Stochastic approximation EM algorithm in nonlinear mixed effects model for viral load decrease during anti-HIV treatment

Adeline Samson[1], Marc Lavielle[2] and France Mentré[1]

[1] INSERM E0357, Department of Epidemiology, Biostatistics and Clinical Research, University Hospital Bichat-Claude Bernard, Paris, France

[2] University Paris-Sud, Orsay, France

Inserm
Institut national de la santé et de la recherche médicale

Estimation in nonlinear mixed effects models (NLMEM)

by maximum likelihood

- No close form of the likelihood because of the non linearity of the model $\Rightarrow$ estimation by maximum likelihood rather delicate
- Usual estimation software: approximation of the likelihood by linearization of the model (FO, FOCE, ...)
  - NONMEM software, **`nlme`** in Splus/R, Proc SAS NLMIXED,...
- Inconsistent estimates produced by these algorithms based on linearization except if both N $\longrightarrow \infty$ and $n_i \longrightarrow \infty$ (Vonesh, *Biometrika*, 1996)
- Inflation of the type I error of the Likelihood Ratio Test and Wald Test using the algorithms based on linearization (Wählby et al., *J Pharmacokinet Pharmacodyn*, 2001; Comets and Mentré, *J Biopharm Stat*, 2001; Ding and Wu, *Stat Med*, 2001; Panhard and Mentré, *Stat Med*, 2004)
- Hypothesis: problems due to linearization

EM algorithms

- Alternative method for maximum likelihood estimation to avoid the linearization (Dempster et al., *JRSSB*, 1977)
- Individual parameters considered as missing data
- Iterative algorithm
  1. E step: expectation of the missing data
     - No analytical form of the distribution of the individual parameters conditionally to the observations and the hyperparameters
     - Conditional expectation of the log-likelihood of the complete data difficult to compute, even when the individual parameters are known
  2. M step: maximization of the likelihood of the complete data
     - Analytically
     - Newton-Raphson algorithm

E step for the NLMEM

- Linearization (Mentré and Gomeni, *J Biopharm Stat*, 1995)
- Simulation of the individual parameters in their unknown distribution
  - Monte Carlo method (MCEM Algorithm) (Walker, *Biometrics*, 1996)
  - Metropolis Hastings procedure (Quintana, Liu and Del Pino, *CSDA*, 1999)
  - Monte Carlo Markov Chain procedure (MCMC) (Gu and Kong, *PNAS*, 1998, Kuhn and Lavielle *PAGE* 2003, *ESAIM PS*, 2004)
- Approximation of the conditional expectation
  - Monte Carlo EM (Wei and Tanner, *JASA*, 1990)
  - Stochastic Approximation EM (SAEM) (Delyon, Lavielle and Moulines, *Annals Stat*, 1999)

$\Rightarrow$ Kuhn and Lavielle (*PAGE* 2003, *ESAIM PS* 2004) proposed to combine SAEM and MCMC

Contents

1. Evaluation of the SAEM algorithm by simulation
2. Estimation of the likelihood by importance sampling
3. Evaluation of the type I error of a LRT
4. Illustration on real data : TRIANON

Models and notations

- Statistical model: $y_{ij} = f(\theta_i, t_{ij}) + \varepsilon_{ij}$
  - y_{ij}: measurement of subject i $(i = 1, \ldots, N)$ at time t_{ij} $(j = 1, \ldots, n_i)$
  - θ_i: p-vector of the parameters of subject i
  - ε_{ij}: measurement error of subject i at time t_{ij}
  - $\varepsilon_{ij}|\theta_i$ homoscedastic or heteroscedastic model
  - $\theta_i = \mu + b_i$ with $b_i \sim \mathcal{N}(0, \Omega)$
  - ψ: vector of the hyperparameters of the model
- Log-likelihood: $L_{obs}(\psi \,|\, y) = \sum_{i=1}^{N} log\,(p(y_i|\psi\,))$

  with $p(y_i|\psi\,) = \int p(\,y_i|\,\theta_i,\,\psi\,)p(\,\theta_i|\,\psi\,)\,d\theta_i$

