Department of Mathematical Sciences

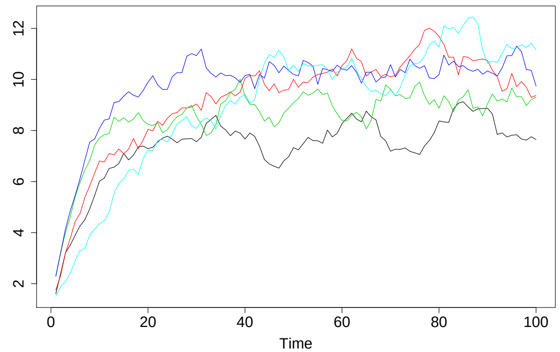
Stochastic differential equations with random effects

Helle Sørensen

Joint with Susanne Ditlevsen

AMS-EMS-SPM, Porto, June 2015

Simulated data



100 discrete-time obs. from five subjects. Same patterns, but

- **The (stationary) levels differ between subjects**
- **Speed of 'mean-reversion' differs between subjects**

Outline

Aim: **Parameter estimation for SDE models with random parameters**

Outline:

- SDE models with linear drift and random parameters
- Estimation approach and simulation results for CIR model
- Data example on growth of pigs

Linear drift SDE's with random parameters

Hierarchical set-up:

- Data: N **equidistant obs. from each of** M **processes:** $Y_i = (Y_{i1}, \ldots, Y_{iN})$ for $i = 1, \ldots, M$ where $Y_{ij} = X_{i,j\Delta}$.

Linear drift SDE's with random parameters

Hierarchical set-up:

- Data: N **equidistant obs. from each of** M **processes:** $Y_i = (Y_{i1}, \dots, Y_{iN})$ for $i = 1, \dots, M$ where $Y_{ij} = X_{i,j\Delta}$.
- For process i: $dX_{i,t} = -b_i(X_{i,t} - a_i)\,dt + \sigma(X_{i,t})\,dW_{i,t}$.

  a_i is the stationary mean, b_i is a mean-reversion parameter.

Linear drift SDE's with random parameters

Hierarchical set-up:

- Data: N **equidistant obs. from each of** M **processes:** $Y_i = (Y_{i1}, \dots, Y_{iN})$ for $i = 1, \dots, M$ where $Y_{ij} = X_{i,j\Delta}$.
- For process i: $dX_{i,t} = -b_i(X_{i,t} - a_i)\,dt + \sigma(X_{i,t})\,dW_{i,t}$.

  a_i is the stationary mean, b_i is a mean-reversion parameter.
- Drift parameters vary between processes: $a_i \sim N(\alpha, \tau_a^2)$, $b_i \sim N(\beta, \tau_b^2)$, a_i and b_i independent (for simplicity).

  $\alpha, \beta, \tau_a, \tau_b$ are **population parameters**

Linear drift SDE's with random parameters

Hierarchical set-up:

- Data: N **equidistant obs. from each of** M **processes:** $Y_i = (Y_{i1}, \dots, Y_{iN})$ for $i = 1, \dots, M$ where $Y_{ij} = X_{i,j\Delta}$.
- For process i: $dX_{i,t} = -b_i(X_{i,t} - a_i)\,dt + \sigma(X_{i,t})\,dW_{i,t}$.

  a_i is the stationary mean, b_i is a mean-reversion parameter.
- Drift parameters vary between processes: $a_i \sim N(\alpha, \tau_a^2)$, $b_i \sim N(\beta, \tau_b^2)$, a_i and b_i independent (for simplicity).

  $\alpha, \beta, \tau_a, \tau_b$ are **population parameters**
- Today (most of the time): σ is unspecified

Linear drift SDE's with random parameters

Hierarchical set-up:

- Data: N **equidistant obs. from each of** M **processes:** $Y_i = (Y_{i1}, \dots, Y_{iN})$ for $i = 1, \dots, M$ where $Y_{ij} = X_{i,j\Delta}$.
- For process i: $dX_{i,t} = -b_i(X_{i,t} - a_i)\,dt + \sigma(X_{i,t})\,dW_{i,t}$.

  a_i is the stationary mean, b_i is a mean-reversion parameter.
- Drift parameters vary between processes: $a_i \sim N(\alpha, \tau_a^2)$, $b_i \sim N(\beta, \tau_b^2)$, a_i and b_i independent (for simplicity).

