Available online at www.sciencedirect.com
ScienceDirect
Computational Statistics & Data Analysis 51 (2007) 6614–6623
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

Nonlinear random effects mixture models: Maximum likelihood estimation via the EM algorithm

Xiaoning Wang[a], Alan Schumitzky[b], David Z. D'Argenio[a,*]

[a]*Department of Biomedical Engineering, University of Southern California, 1042 Downey Way, DRB 140, Los Angeles, CA 90089, USA*
[b]*Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA*

Received 29 September 2006; accepted 9 March 2007
Available online 13 March 2007

Abstract

Nonlinear random effects models with finite mixture structures are used to identify polymorphism in pharmacokinetic/pharmacodynamic (PK/PD) phenotypes. An EM algorithm for maximum likelihood estimation approach is developed and uses sampling-based methods to implement the expectation step, that results in an analytically tractable maximization step. A benefit of the approach is that no model linearization is performed and the estimation precision can be arbitrarily controlled by the sampling process. A detailed simulation study illustrates the feasibility of the estimation approach and evaluates its performance. Applications of the proposed nonlinear random effects mixture model approach to other population PK/PD problems will be of interest for future investigation.

Keywords: Finite mixture models; Mixed effects models; Pharmacokinetics/pharmacodynamics

1. Introduction

There is substantial variability in the way individuals respond to medications, in both treatment efficacy and toxicity. The sources of a drug's underlying pharmacokinetic (PK) and pharmacodynamic (PD) variability can include demographic factors (such as age, sex, weight), physiological status (such as renal, liver, cardiovascular function), disease states, genetic differences, interactions with other drugs and environmental factors. In their seminal work, Sheiner et al. (1972) proposed a parametric nonlinear mixed-effects modeling framework for quantifying both within and between subject variability in a drug's PKs, and developed an approximate maximum likelihood solution to the problem. Since the introduction by Beal and Sheiner (1979) of the general purpose software package NONMEM implementing this approach, other approximate maximum likelihood algorithms have been introduced to solve the nonlinear random and mixed effects modeling problem (see Davidian and Giltinan, 1995 for an extensive review). An exact maximum likelihood (i.e., no linearization) solution to the parametric population modeling problem based on the EM algorithm has also been proposed by Schumitzky (1995) and fully developed and implemented by Walker (1996). The population modeling framework has had a significant impact on how PK (and PD) variability is quantified and studied during drug

* Corresponding author. Tel.: +1 213 740 0341; fax: +1 213 740 0343.
E-mail address: dargenio@bmsr.usc.edu (D.Z. D'Argenio).

0167-9473/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2007.03.008

development, and on the identification of important covariates associated with a drug's inter-individual kinetic/dynamic variability.

While population models incorporating measured covariates have proven to be useful in drug development, it is recognized that genetic polymorphisms in drug metabolism and in the molecular targets of drug therapy, for example, can also have a significance influence on the efficacy and toxicity of medications (Evans and Relling, 1999). There is, therefore, a need for population modeling approaches that can extract and model important subpopulations using PK/PD data collected in the course of drug development trials and other clinical studies, in order to help identify otherwise unknown genetic determinants of observed PK/PD phenotypes. The nonparametric maximum likelihood approach for nonlinear random effects modeling developed by Mallet (1986), as well as the nonparametric Bayesian approaches of Wakefield and Walker (1997) and Rosner and Muller (1997), and the smoothed nonparametric maximum likelihood method of Davidian and Gallant (1993) all address this important problem. In this paper we propose a parametric approach using finite mixture models to identify subpopulations with distinct PK/PD properties.

An EM algorithm for exact maximum likelihood estimation of nonlinear random effects finite mixture models is introduced, extending the previous work of Schumitzky (1995) and Walker (1996). The EM algorithm has been used extensively for linear mixture model applications (see McLachlan and Peel, 2000 for a review). The algorithm for nonlinear mixture models presented below has an analytically tractable M step, and uses sampling-based methods to implement the E step. Section 2 of this paper describes the finite mixture model within a nonlinear random effects modeling framework. Section 3 gives the EM algorithm for the maximum likelihood estimation of the model. Section 4 addresses individual subject classification, while an error analysis is presented in Section 5. A detailed simulation study of a PK model is presented in Section 6. Section 7 contains a discussion.

2. Nonlinear random effects finite mixture models

A two-stage nonlinear random effects model that incorporates a finite mixture model is given by

$$\mathbf{Y}_i|\boldsymbol{\theta}_i, \boldsymbol{\beta} \sim N(\mathbf{h}_i(\boldsymbol{\theta}_i), \mathbf{G}_i(\boldsymbol{\theta}_i, \boldsymbol{\beta})), \quad i = 1, \ldots, n \tag{1}$$

and

$$\boldsymbol{\theta}_i, \ldots, \boldsymbol{\theta}_n \sim_{i.i.d} \sum_{k=1}^{K} w_k N(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), \tag{2}$$

where $i = 1, \ldots, n$ indexes the individuals and $k = 1, \ldots, K$ indexes the mixing components.

At the first stage represented by (1), $\mathbf{Y}_i = (y_{1i}, \ldots, y_{m_i i})^{\mathrm{T}}$ is the observation vector for the ith individual ($\mathbf{Y}_i \in R^{m_i}$); $\mathbf{h}_i(\boldsymbol{\theta}_i)$ is the function defining the PK/PD (PK/PD) model, including subject specific variables (e.g., drug doses), and $\boldsymbol{\theta}_i$ is the vector of model parameters (random effects) ($\boldsymbol{\theta}_i \in R^p$). In (1) $\mathbf{G}_i(\boldsymbol{\theta}_i, \boldsymbol{\beta})$ is a positive definite covariance matrix ($\mathbf{G}_i \in R^{m_i \times m_i}$) that may depend upon $\boldsymbol{\theta}_i$ as well as on other parameters $\boldsymbol{\beta}$ (fixed effects) ($\boldsymbol{\beta} \in R^q$).

At the second stage given by (2), a finite mixture model with K multivariate normal components is used to describe the population distribution. The weights $\{w_k\}$ are nonnegative numbers summing to one, denoting the relative size of each mixing component (subpopulation), for which $\boldsymbol{\mu}_k (\boldsymbol{\mu}_k \in R^p)$ is the mean vector and $\boldsymbol{\Sigma}_k (\boldsymbol{\Sigma}_k \in R^{p \times p})$ is the positive definite covariance matrix.

Letting ϕ represent the collection of parameters, $\{\boldsymbol{\beta}, (w_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), k = 1, \ldots, K\}$, the population problem involves estimating ϕ given the observation data$\{\mathbf{Y}_1, \ldots, \mathbf{Y}_n\}$. The maximum likelihood estimate (MLE) can be obtained by maximizing the overall data likelihood L with respect to ϕ. Under the $i.i.d.$ assumption of the individual parameters $\{\boldsymbol{\theta}_i\}$, L is given by the expression

$$L(\phi) = \prod_{i=1}^{n} \int p(\mathbf{Y}_i|\boldsymbol{\theta}_i, \boldsymbol{\beta}) \sum_{k=1}^{K} w_k p(\boldsymbol{\theta}_i|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \, \mathrm{d}\boldsymbol{\theta}_i. \tag{3}$$

The MLE of ϕ is defined as ϕ_{ML} with $L(\phi_{ML}) \geqslant L(\phi)$ for all ϕ in the parameter space.