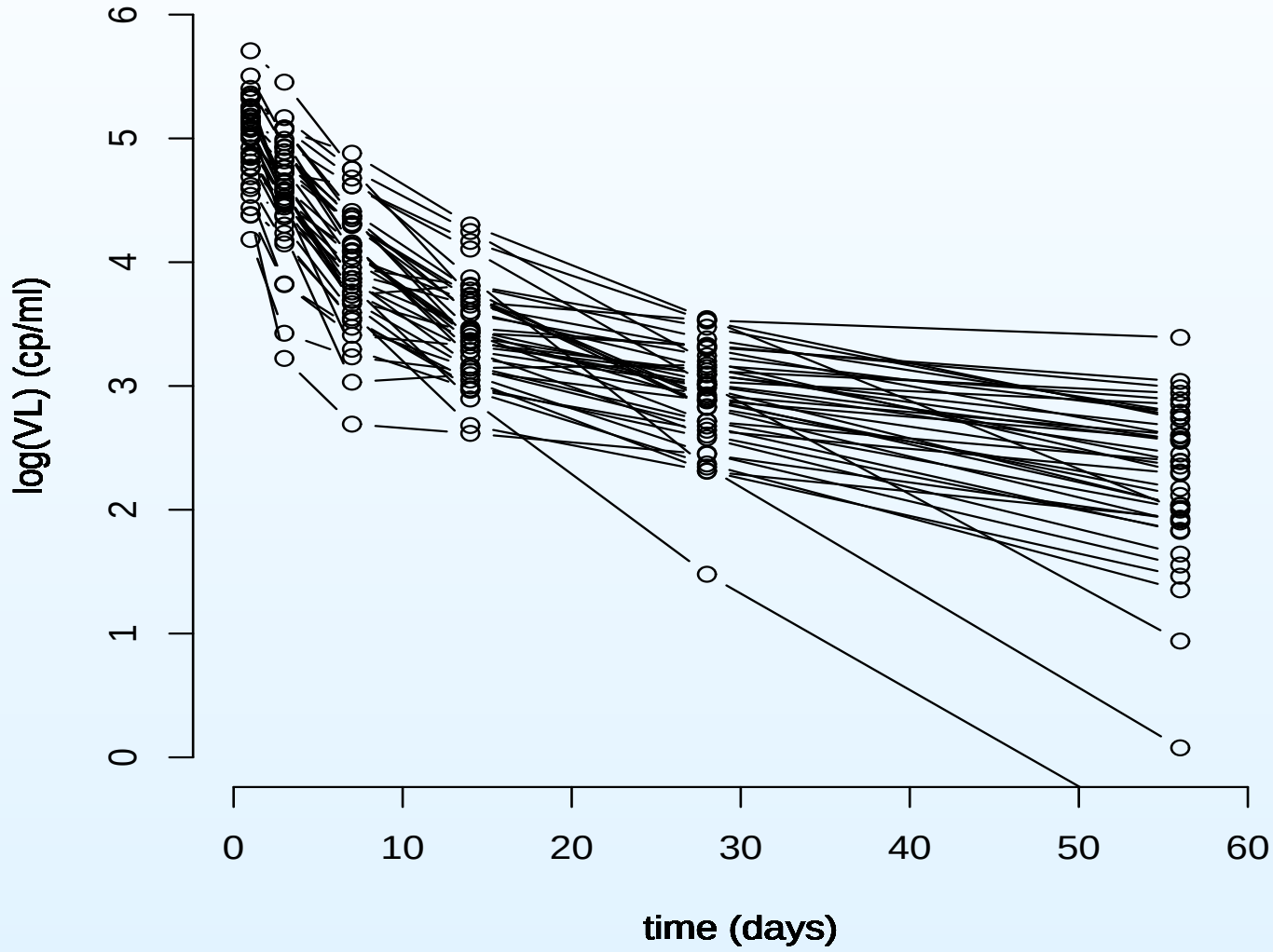
1.1 Evaluation of SAEM: simulation example

