitable function Φ. We now put this problem into the stochastic control framework developed above.

First define the "controls" $u_n = s_{n+1}$ with control constraints: $0 \leq u_0 \leq u_1 \leq \ldots \leq u_N \leq T$. The measurement equation corresponding to Eq. (2) is obtained by setting $y_n = y(s_n)$ and $v_n = \epsilon_n$ in Eq. (30):

$$y_n = h_n(u_n, \phi) + v_n \tag{32}$$

(Note that there are no "state" variables in this equation.)

The main purpose of the "state" equation is to get the cost function in the form of Eq. (10). To this end, define the "matrix" state variables $\{x_n\}$ by the state equation:

$$x_{n+1} = x_n + \eta(\phi, u_n)^T \eta(\phi, u_n), \qquad x_0 = 0 \qquad n = 0, 1, \ldots, N$$

(Note that the state equation is deterministic given ϕ.) If we define:

$$C(X, U, \phi) = -\Phi(x_{N+1})$$

then the resulting optimization problem becomes one of minimizing a "terminal cost":

$$J(U) = -\int \Phi(x_{N+1}) p(\phi) d\phi \tag{33}$$

over all admissible controls U.

In the PK/PD literature, the D-optimality problem (and its variants) have not been considered explicitly from such a stochastic control framework. However a number of previous works can be interpreted in this light. For example, in D'Argenio [27] and Walter-Pronzato [28], optimal "open loop" controllers are obtained; and in D'Argenio [29] an optimal "separation principle" controller is derived. (See also [30] for a recent survey on this subject.)

Considered as a terminal cost problem, the optimal sampling schedule design should be ideally suited for optimal policies. In [31] an iteration in policy space approach is suggested. A "natural" nominal admissible control policy is available. Namely:

$$U_0 \equiv (\mu_0, \ldots, \mu_N)$$

is such that, at each stage n, the remaining points $\{\mu_i, i \geq n\}$ are equally distributed in the open interval (μ_{n-1}, T), where $\mu_{-1} \equiv 0$.

It is hoped that this control perspective brings new insight to optimal sampling design.

ACKNOWLEDGMENTS

The writing of the paper benefited in many ways from many people; and the author wishes to express his sincere appreciation to David Bayard, David D'Argenio, Marianne Hubner, Roger Jelliffe, Mark Milman and Poornima Raghu. Further, this research was supported in part by National Institutes of Health grants RR01629 and P41-RR01861.

