Stochastic models

Growth models and some estimation methods

Adeline Leclercq Samson

Laboratoire Jean Kuntzmann, Université Joseph Fourier, Grenoble
Growth data

- Data measured in (pre)-clinical trials
  - Height, weight of subjects
  - Tumor volume, tumor size
  - Circulating biomarkers

- Longitudinal data
  - Several subjects \( i = 1, \ldots, n \)
  - Repeated measures at times \( t_{ij}, \ j = 0, \ldots, J \)
  - Observations \( y_{ij} \) at time \( t_{ij} \)
  - Measurement noise

- Mixed-effect models

The FDA has recommended the use of mixed-effect models to analyze longitudinal data of tumor response to treatment
Mixed-effect model

- Repeated observations:

\[ y_{ij} \text{ at time } t_{ij}, \quad i = 1, \ldots, n, \quad j = 0, \ldots, J \]

- Standard regression model

\[
\begin{align*}
y_{ij} &= f(\phi_i, t_{ij}) + g(\phi_i, t_{ij})\varepsilon_{ij} \\
\varepsilon_{ij} &\sim_{iid} \mathcal{N}(0, \sigma^2) \\
\phi_i &\sim_{iid} \mathcal{N}(\mu, \Omega)
\end{align*}
\]

- \( f \): parametric regression function
- \( \phi_i \): "biological" or "physiological" random parameters
- \( g \): error model [homoscedastic \( g = 1 \) or heteroscedastic \( g = f \)]

- Parameters to be estimated

\[ \theta = (\mu, \Omega, \sigma) \]
Likelihood

- Notations
  - $y_i = y_{i0:J} = (y_{i0}, \ldots, y_{iJ})$: data vector of subject $i$
  - $y = (y_1, \ldots, y_n)$: global data vector

- Likelihood function

\[
L(y; \theta) = \prod_{i=1}^{n} p(y_i; \theta) = \prod_{i=1}^{n} \int p(y_i, \phi_i; \theta) d\phi_i
\]

\[
= \prod_{i=1}^{n} \int p(y_i|\phi_i; \theta) p(\phi_i; \theta) d\phi_i
\]

⇒ If $f$ not linear with respect to $\phi_i$, likelihood not explicit
Maximum likelihood estimation

- Approximation of the likelihood
  - Linearization of the likelihood [Pinheiro and Bates, 2000]

- Numerical computation of the likelihood
  - Gaussian quadrature [Davidian and Giltinan, 1995; Guedj et al, 2007; Picchini et al, 2010]
  - Monte Carlo EM algorithm (MCEM) [Wei and Tanner, 1991]
  - Stochastic Approximation EM algorithm (SAEM) [Kuhn and Lavielle, 2005]

- Bayesian approach
  - Prior choice on $\theta$
  - MCMC algorithms [Spiegelhater et al, 1992]
  - Posterior distribution $p(\theta|y)$
Gaussian quadrature methods

[Davidian and Giltinan, 1995; Guedj et al, 2007; Picchini et al, 2010]

- Gauss-Hermite quadrature of order $R$
  - Individual likelihood
    \[ L(y_i; \theta) = \int p(y_i|\phi_i; \theta)p(\phi_i; \theta)d\phi_i \]
  - Approximation
    \[ L^G(y_i; \theta) = \sum_{r=1}^{R} \pi_r p(y_i|\omega \sqrt{2r}z_r/\mu; \theta) \]
    - $z_r, r = 1, \ldots, R$ zeros of the Hermite polynomials $H_R(\cdot)$ of degree $R$
    - $\pi_r$ adequate weights
  - Convergence of $L^R$ to the true likelihood when $R$ tends to infinity

- Software: SAS
Stochastic Approximation EM (SAEM) Algorithm

[Denpster, Laird, Rubin, 1977; Delyon, Lavielle and Moulines, 1999; Kuhn, Lavielle, 2005]