- Model of the decrease of viral load after beginning of anti HIV treatment (Ding and Wu, *Stat Med*, 2001)

$$f(\theta_i, t_{ij}) = \log_{10}(P_{1,i}\exp(-\lambda_{1,i}\, t_{ij}) + P_{2,i}\exp(-\lambda_{2,i}\, t_{ij}))$$

- Four parameters: $\ln P_1$, $\ln P_2$, $\ln\lambda_1$, $\ln\lambda_2$
- Additive random effects ($\omega_k^2 = 0.3$, $k = 1, \ldots, 4$)
- Additive error with constant variance ($\sigma = 0.065$)
- Six identical sampling times: 1, 3, 7, 14, 28 and 56 days
- $N = 40$ and $N = 200$

1.2 Example of simulated data, N=40



Inserm
Institut national
de la santé et de la recherche médicale

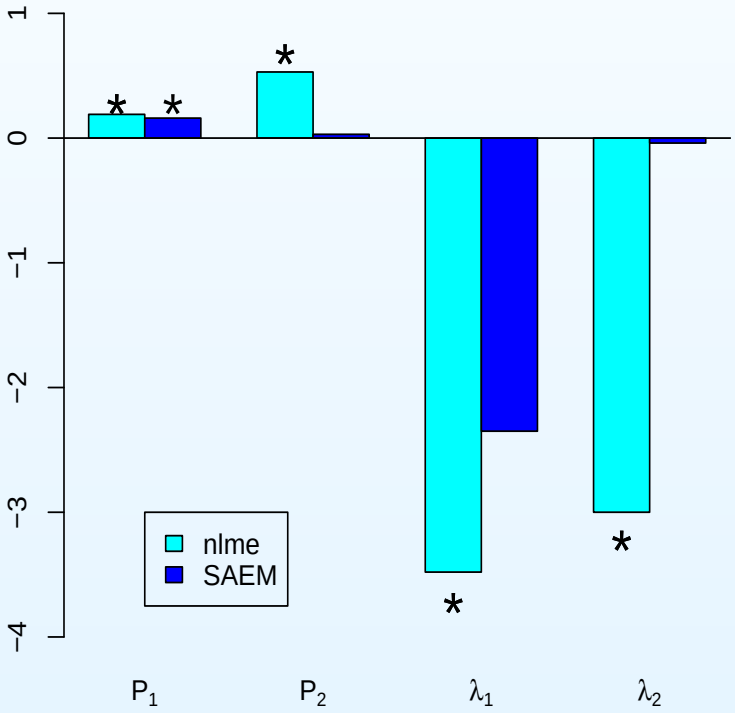
1.3 Evaluation settings

- Simulation of 100 trials with $N = 40$ or $N = 200$ subjects
- **`nlme`** function of R 1.7 software
- SAEM function implemented in R 1.7 software
- Evaluation of the estimation properties for both algorithms
  - Relative Biases
  - Relative RMSE
- Test whether the biases are significantly different from zero by a Student's test on the 100 replications