REFERENCES

1. A. .H. Jazwinski. Stochastic Processes and Filtering theory. Academic Press, New York, 1970.
2. A. Racine-Poon and A. F. M. Smith. Population models. In D. A. Berry and Marcel Dekker (Eds.), *Statistical Methodology in the Pharmaceutical Sciences*, New York, 1990, pp. 139–162.
3. J. L. Steimer, A. Mallet, and F. Mentre. Estimating interindividual pharmacokinetic variability. In M. Rowland *et al.* (Eds.) *Variability in Drug Therapy: Description, Estimation, and Control*, Raven Press, New York, 1985, pp. 65–111.
4. R. W. Jelliffe. A simulation study of factors affecting aminoglycoside therapeutic precision. In C. Cobelli and L. Mariani (Eds.), *Proc. First Symposium on Modeling and Control in Biomedical Systems*, Venice, Italy, 1988, pp. 86–88.
5. R. W. Jelliffe, A. Schumitzky, and M. Van Guilder. A simulation study of factors affecting aminoglycoside therapeutic precision, Technical Report: 90-3 Laboratory of Applied Pharmacokinetics, USC School of Medicine, Los Angeles, California, 1990.
6. D. P. Bertsekas. Dynamic Programming: Deterministic and Stochastic Models. Prentice Hall, Englewood Cliffs, N.J., 1987.
7. R. Bellman. Adaptive Control Processes: A Guided Tour. Princeton University Press, Princeton, N.J., 1961.
8. D. Bayard. A forward method for optimal stochastic control of nonlinear systems. In *Proc. 27th IEEE Conference on Decision and Control, Austin, Texas*, 1988. To appear: *IEEE Trans. Aut. Control.*
9. D. Bayard. Aspects of stochastic adaptive control synthesis. Doctoral Thesis, Electrical Engineering Department, State University of New York at Stony Brook, 1984.
10. Y. Bar-Shalom. Stochastic dynamic programming: caution and probing. *IEEE Trans. Aut. Control.*, AC-10:1184–1195 (1981).
11. D. Bayard and M. Eslami. Implicit dual control for general stochastic systems. *Opt. Control Applications Methods*, 6: 265–279 (1985).
12. W. F. Powers, P. H. Abbrecht, and D. G. Covel. Systems and microcomputer approach to anticoagulant therapy. *IEEE Trans. Biomed. Eng.*, BME 27:520–523 (1980).
13. S. Amrani, E. Walter, Y. Lecourtier, and R. Gomeni. Robust control of uncertain pharmacokinetic models. *Proc. IFAC 9th Triennial World Congress, Budapest, Hungary* 3079–3083 (1984).
14. A. Schumitzky, M. Milman, P. Khademi, and R. Jelliffe. Approximate optimal closed loop control of pharmacokinetic systems. In *Proc. IFAC Workshop on Decision Support for Patient Management: Measurement, Modeling, and Control*, British Medical Informatics Society, London, 1989, pp. 338–347.
15. J. Gaillot, J-L. Steimer, A. Mallet, J. Thebault, and A. Beider. A prior lithium dosage regimen using population characteristics of pharmacokinetic parameters. *J. Pharmacokin. Biopharm.* 7:579–628 (1979).
16. O. Richter and D. Reinhardt. Methods for evaluating optimal dosage regimens and their application to theophylline. *Int. J. Clin. Pharmacol. Ther. Toxicol.*. 20:564–575 (1982).
17. A. Mallet, F. Mentre, J. Giles, A. W. Kelman, A. H. Thomson, S. M. Bryson, and B. Whiting. Handling covariates in population pharmacokinetics with an application to gentamicin. *Biomed. Meas. Infor. Contr.*, 2:138–146 (1988).
18. D. Katz and D. Z. D'Argenio. Stochastic control of pharmacokinetic systems: Open loop strategies. In C. Cobelli and L. Mariani (Eds.), *Proc. First Symposium on Modelling and Control in Biomedical Systems*, Venice, Italy, 1988, pp. 560–566.
19. D. Z. D'Argenio and D. Katz. Implementation and evaluation of control strategies for individualizing dosage regimens with application to the aminoglycoside antibiotics. *J. Pharmacokin. Biopharm.* 14:523–37 (1986).
20. L. B. Sheiner, B. Rosenberg, and K. L. Melmon. Modelling of individual pharmacokinetics for computer-aided drug dosage. *Computers and Biomedical Research*, 5:441–459 (1972).
21. R. W. Jelliffe. Clinical applications of pharmacokinetics and control theory: planning, monitoring, and adjusting dosage regimens of aminoglycosides, lidocaine, digitoxin, and digoxin, In R. Maronde (Ed.), *Selected Topics in Clinical Pharmacology*, Springer-Verlag, New York, 1986, pp. 26–82.
22. R. W. Jelliffe, A. Schumitzky, and L. Hu, M. Liu. PC Computer programs for Bayesian adaptive control of drug dosage regimens. Technical Report: 90-5 Laboratory of Applied Pharmacokinetics, USC School of Medicine, Los Angeles, California, 1990.

23. S. Vozeh and J. L. Steimer. Feedback control methods for drug dosage optimization. *Clin. Pharmacokinetics*, 10:457–476 (1985).

24. A. Schumitzky. Stochastic control of pharmacokinetic systems. In R. Maronde (Ed.), *Selected Topics in Clinical Pharmacology and Therapeutics*, Springer-Verlag, New York, 1986, pp. 13–25.

25. A. Schumitzky. Adaptive control in drug therapy. In H. Ducrot *et al.* (Eds.) *Computer Aid to Drug Therapy and to Drug Monitoring*, Berne, Switzerland, March 6-10, 1978, North Holland, Amsterdam, 1978, pp. 357–360.

26. Y. Bar-Shalom and E. Tse. Concepts and methods in stochastic control. In C. Leondes (Ed), *Control and Dynamic Systems*, vol. 12, Academic Press, New York, 1976, pp. 99–172.

27. D. Z. D'Argenio. Incorporating prior parameter uncertainty in the design of sampling schedules for pharmacokinetic parameter estimation experiments. *Math. Biosciences*, 99:105–118 (1990).

28. E. Walter and L. Pronzato. Robust experimental design via stochastic approximation. *Math. Biosciences*, 75:103–120 (1985).

29. D. Z. D'Argenio. Optimal sampling times for pharmacokinetic experiments. *J. Pharmacokin. and Biopharm.*, 9:739–56 (1981).

30. L. Pronzato and E. Walter. Qualitative and quantitative experiment design for phenomenological models - A survey. *Automatica*. 26:195–213 (1990).

31. D. Bayard and A. Schumitzky. A stochastic control approach to optimal sampling schedule design. Technical Report: 90-1 USC/Laboratory of Applied Pharmacokinetics, Los Angeles, California, 1990.