- SAEM algorithm
  - **E step**
    - **S step**: simulation of \((\phi_m)\) under distribution \(p(\phi|y; \hat{\theta}_m)\) with MCMC algorithm
  - **SA step**: approximation of
    \[
    Q_{m+1}(\theta) = \mathbb{E} \left[ \log p(y, \phi; \theta) | y, \hat{\theta}_m \right]
    \]
    with a stochastic approximation scheme of step size \(\alpha_m\)
    \[
    Q_{m+1}(\theta) = (1 - \alpha_m)Q_m(\theta) + \alpha_m \log p(y, \phi_m; \theta)
    \]
  - **M step**: update of \(\hat{\theta}_m\)
    \[
    \hat{\theta}_{m+1} = \arg \max_\theta Q_{m+1}(\theta)
    \]
  - Convergence of \(\hat{\theta}_m\) to the maximum likelihood estimator

- Software: Monolix

A. Samson Stochastic models Workshop, Bordeaux, 11/10/2013 7 / 29
Growth models (regression function $f$)

- Several classes of models
  - Standard growth functions
    - Logistic, Gompertz, Richards, Weibull [Zimmerman and Nunez-Anton, 2001]
    - Monotone increase
  - Phenomenological models
    - System theory [Wiener, 1948; Bertalanffy, 1960; Bastogne et al; 2009]
    - Holistic representation, black-box models
  - Mechanistic models
    - Partial differential equations [Ribba, Colin, Schnell 2006; Colin et al, 2013; Lagaert PhD]
    - Dynamic of angiogenesis
Growth data

[Donnet, Foulley, Samson, 2010]

- Population
  - 50 pigs
  - 11 weight measures per subject

- Gompertz function
  - \( f(\phi, t) = Ae^{-Be^{-Ct}} \)
  - \( \phi = (A, B, C) \)

Population prediction \( f(\mu, t) \), Individual prediction \( f(\phi_i, t) \)
Growth inhibition study

[Bastogne, Samson et al, 2010]

- **Data**
  - **Population**
    - 96 Female mice, with tumor implantation
    - Treatments: no treatment (NT) or radiotherapy (RT) or concomitant radiochemotherapy (RCT) or photodynamic therapy (PDT)
  - **Measurements**
    - $v(t)$ tumor volume at time $t$
    - $y_{ij} = 3 \sqrt{\frac{6v(t)}{\pi}}$

- **Linear-Exponential-Linear Model (phenomenological)**

\[
f(t, \phi) = s_0 \left[ \underbrace{1 + at}_{natural growth} - bt - k_2 T \left(1 - e^{-\left(t-\tau\right)/T}\right) \mathbf{1}_{t \geq \tau} - k_3 \left(t - \tau\right) \mathbf{1}_{t \geq \tau} \right]
\]

(1)
Evaluation of the treatments

- Treatment effects
  - PDT on transient decrease
  - Duration
  - Dose of radiation

---

(a) Radiotherapy responses

(b) Concomitant radiochemotherapy responses

(c) Photodynamic therapy responses
Hypothesis

- Angiogenesis inhibitors could contribute to normalize vasculature

Data

- Sunitinib oral molecule
- 30 subjects
- 27 points per subject up to 100 days
Model

\[ \frac{dS}{dt} = -k_S S \]

\[ \frac{dV}{dt} = \lambda V \left( 1 - \frac{V}{K} \right) \]

\[ \frac{dK}{dt} = bV^\gamma - \beta k_s S K \]

- 7 parameters: \( k_S, \lambda, b, \gamma, \beta, V_0, K_0 \)
## Results

### Treatment effects

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value (s.e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_S$</td>
<td>2.12 (-)</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.02 (4%)</td>
</tr>
<tr>
<td>$b$</td>
<td>0.00168 (4%)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>2 (-)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.0237 (9%)</td>
</tr>
<tr>
<td>$V_0$</td>
<td>1.76 (7%)</td>
</tr>
<tr>
<td>$K_0$</td>
<td>7.43 (1%)</td>
</tr>
</tbody>
</table>

### Predictions

![Graph showing predicted vs observed tumor diameter](image-url)
Limits of complex growth models

- Complex deterministic models
  - Ordinary differential equations
    - Large number of equations: difficulty to solve the system with discrete numerical scheme. Example: [Lignet et al, 2013] 37 equations, 78 parameters
    - Large number of parameters, some of them have to be fixed
  - Partial Differential Equations
    - Few parameters
    - Computationally intensive to obtain one realization of the solution