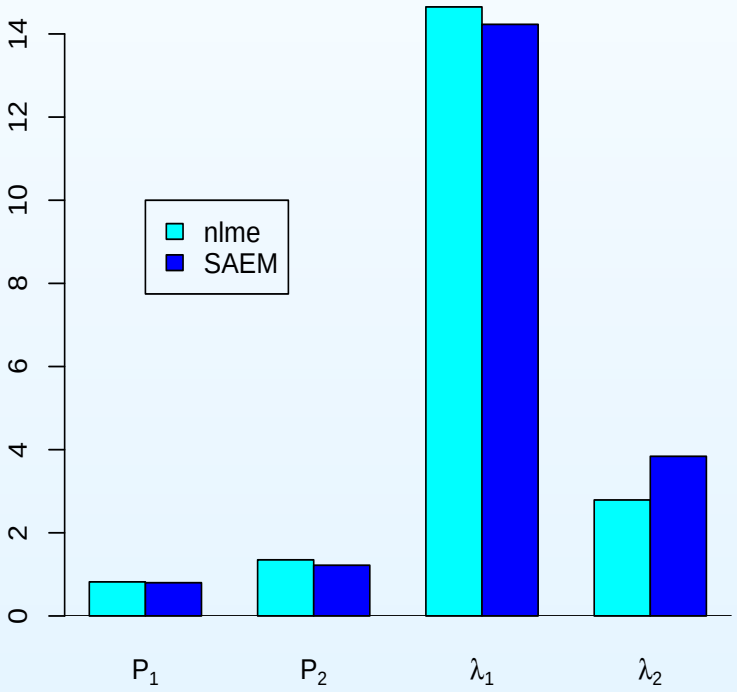
Inserm
Institut national
de la santé et de la recherche médicale

1.4 Results N=40, fixed effects

Relative Biases (%)



Relative RMSE (%)

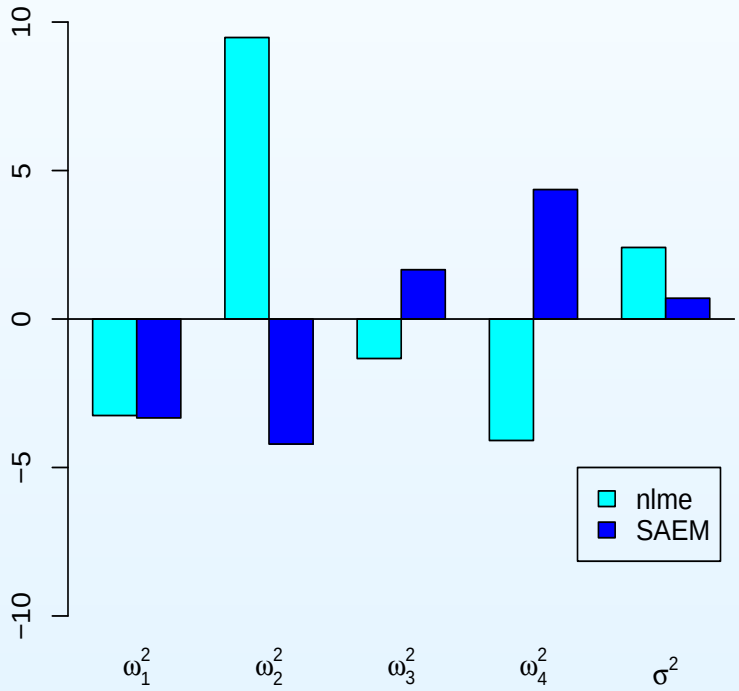


(* p<0.05)

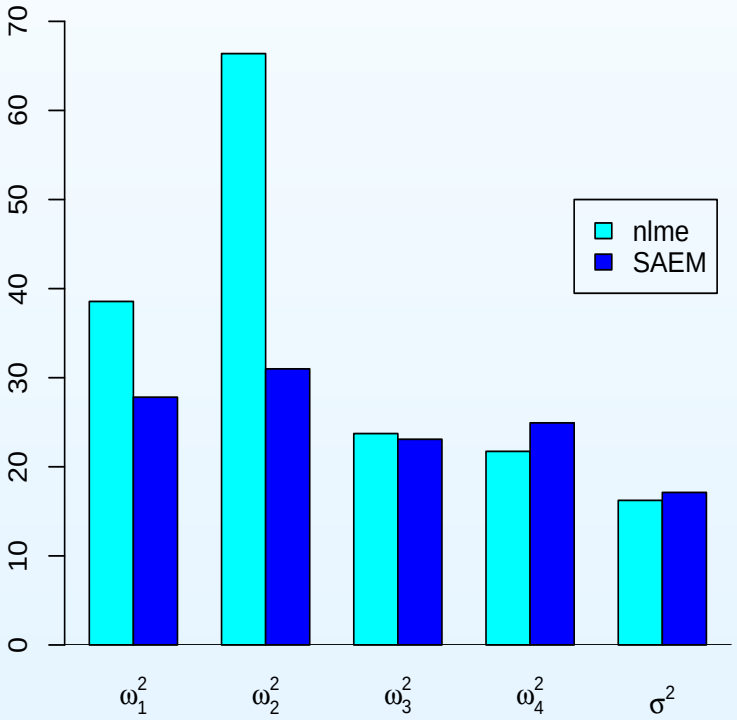
Inserm
Institut national
de la santé et de la recherche médicale

