AN ADAPTIVE GRID NON-PARAMETRIC APPROACH TO PHARMACOKINETIC AND DYNAMIC (PK/PD) POPULATION MODELS

Technical Areas: Pharmacoinformatics, Decision Support Systems, Software Systems in Medicine

R. Leary[1], R. Jelliffe[2], A. Schumitzky[2], and M. Van Guilder[2]
[1]San Diego Supercomputer Center, University of California, San Diego, 9500 Gilman Drive, La Jolla CA 92903-0505, (858) 534-5123, fax: (858) 534-8303, leary@sdsc.edu
[2]Laboratory of Applied Pharmacokinetics, USC School of Medicine, 2250 Alcazar St, Los Angeles, CA 90033, USA, (323) 442-1300, jelliffe@hsc.usc.edu

AN ADAPTIVE GRID NON-PARAMETRIC APPROACH TO PHARMACOKINETIC AND DYNAMIC (PK/PD) POPULATION MODELS

ABSTRACT

Our NPEM software for non-parametric PK/PD population modeling employs the classical EM optimization algorithm to compute a maximum likelihood density on a fixed large multidimensional grid. The resulting support points and probabilities define a discrete distribution that approximates the maximum likelihood estimate of the true density, where the quality of the approximation is determined by the fineness of the grid. In order to achieve good resolution, a large number of grid points must be chosen, which can lead to high computational demands requiring a large-scale parallel supercomputer. Here we describe an improved method that uses a sequence of adaptively refined grids, as well as a new, state-of-the-art interior point algorithm for solving the associated maximum likelihood problem on each successive grid.

Together the combination of the adaptive grid idea with the interior point optimization algorithm is faster, often by factors of 100 or more, than the original NPEM method. Also, the new method requires much less memory, thus making many computations feasible on a PC or workstation that previously required supercomputer resources. Finally, the new algorithm easily and naturally accommodates the separation and simultaneous maximum likelihood estimation of both intra-individual and inter-individual variability, thus overcoming a major limitation of the original NPEM program.

INTRODUCTION

PK/PD models of the dose/response relationship for a given drug are often formulated in terms of differential equations describing the flow of the drug among various body compartments. Model parameters relating to flow rates and compartment volumes vary from individual to individual. Past experience with drugs can be stored in terms of the multidimensional probability distribution of the model parameters in the target population. Such a probability distribution can serve, for example, as the Bayesian prior to design the initial dose regimen for the next patient who appears to belong to that population.

This paper describes improvements in the statistical and computational methodology for non-parametric estimation of population probability distributions, given the data of past doses and responses (serum concentrations, etc.) from a sample of subjects in the target population. The advantages of non-parametric relative to parametric statistical estimation in the context of PK/PD population modeling are described in [1-3]. Here we simply note that non-parametric techniques are able to discover and quantitatively describe, for example, unsuspected and possibly genetically based sub-populations that can commonly give rise to multimodal population distributions. Such behavior cannot be captured by parametric techniques that rely on assumptions of unimodal normal or log-normal functional forms.

NONPARAMETRIC (NP) MAXIMUM LIKELIHOOD POPULATION MODELS

The NP approach to PK/PD modeling was introduced independently by Lindsay [2] and Mallet [3]. They emphasized use of maximum likelihood (ML) estimators, which have many optimal mathematical properties, including asymptotic consistency. This implies that

as sample size becomes large, the estimated population distribution converges to the true distribution. In can be shown under mild assumptions that the maximum likelihood distribution based on a sample of *N* subjects is in fact a discrete distribution with at most *N* support points in the model parameter space. The ML problem on the continuous multidimensional model parameter space can be approximated by gridding the feasible region of model parameter space, and then solving for the ML distribution on the finite grid using the classical EM (Expectation Maximization) algorithm, as developed for this case by Schumitzky[1]. This forms the basis for the NPEM algorithm currently implemented in the USC*PACK collection of PK/PD software [4].

