Diffusion-driven Models for Physiological Processes

Konstantinos Kalogeropoulos[1], Nikolaos Demiris[2], and Omiros Papaspiliopoulos[3]

[1] Department of Engineering - Signal Processing, University of Cambridge, CB2 1PZ, Cambridge UK. (e-mail: kk384@cam.ac.uk)

[2] Medical Research Council Biostatistics Unit, CB2 0SR, Cambridge UK. (e-mail: nikolaos.demiris@mrc-bsu.cam.ac.uk)

[3] Department of Economics and Business, Universitat Pompeu Fabra, 08005, Barcelona Spain. (e-mail: omiros.papaspiliopoulos@upf.edu)

Abstract. Diffusion driven models for physiological processes provide stochastic extensions to existing models formulated via systems of ordinary differential equations. The system noise is modelled explicitly and disentangled from measurement error. Thus, the additional flexibility leads to parsimonious models and offers a natural prediction framework. We present a general inference framework through data augmentation, which encompasses inherent features of the problem such as measurement error and variability across individuals. The methodology is illustrated on simulated data.

Keywords. PK/PD models, Data augmentation, MCMC, Multilevel models, SDEs.

1 Introduction

The use of ordinary differential equations (ODEs) provides a natural modelling framework for a number of continuous time phenomena across several scientific disciplines. In this paper we focus our attention to pharmacokinetic/pharmacodynamic (PK/PD) models; see for example Racine-Poon and Wakefield (1998) and more general physiological processes, as in Huang et al. (2006). Under such formulations, the fitted model may be interpreted as a system of ODEs driven by the estimated parameter vector, whereas an error parameter is often employed to capture the deviations from the observations. However, as these deviations stem from multiple sources, such as measurement error, inherent system noise and structural mis-specifications, highly complex systems are often required to achieve a satisfactory fit. Models of this kind may not always be stable and the estimation of their parameters is often difficult, particularly when a likelihood-based approach is adopted. Furthermore, it is not straightforward to carry out prediction tasks in the absence of dynamic noise.

This paper explores stochastic versions of the above mentioned models to address these issues. The use of stochastic differential equations (SDEs) preserves the mean behavior of the model and naturally disentangles system noise from observation error. It also introduces a natural probability framework which may be used for model assessment and development. In other words, diffusion-driven models are expected to provide greater flexibility and lead to more parsimonious and stable model formulations. They also offer a natural prediction framework which allows for assessment of the uncertainty around the forecasts.

The task of likelihood-based inference on diffusion processes is particularly challenging and has received a remarkable amount of attention in the recent literature; Most available methodologies, see for example Beskos et al. (2006) and Ait-Sahalia (2007), take advantage of the Markov property and approach the likelihood through the transition density. However, under the observation regimes of the applications considered in this paper, the observed process is not Markov. More specifically, the observations may exist only for some of the diffusion components, are amenable to measurement error, may be irregularly spaced, and refer to functionals of the diffusion. Data augmentation schemes for diffusion processes in the spirit of Roberts and Stramer (2001), henceforth denoted as RS, may potentially cover all of these cases. Nevertheless, as noted in RS, likelihood reparametrisation is essential to avoid degenerate Markov chain Monte Carlo (MCMC) algorithms.

The outline of the paper is the following. In Section 2 we present a standard PK/PD model and a general diffusion extension of it. Section 3 provides an appropriate reparametrisation of the likelihood which is essential for an irreducible MCMC scheme, which is implemented on a simulated data example in Section 4. Section 5 concludes with some relevant discussion.

2   Stochastic Physiological Models

Consider a first order absorption elimination PK/PD model which could be represented by the graph of Figure 1, where the drug is administered to the patient (compartment A), absorbed to the compartment X with rate K_a, and eliminated with rate K_e. The available (noisy) observations are longitudinal drug concentrations (amount of drug relative to volume V) and correspond to the compartment X. The corresponding system of ODEs may be written as

$$dA_t = -K_a A_t dt, \ A_0 = Dose,$$
$$dX_t = \left(\frac{K_a A_t}{V} - K_e X_t \right) dt,$$

or else
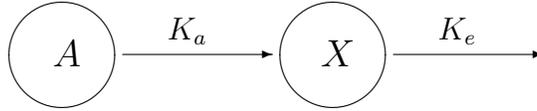
Fig. 1. Graphical compartmental representation of the a first order absorption elimination PK/PD model.