1.5 Results N=40, variance parameters

Relative Biases (%)



Relative RMSE (%)

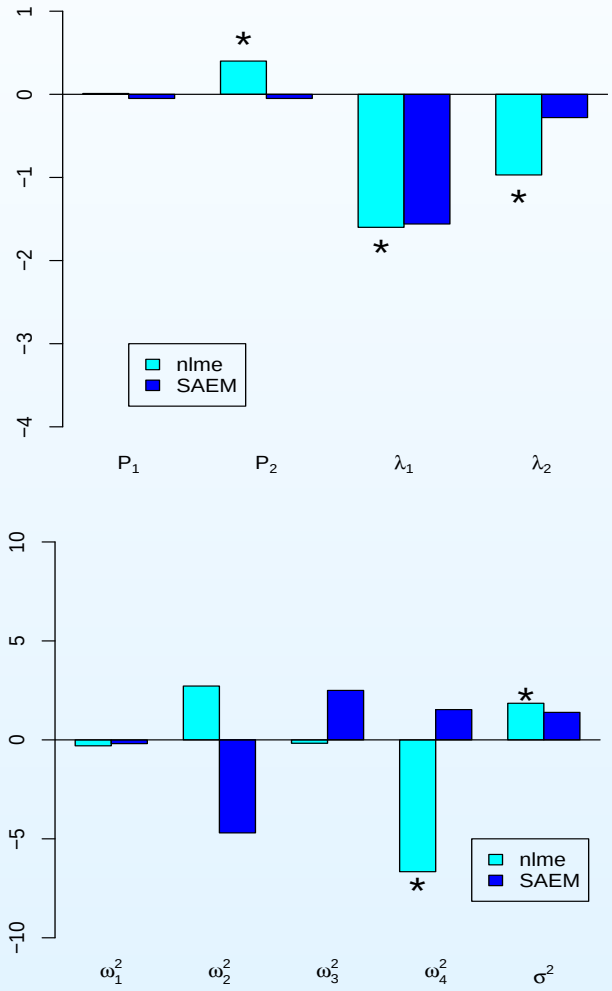


(* p<0.05)

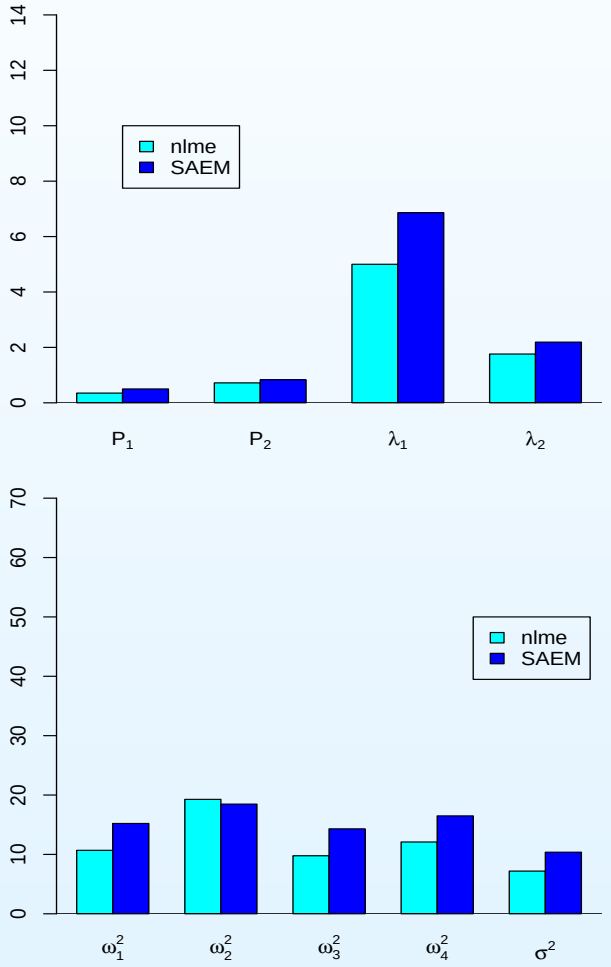
Inserm
Institut national
de la santé et de la recherche médicale