  $\alpha, \beta, \tau_a, \tau_b$ are **population parameters**
- Today (most of the time): σ is unspecified

Parameter of interest: $\theta = (\alpha, \beta, \tau_a, \tau_b)$

Aim: **Estimation of θ!**

Approximation to likelihood

The contribution from process i to the **correct likelihood** is

$$L_i = \int \prod_{j=1}^{N} p(Y_{ij} | Y_{i,j-1}, a_i, b_i, \sigma) p(a_i, b_i | \alpha, \beta, \tau_a, \tau_b)\, d(a_i, b_i)$$

Our approach: Approximate $p(Y_{ij} | Y_{i,j-1}, a_i, b_i, \sigma)$ by a Gaussian density, but leave the rest unchanged.

Approximation to likelihood

The contribution from process i to the **correct likelihood** is

$$L_i = \int \prod_{j=1}^{N} p(Y_{ij} | Y_{i,j-1}, a_i, b_i, \sigma) p(a_i, b_i | \alpha, \beta, \tau_a, \tau_b)\, d(a_i, b_i)$$

Our approach: Approximate $p(Y_{ij} | Y_{i,j-1}, a_i, b_i, \sigma)$ by a Gaussian density, but leave the rest unchanged.

- Use the **true conditional expectation** as mean, i.e.

$$\mathrm{E}(Y_{i,j} | Y_{i,j-1} = y, a_i = a, b_i = b)$$

Approximation to likelihood

The contribution from process i to the **correct likelihood** is

$$L_i = \int \prod_{j=1}^{N} p(Y_{ij} | Y_{i,j-1}, a_i, b_i, \sigma) p(a_i, b_i | \alpha, \beta, \tau_a, \tau_b)\, d(a_i, b_i)$$

Our approach: Approximate $p(Y_{ij} | Y_{i,j-1}, a_i, b_i, \sigma)$ by a Gaussian density, but leave the rest unchanged.

- Use the **true conditional expectation** as mean, i.e.

$$\mathrm{E}(Y_{i,j} | Y_{i,j-1} = y, a_i = a, b_i = b) = a + e^{-b\Delta}(y - a)$$

Approximation to likelihood

The contribution from process i to the **correct likelihood** is

$$L_i = \int \prod_{j=1}^{N} p(Y_{ij} | Y_{i,j-1}, a_i, b_i, \sigma) p(a_i, b_i | \alpha, \beta, \tau_a, \tau_b)\, d(a_i, b_i)$$

Our approach: Approximate $p(Y_{ij} | Y_{i,j-1}, a_i, b_i, \sigma)$ by a Gaussian density, but leave the rest unchanged.

- Use the **true conditional expectation** as mean, i.e.

$$\mathrm{E}(Y_{i,j} | Y_{i,j-1} = y, a_i = a, b_i = b) = a + e^{-b\Delta}(y - a)$$

- For convenience, use an approx. to the conditional variance:

$$\mathrm{Var}(Y_{i,j} | Y_{i,j-1} = y, a_i = a, b_i = b) \approx c_1 (c_2 + y^{c_3})^2$$

Why?