$$dX_t = \left(\frac{Dose\ K_a \exp(-K_a t)}{V} - K_e X_t \right) dt. \tag{1}$$

An SDE extension of (1) is the following

$$dX_t = \left(\frac{Dose\ K_a \exp(-K_a\ t)}{V} - K_e X_t \right) dt + \sigma X_t^{\gamma} dB_t, \tag{2}$$

where B_t is a standard Brownian motion. The volatility of the diffusion above (σX_t^{γ}) reflects the system noise. It contains a scale parameter σ and the parameter γ which determines the distribution X_t. For instance if $\gamma = 0$ the transition density of X is known to be Gaussian, whereas for $\gamma = 0.5$ it is a non-central $\mathcal{X}^2$ distribution (Cox et al., 1985). Alternative SDE formulations with additional Brownian motion components are also possible.

Another feature of such processes is their hierarchical structure due to the inherent differences across the different individuals. Typically, these are incorporated into the model through random effects on the model parameters. Let X_{it} denote a diffusion corresponding to individual i and $\mathbb{P}_{\theta_i}(X_{it})$ denote its distribution. As this distribution is not available in closed form, a data augmentation scheme may be employed to impute intermediate points that may be used for accurate likelihood approximations through Girsanov formula. These points are drawn from the conditional diffusion distribution on $x_{ij}, \mathbb{P}_{\theta_i}(X_{it}|xij)$. The parameters of X_{it}, denoted by θ_i, may or may not be subject specific. Let x_{ij} to be the true diffusion points of X_{it} corresponding to the noisy observations y_{ij} at times t_{ij}. A possible model, with Gaussian random effects on K_e and Gaussian additive measurement error, is shown below

$$y_{ij} \sim \mathcal{N}(x_{ij}, \tau^2) \tag{3}$$

$$x_{ij} \sim \mathbb{P}_{\theta_i}(X_{it}), \ \theta_i = (K_{ei}, K_a, V, \sigma, \gamma) \tag{4}$$

$$X_{it}|x_{ij} \sim \mathbb{P}_{\theta_i}(X_{it}|x_{ij}) \tag{5}$$

$$K_{ei} \sim \mathcal{N}(\mu_k, \sigma_k^2) \tag{6}$$

The choice of Bayesian inference through data augmentation provides a general framework which is also a natural choice for handling hierarchical models. Alternative methodologies for such models were given by Overgaard et al. (2005), Tornøe et al. (2005) through an algorithm based on an extended Kalman filter, and by Donnet and Samson (2006) through a stochastic EM algorithm. Both of these approaches were restricted to models of Gaussian system noise ($\gamma = 0$).

3   Likelihood Reparametrisation

In order to evaluate the likelihood we need to draw intermediate paths of X between x_{ij} from $\mathbb{P}_{\theta_i}(X_{it}|x_{ij})$. However, as the level of augmentation increases, the paths of X_{it} tend to have perfect correlation with the volatility parameters σ and γ. We proceed by applying the RS reparametrisation for each X_{it}, via the following 2-step transformation

1. $U_{it} = \sigma^{-1}(1 - \gamma)^{-1}X_{it}^{1-\gamma}$, if $\gamma \neq 1$, or
   $U_{it} = \sigma^{-1}\log(X_{it})$, if $\gamma = 1$.
2. $Z_s = \{(s - t_{ij-1})x_{ij} + (t_{ij} - s)x_{ij-1}\}(t_{ij} - t_{ij-1})^{-1}, \ \forall i, j.$

Under the RS framework, the likelihood parametrised in terms of x_{ij}, $Z_{it}|x_{ij}$, and parameters, leads to MCMC algorithms that do not degenerate as the level of augmentation increases.

The above reparametrisation may be viewed as an extension to the RS approach to examples of diffusions observed with error, cast in terms of the specific PK/PD context. The additional feature compared to alternative methodologies for diffusions with unobserved paths, such as the stochastic volatility formulations in Chib et al (2005) and Kalogeropoulos (2007), is the extra conditioning on the latent diffusion observations x_{ij}. Their presence ensures that the proposed diffusion paths are always close to the observations y_{ij} and results in substantially higher acceptance rates.