1.6 Results N=200

Relative Biases (%)



Relative RMSE (%)



Inserm
Institut national
de la santé et de la recherche médicale

2.1 Evaluation of the likelihood without linearization

- Approximation of each $p(y_i|\psi\,)$
- For $t = 1, \ldots, T$, simulation of sample $\theta_i^{(t)}$
- Monte Carlo integration

$$p(y_i|\hat{\psi}\,) \approx \frac{1}{T}\sum_{t=1}^{T} p(y_i|\theta_i^{(t)}, \hat{\psi})$$

with $\theta_i^{(t)}$ sampled in the current population distribution $p(., \hat{\psi})$

$\Rightarrow$ No stability of the evaluation even with a very large T

- Importance sampling procedure

$$p(y_i|\hat{\psi}\,) \approx \frac{1}{T}\sum_{t=1}^{T} \frac{p(y_i|\theta_i^{(t)}, \hat{\psi})p(\theta_i^{(t)}, \hat{\psi})}{h_i(\theta_i^{(t)}, \hat{\psi})}$$

with $\theta_i^{(t)}$ sampled in an instrumental distribution $h_i(., \hat{\psi})$

2.2 Importance sampling method

- Choice of the instrumental distribution
  - Gaussian approximation of the individual posterior distribution of θ_i given y_i and $\widehat{\psi}$