- Use the correct conditional mean in the Gaussian approx. **Same interpretation of parameters of interest**.
- Analogy to **martingale estimating functions** for a single SDE, where Gaussian approx. leads to consistent estimators.
- Can think of the approximating model as a nonlinear mixed effects (NLME) model, and **rely on well-tested software!**

R: Data and `nlme` command

```
> head(myData)
  myTime         y      yLag subject Delta
1      1 0.6893262 0.0000000       1     1
2      2 1.2101795 0.6893262       1     1
3      3 1.3890037 1.2101795       1     1
4      4 1.8931005 1.3890037       1     1

m1 <- function(x,t,alpha,beta) exp(-beta*t)*(x-alpha)+alpha

fit <- nlme(y ~ m1(yLag, Delta, alpha, beta),
            fixed = list(alpha+beta~1),
            random = list(subject=alpha+beta~1),
            weights=varConstPower(form=~yLag))
            data=myData, start=myStart)
```


Simulation study for the square-root process

Square-root (Cox-Ingersoll-Ross) process with random effects:

- SDE for subject i: $dX_{i,t} = -b_i(X_{i,t} - a_i)\,dt + \sigma\sqrt{X_{i,t}}\,dW_{i,t}$
- $a_i \sim N(\alpha, \tau_a^2)$, $b_i \sim N(\beta, \tau_b^2)$, a_i and b_i independent.

Simulation study for the square-root process

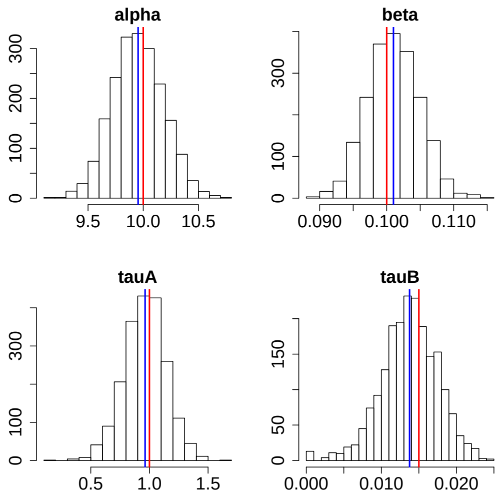
Square-root (Cox-Ingersoll-Ross) process with random effects:

- SDE for subject i: $dX_{i,t} = -b_i(X_{i,t} - a_i)\,dt + \sigma\sqrt{X_{i,t}}\,dW_{i,t}$
- $a_i \sim N(\alpha, \tau_a^2)$, $b_i \sim N(\beta, \tau_b^2)$, a_i and b_i independent.

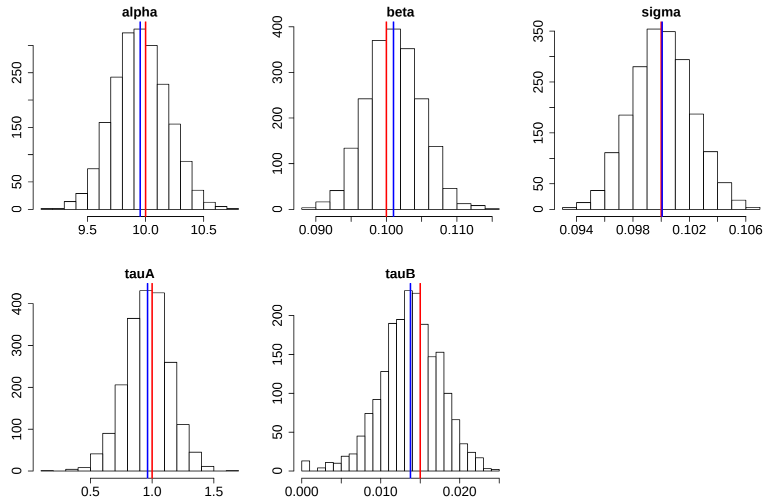
Set-up:

- $M = 25$ (no. of subjects)
- $N = 50$ (no. of observations per subject)
- $\Delta = 1$ (distance between observations)
- $a_i \sim N(10, 1)$, $b_i \sim N(0.1, 0.015^2)$, $\sigma = 0.1$
- 2000 simulated data sets