4   Application to simulated data

We simulated data from the model in (3), (4), (5), (6), with fixed gamma of 0.5, by using a high frequency Euler approximation scheme. The number of individuals and of observations per individual was 20 corresponding to a

realistic sample size. A single drug dose was assumed and each individual was observed up to a time horizon of 4 hours. To complete the model formulation, improper flat priors were assigned to the hyperparameters $K_a, V, \mu_k, \sigma_k, \sigma, \gamma$, and the diffusion paths of X_t were constrained to be strictly positive.

The parameters x_{ij} where updated with single site random walk metropolis steps. For the updates of the diffusion paths between x_{ij}, an independence sampler was used with Brownian bridge proposals. The acceptance rates were quite high, roughly 80%, to consider alternative proposals. The parameters V^{-1}, μ_k, σ_k, and τ^2 where updated with Gibbs steps, whereas for K_a and σ random walk Metropolis steps were used.

The level of augmentation is controlled by the number of imputed points between successive latent diffusion observations x_{ij}, denoted by m. There was no sign of increasing MCMC inefficiency in m in the relative autocorrelation plots of the volatility parameter σ. The convergence of the likelihood approximation was assessed through density plots of the log-likelihood which become virtually indistinguishable for an m of 80 and higher. Table 1 contains posterior summaries of the parameters. They appear to be in good agreement with the values the data where simulated from.

Parameter	True Value	Posterior mean	Posterior SD	Posterior median
K_a	1.0	0.996	0.038	0.996
V	1.4	1.340	0.047	1.339
μ_k	0.9	0.975	0.070	0.975
σ_k	0.3	0.265	0.057	0.259
σ	0.7	0.750	0.062	0.748
τ	1.0	1.057	0.085	0.057

Table 1. Summaries of the posterior draws for $m = 80$.

5 Discussion

The use of diffusion driven models in PK/PD applications and physiological processes provide stochastic extensions to existing formulations based on deterministic ODE systems, that are expected to be beneficial for inference and prediction purposes. In this paper we offer a data augmentation scheme that deals naturally with missing data, imbalanced designs, measurement error and heterogeneity across individuals. Our methodology may be used to sample from the posterior of the diffusion points, parameters and functionals thereof such as the area under the concentration vs time curve used extensively in bioequivalence studies.

Future steps include the higher dimensional SDE extension with more than one Brownian components on compartments that are indirectly observed, rather than observed with error. Another outstanding issue is the

joint estimation of σ and γ in the MCMC scheme, which will provide an implicit model averaging integrating out distributional assumptions on the system noise. Additional computational tools, such as the Metropolis adjusted Langevin algorithm, adaptive strategies, more sophisticated path proposals, are available in our disposal to construct more efficient and easy to implement MCMC algorithms.

References

Aït-Sahalia, Y. (2007). Closed form likelihood expansions for multivariate diffusions. To appear in the Annals of Statistics.

Beskos, A., Papaspiliopoulos, O., Roberts, G., and Fearnhead, P. (2006). Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3):333–382. With comments and a reply by the authors.

Chib, S., Pitt, M. K., and Shephard, N. (2005). Likelihood based inference for diffusion models. Submitted.

Cox, J. C., Ingersoll, J. E., and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, 53:385–407.

Donnet, S. and Samson, A. (2006). Parametric inference for mixed models defined by stochastic differential equations. To appear in ESAIM PS.

Huang, Y., Liu, D., and Wu, H. (2006). Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system. *Biometrics*, 62(2):413–423.

Kalogeropoulos, K. (2007). Likelihood based inference for a class of multidimensional diffusions with unobserved paths. *Journal of Statistical Planning and Inference*, 137:3092–3102.

Overgaard, R., Jonsson, N., Tornøe, C., and Madsen, H. (2005). Non-linear mixed-effects model with stochastic differential equations: Implementation of an estimation algorithm. *J Pharmacokinet. Pharmacodyn.*, 32:85–107.

Racine-Poon, A. and Wakefield, J. (1998). Statistical methods for population pharmacokinetic modelling. *Statistical Methods in Medical Research*, 7:63–84.

Roberts, G. and Stramer, O. (2001). On inference for partial observed nonlinear diffusion models using the metropolis-hastings algorithm. *Biometrika*, 88(3):603–621.

Tornøe, C., Overgaard, R., Agersø, H., Nielsen, H., Madsen, H., and Jonsson, N. (2005). Stochastic differetial equations in nonmem: implementation application and comparison with ordinary differential equations. *Pharm. Res.*, 22:1247–1258.