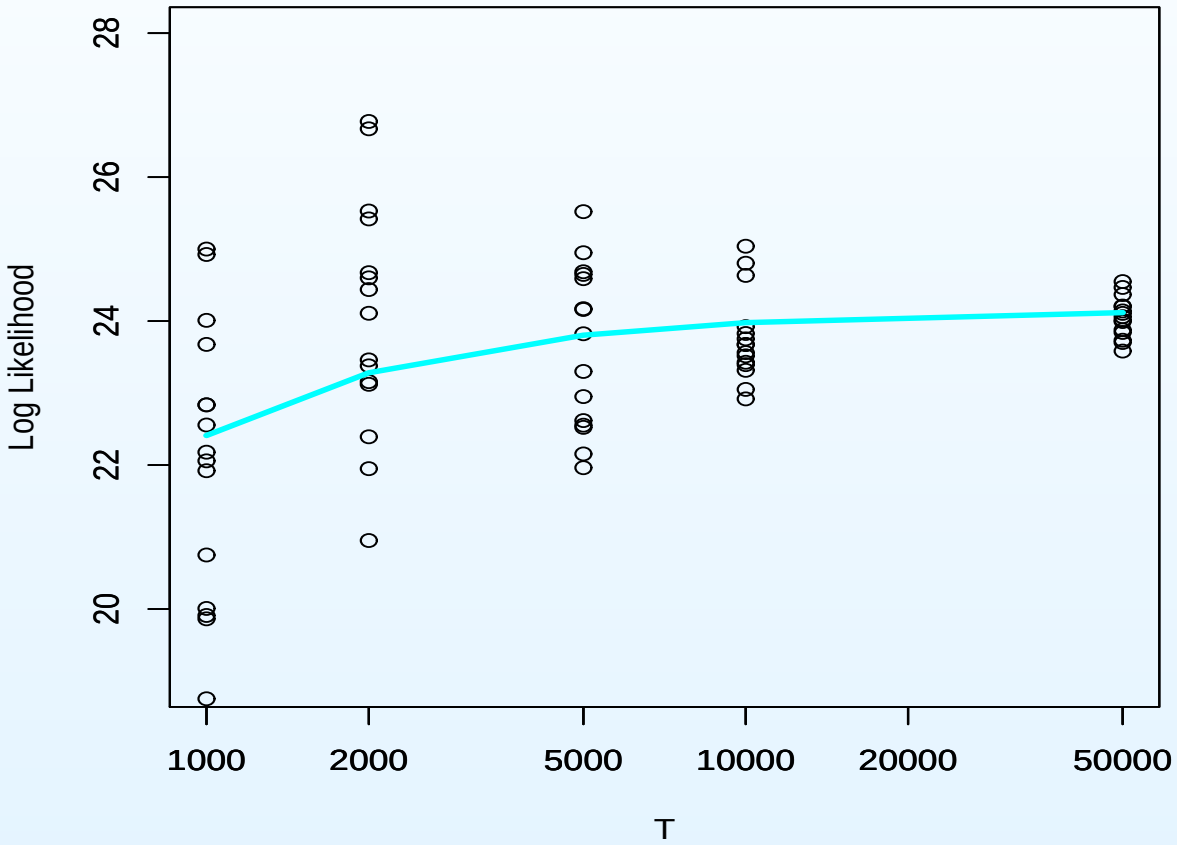
$$h_i(.,\hat{\psi}) = \mathcal{N}(\hat{\mu}_i^{post}, \hat{\Omega}_i^{post})$$

  - $\hat{\mu}_i^{post}$ the posterior individual mean
  - $\hat{\Omega}_i^{post}$ the posterior individual variance
  - Estimation of $\hat{\mu}_i^{post}$ and $\hat{\Omega}_i^{post}$ by the empirical mean and variance of the 250 last θ_i simulated during the MCMC procedure

- Choice of T: the number of simulated samples

2.3 Study of the influence of T

Repeated evaluations of the log likelihood for one trial (N=40) with different seeds and values of T



Using the R software: T=50,000 is a good choice

3.1 LRT

- LRT for a treatment effect on the first viral decay rate $\ln\lambda_1$
- Two groups of treatment of same size
- Evaluation of the likelihood
  - SAEM: likelihood estimated by importance sampling without linearization
  - **`nlme`**: - linearized likelihood obtained by **`nlme`**
    - likelihood estimated by importance sampling without linearization

$\Rightarrow$ Three LRT evaluated

- Evaluation of the type I error with a critical value of 3.84 (level of 5% for a χ^2 with 1 d.f.)