- Alternative: Stochastic models
  - Lot of noise for some biomarkers (different from measurement noise)
  - Reduction of the dimension of the system [Mortensen et al, 2007; Donnet and Samson, 2013]
  - Introduction of a stochastic part to "absorb" all details that are not modeled
Stochastic Differential Equation

- **SDE in biology**
  - Pharmacocinetics [Ditlevsen et al, 2005; Ditlevsen, Samson, 2013; Donnet, Samson, 2013]
  - Neurobiology [Hopfner and Broda, 2005; Picchini et al, 2008; Ditlevsen, Samson, 2013]
  - Growth [Donnet et al, 2010]

- **New source of variability**
  - Variability around the deterministic model: what is not modeled by the deterministic part
  - Within-subject variability: variability in time
Variability around the deterministic model

- Ordinary differential equation
  - $dX_t = \left( -\frac{X_t}{\tau} + \phi \right) \, dt$
  - Deterministic solution

![Graph showing depolarization over time](image-url)
Variability around the deterministic model

- Stochastic differential equation
  - \( dX_t = \left( -\frac{X_t}{\tau} + \phi \right) dt + \sigma dB_t \) with \( B_t \) a Brownian motion
  - Stochastic solution
Variability around the deterministic model

- Stochastic differential equation
  - $dX_t = \left( -\frac{X_t}{\tau} + \phi \right) dt + \sigma dB_t$ with $B_t$ a Brownian motion
- Stochastic solution
Variability around the deterministic model

- Stochastic differential equation
  - $dX_t = \left(-\frac{X_t}{\tau} + \phi \right) dt + \sigma dB_t$ with $B_t$ a Brownian motion
- Stochastic solution
Stochastic Differential Equation with random parameters

\[ dX_t = a(X_t, \phi)dt + b\gamma(X_t, \phi)dB_t, \quad X_0 = x_0 \]

- SDE with random parameters

\[ dX_{it} = a(X_{it}, \phi_i)dt + b\gamma(X_{it}, \phi_i)dB_{it}, \quad X_{i0} = x_0 \]

with
- \((\phi_i)\) random variables
- \((B_{it})\) independent brownian motions

- SDE mixed models

\[ y_{ij} = X_{itij} + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim iid \mathcal{N}(0, \sigma^2) \]
\[ dX_{it} = a(X_{it}, \phi_i)dt + b\gamma(X_{it}, \phi_i)dB_{it}, \quad X_{i0} = x_0 \]
\[ \phi_i \sim iid \mathcal{N}(\mu, \Omega) \]

Parameters to be estimated: \(\theta = (\mu, \Omega, \gamma, \sigma)\)
SDE mixed models

\[ y_{ij} = X_{it_{ij}} + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim_{iid} \mathcal{N}(0, \sigma^2) \]
\[ dX_{it} = a(X_{it}, \phi_i)dt + b\gamma(X_{it}, \phi_i)dB_{it}, \quad X_{i0} = x_0 \]
\[ \phi_i \sim_{iid} \mathcal{N}(\mu, \Omega) \]

Notations
- \( X_i = X_i,0:J = (X_{t_{i0}}, \ldots, X_{t_{ij}}) \): hidden diffusion of subject \( i \)
- \( X = (X_1, \ldots, X_n) \)

Likelihood for subject \( i \)

\[ p(y_i; \theta) = \int p(y_i, X_i, \phi_i; \theta) dX_i d\phi_i \]
\[ = \int p(y_i|X_i; \theta)p(X_i|\phi_i; \theta)p(\phi_i; \theta)dX_i d\phi_i \]
Likelihood

Girsanov formula gives

\[ p(X_t; \theta) = \int \exp \left( \int \frac{a(X_i(s), \phi_i)}{b^2(X_i(s), \phi_i)} \, dX_i(s) - \frac{1}{2} \int \frac{a^2(X_i(s), \phi_i)}{b^2(X_i(s), \phi_i)} \, ds \right) p(\phi_i; \theta) \, d\phi_i \]