2000 sim., $M = 25$, $N = 50$, $\Delta = 1$



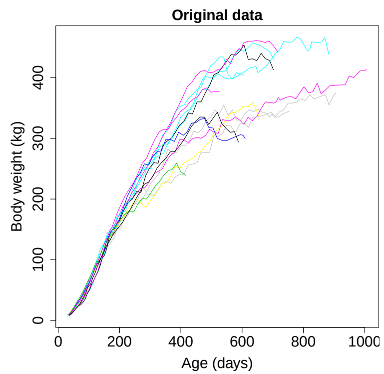
Also possible to compute an estimate of σ



Simulation results

- Only small bias
- Approximately Gaussian distributions
- Coverage of 95% confidence intervals for α and β from NLME model is about 93% ($N = 50$, $\Delta = 1$, $M = 25, 50$)
- We are also able to estimate σ in the CIR model with very little bias
- Similar results for the Ornstein-Uhlenbech process and the Jacobi processes

Growth data from pigs



- Body weight from weaning to (near) maturity from 13 pigs
- Weekly obs. until 200 days of age, then every second week
- Interested in **population average** of mature body weight

Bertalanffy-Richards growth model

Bertalanffy-Richards ODE for growth — for some power γ:

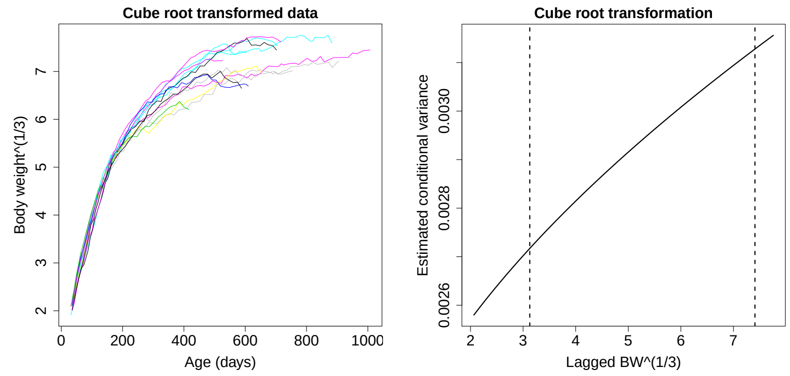
$$\frac{dx^\gamma}{dt} = -\beta(x^\gamma - \alpha)$$

SDE version:

$$dX_t^\gamma = -\beta(X_t^\gamma - \alpha)\,dt + \sigma(X_t^\gamma)\,dW_t$$

Let us try $\gamma = 1/3$, ie. **model $BW^{1/3}$ as linear drift diffusions.**

Transformed data and estimated cond'l variance



- Estimated conditional variance increases only about 15% over the 95% central part of the data. Not far from constant.
- This suggests an **Ornstein-Uhlenbeck process**

Model and estimates

AIC values suggest a fixed (non-random) β, leading to:

- $dX_{i,t}^{1/3} = -\beta(X_{i,t}^{1/3} - a_i)\,dt + \sigma\,dW_{i,t}$
- $a_i \sim N(0, \tau_a^2)$.

Gaussian transitions and constant conditional variance. Then NLME with constant variance ($v \equiv c$) gives the **true likelihood.**

Model and estimates

AIC values suggest a fixed (non-random) β, leading to:

- $dX_{i,t}^{1/3} = -\beta(X_{i,t}^{1/3} - a_i)\,dt + \sigma\,dW_{i,t}$
- $a_i \sim N(0, \tau_a^2)$.

Gaussian transitions and constant conditional variance. Then NLME with constant variance ($v \equiv c$) gives the **true likelihood.**

Results:

- ML estimates: $\hat{\alpha} = 7.30$, $\hat{\tau}_a = 0.28$, $\hat{\beta} = 0.0061$.
- Estimate and 95% confidence interval for median mature body weight, i.e. for α^3: 362 kg (388–417)

Concluding remarks

Conclusions so far:

- Simple strategy, relies on well-tested software
- Nice simulation results — also for other linear drift processes
- Preliminary investigations indicate nice asymptotic behaviour as $M, N \to \infty$ (fixed Δ)
- Procedure may help to identify appropriate model