As in the continuous case, an ML solution can be found on the grid with at most *N* support points. As the number of grid points increases, the solution on the grid will converge to the continuous solution. However, this convergence can be slow. For example, in a 5-dimensional model parameter space, fairly typical of many PK/PD population models, approximately 100^5 = 10 billion grid points are needed to resolve the grid to within 1% of the specified range of each parameter. Moreover, each doubling of the grid solution requires a increase in the number of grid points by a factor of 2^5=32. Thus the primary cost of the NPEM algorithm is usually not in the EM algorithm itself, but rather in the initial phase in which the differential equations must be solved for each sample subject and grid point.

THE ADAPTIVE GRID IMPLEMENTATION

Recently we have modified NPEM to implement a considerably more efficient way of reducing the discretization error than simply using the largest computationally feasible grid. In the modified approach, the original NPEM algorithm is applied to a modestly sized grid to produce an ML solution for that grid with at most *N* support points. A new small grid is then created, discarding all but the *N* optimal support points in the old grid and adding a modest number of additional points locally around each of these support points. The ML solution is then found on the refined grid, producing a new set of at most *N* optimal support points. The process is iterated while gradually reducing the size of the refinement regions around each successive set of optimal support points. This modified NPAG (Non-Parametric Adaptive Grid) algorithm uses far fewer total grid points than NPEM to achieve a given quality of solution. Moreover, we have replaced the EM algorithm that finds the ML solution on each grid with a more efficient convex primal-dual interior point optimization algorithm developed recently by Burke [5].

Together the combination of the adaptive grid and the improved interior point optimization technology have resulted in a great reduction in the computational resources, often by factors of 100 or more in both memory and CPU time, to achieve a given quality result. For example, a recent NPEM test computation on a 5-parameter 8-subject Michaelis-Menten model of pipericillin required over 2000 CPU processor-hours and 10 gigabytes of memory on the 1152-processor Blue Horizon IBM SP supercomputer at the San Diego Supercomputer Center. A slightly better quality solution was achieved in less than 3 hours with NPAG on a single processor 833 MHz IBM PC using 6 megabytes of memory.

SEPARATION OF INTRA-INDIVIDUAL AND INTER-INDIVIDUAL VARIABILITY

The non-parametric model parameter distribution function quantifies the inter-individual variability of the dose/response model. However, the apparent dose/response behavior within a given individual may also vary due do such external factors as serum assay error, dose and level timing errors, model mis-specification, as well as any internal physiological functional variability over the time scale of the measurements. NPEM requires the user to enter a fixed explicit intra-individual error model with a standard deviation that is a polynomial function of the response level. This polynomial plays a crucial role in the overall maximum likelihood optimization, and a poor error model may seriously degrade the quality of the non-parametric model parameter distribution. A better procedure is to include the parameters of the error model in the maximum likelihood computation itself, so they are in effect determined by the data. The structure of NPEM does not permit such a coupled optimization, and the burden this places on the user to input a reasonable error model polynomial is a major limitation of the NPEM approach. In contrast, NPAG allows the error parameters to be easily included in the sequential maximum likelihood optimization as that computation proceeds over the successively refined grids.

CONCLUSION

A new adaptive grid variant of the nonparametric NPEM PK/PD population modeling algorithm has been implemented which greatly enhances both its computational efficiency and functional capability. Many computations previously requiring supercomputer resources can now be performed on PCs. Additionally the scope of the underlying mathematical optimization has been enlarged to include separation and estimation of intra- and inter-individual variability.

Supported by NIH grants LM05401 and RR11526.

REFERENCES

1. Schumitzky A: Nonparametric EM Algorithms for Estimating Prior Distributions. Applied Math. and Computation, 45: 143-157, 1991.
2. Lindsay B: The Geometry of Mixture Likelihoods: a General Theory. Ann. Statist. 11: 86-94, 1983.
3. Mallet A: A Maximum Likelihood Estimation Method for Random Coefficient Regression Models. Biometrika 73: 645-656, 1986.
4. Jelliffe R, Schumitzky A, Van Guilder M, and Jiang F: User Manual for Version 10.7 of the USC*PACK Collection of PC programs, USC Laboratory of Applied Pharmacokinetics, December 1, 1995.
5. Burke J: personal communication, 2000.