But explicit only for linear drift and known volatility

Alternative: discretization of the SDE

\[ p(y_t; \theta) = \int \int p(y_i|X_i; \theta)p(X_i|\phi_i; \theta)p(\phi_i; \theta) \, dX_i \, d\phi_i \]

\[ = \int \int \prod_{j=0}^{J} p(y_{ij}|X_{t_{ij}}; \theta) \]

\[ \times \prod_{j=1}^{J} p(X_{t_{ij}}|X_{t_{ij-1}}, \phi_i; \theta)p(\phi_i; \theta) \, dX_i \, d\phi_i \]
Estimation methods based on approximations

- Approximation of the conditional distribution
  - Extended Kalman filter
  - Stochastic or deterministic maximisation algorithms
    - [Tornoe et al 2005; Overgaard et al 2005; Delattre and Lavielle, 2013]

- Approximation of the likelihood
  - Gaussian quadrature [Picchini et al, 2010]
  - Laplace approximation [Picchini and Ditlevsen, 2011]
  - Hermite expansion of the transition density if needed [Picchini et al, 2010]
  - Simulation of the hidden SDE [Donnet and Samson, 2008; Donnet et al, 2010; Donnet, Samson, 2013]
Estimation methods based on a simulation step

- Estimation algorithms
  - Bayesian [Donnet, Foulley, Samson, 2010]
  - SAEM [Donnet and Samson, 2008; Donnet, Samson, 2013]

- Simulation step
  - For $i = 1, \ldots, n$, simulation of
    \[
    (X_{im}, \phi_{im}) \sim p(X_{i,0:J}, \phi_i|y_{i,0:J}; \hat{\theta}_m)
    \]
  - Gibbs algorithm
    - $p(\phi_i|y_{i,0:J}, X_{i,0:J}; \hat{\theta}_m)$: standard by Metropolis-Hastings
    - $p(X_{i,0:J}|y_{i,0:J}, \phi_i; \hat{\theta}_m)$: block decomposition and iterative simulation
    - $\Rightarrow$ slow convergence of the chain
  - Particle filter coupled with MCMC (PMCMC)
    - [Del Moral et al, 2001; Doucet et al, 2001; Chopin, 2004; Andrieu et al, 2010]
    - Metropolis Hastings algorithm targeting directly $p(X_{i,0:J}, \phi_i|y_{i,0:J}; \hat{\theta}_m)$
Improvement of the predictions

Subject 1

Subject 4

ODE

SDE
Meta-models: another alternative

- Complex mechanistic models
  - Approximation of the solution of the model at each iteration of the estimation method
  - Partial differential equation solution: difficult to obtain

- Meta-model
  - Precomputation on a pre-defined grid
    - Precise evaluation of the solution on the points of the grid
  - Approximation
    - Nearest neighborhood approximation [Barthelemy, Lavielle, submitted]
    - Linear approximation on the grid [Grenier, Louvet, Vigneaux, submitted]
  - Computational cost gain (PDE example) [Grenier, Louvet, Vigneaux, submitted]
    - Exact SAEM: 23 days
    - Interpolation with heterogeneous grid: 26 min
Conclusion

- **Mechanistic models**
  - Mechanistic description
    - Angiogenesis dynamic
    - Action of molecules
  - Good fitting on real data
  - Need powerful statistical methods because large number of parameters/random effects

- **Stochastic models**
  - Advantages
    - Less parameters, less equations
    - Allow stochastic individual variations around the deterministic theoretical model
    - Improvement of the predictions
  - Need specific statistical tools to filter the stochastic process
  - Need large number of data → pre-clinical data
Perspectives

- Meta-models
  - Nearest neighborhood or linear approximation
    - Estimation on a approximated model
    - Convergence to an approximate MLE
  - More sophisticated meta-models
    - Gaussian process, Reproducing Kernel Hilbert space (RKHS), ...
    - Convergence to the exact MLE joint work with P. Barbillon and C. Barthelemy

- Non parametric estimation
  - Density of the random effects [Comte et al, 2012]
  - Drift function of the stochastic model [Cattiaux et al, submitted]