3.2 Evaluation of the type I error of the LRT (p=5 %)

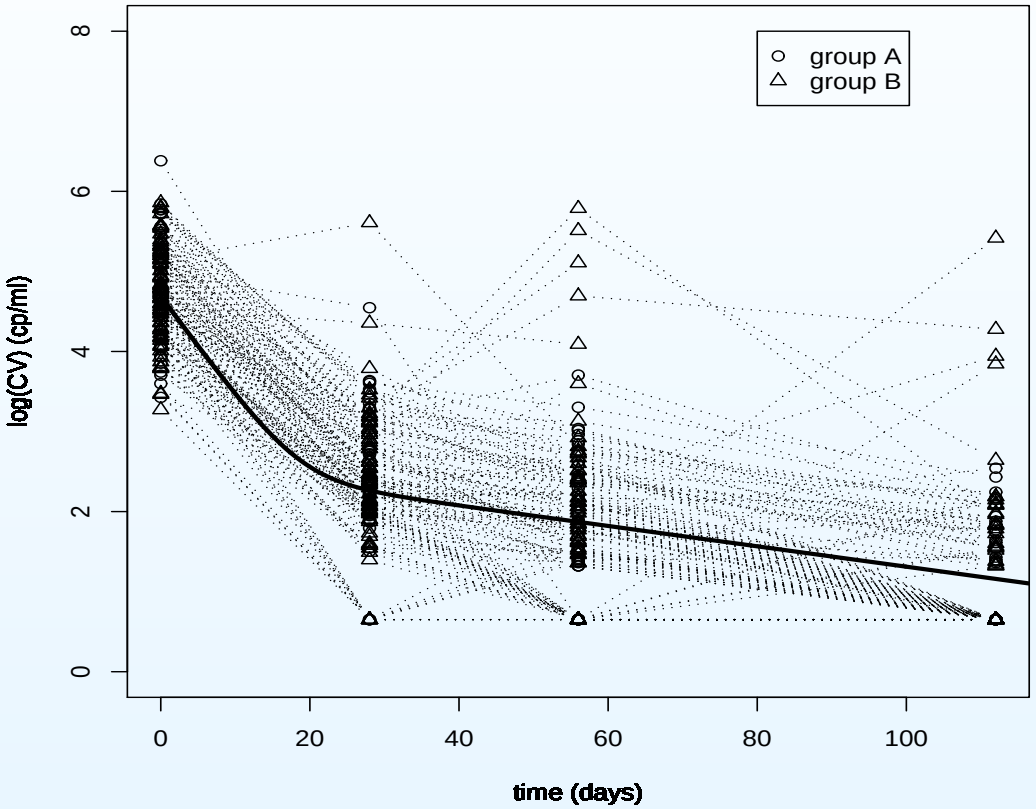
	N=40	N=200
nlme	13 %	18 %
nlme IS	12 %	16 %
SAEM	7 %	5 %

- Inflation of the type I error with **nlme**
- Same inflation with the use of a likelihood estimated without linearization
- Accurate level with SAEM
- Inflation can be explained by the use of algorithms based on the linearization

4.1 Illustration of SAEM on real data (TRIANON)

- AIDS clinical trial supported by the Agence Nationale de Recherche sur le Sida in France (ANRS81)
- 144 HIV-1 infected patients treated during 72 weeks
  - Treatment A : Lamivudine + d4T + Indinavir (71 patients)
  - Treatment B : Nevirapine + d4T + Indinavir (73 patients)
- Analysis of the initial decrease of the viral load (D0, D28, D56, D112)
- Data below the LOQ fixed to LOQ/2

4.2 Results



- No convergence of the **`nlme`** function
- No significant differences with SAEM between the two treatments (LRT on λ_1 and λ_2)

Inserm
Institut national de la santé et de la recherche médicale

4.3 Estimated parameters with SAEM

Parameter	Estimates (SE %)	
$\ln P_1$	10.92	(3.1 %)
$\ln P_2$	6.35	(13 %)
$\ln \lambda_1$	-0.78	(5.3 %)
$\ln \lambda_2$	-3.38	(0.7 %)
ω_1	0.26	(16 %)
ω_2	1.47	(13 %)
ω_3	0.31	(18 %)
ω_4	0.09	(17 %)
σ	0.62	(6.5 %)

Conclusion

- Rapidity and stability of the convergence of the SAEM algorithm
  - Robust to the choice of the initial values
- Less biased than **`nlme`** (especially for small number of subjects)
- Evaluation of the likelihood by importance sampling
  - Stable with a large T
  - Application to the LRT with no inflation of the type I error (in this example)
  - Application to the evaluation of the SE (not shown here)
- SAEM is a good alternative for maximum likelihood estimation in NLMEM
- SAEM now implemented in a MATLAB function developped by Marc Lavielle
- Demonstration during the 'Software demonstration' by Marc Lavielle on several examples (simulated and real data sets)