Concluding remarks

Conclusions so far:

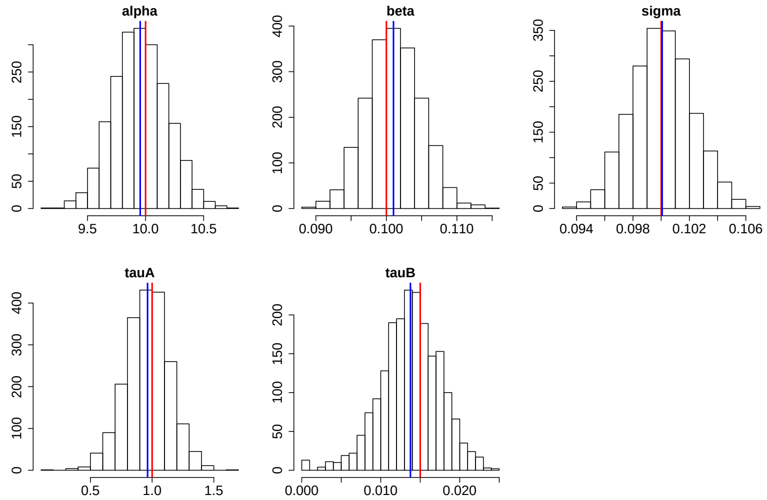
- Simple strategy, relies on well-tested software
- Nice simulation results — also for other linear drift processes
- Preliminary investigations indicate nice asymptotic behaviour as $M, N \to \infty$ (fixed Δ)
- Procedure may help to identify appropriate model

What is next?

- Use correct conditional variance whenever known.
- Crucial that the correct conditional mean is used in the NLME model. What about models with non-linear drift?
- What if the processses are measured with noise?

Thank you for your attention

Also possible to compute an estimate of σ



Asymptotic considerations

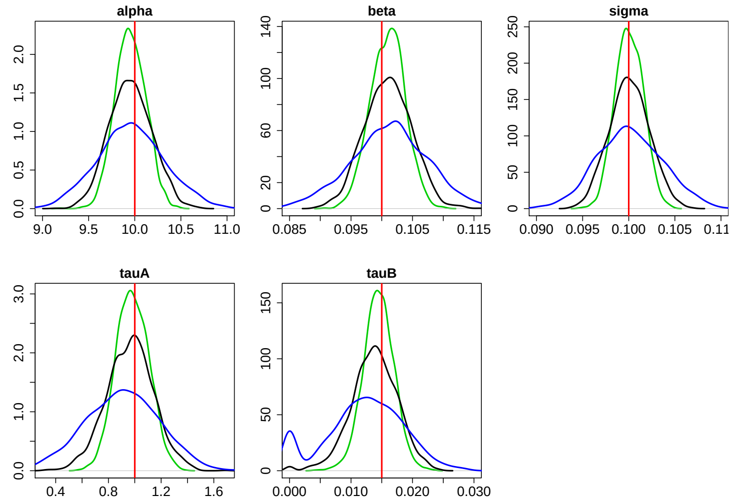
Can we say anything about **asymptotic properties** of $\hat{\theta}$?

Asymptotic set-up:

- Obviously we need $M \to \infty$ (no. of subjects)
- But how about Δ and N?
- In particular: What happens for $M \to \infty$ but fixed Δ, fixed N?

Let us look at some numerical experiments.

$M = 10, 25, 50,\ N = 50,\ \Delta = 1$



What happens for $M \to \infty$ but fixed Δ, fixed N?

$$
\begin{aligned}
\frac{1}{M}\log \tilde{L}_M &= \frac{1}{M}\sum_{i=1}^{M} \log \int \prod_{j=1}^{N} \tilde{p}(Y_{ij}|Y_{i,j-1}, a_i, b_i, c)p(a_i, b_i|\alpha, \beta, \tau)\, d(a_i, b_i) \\
&\to \mathrm{E}\left[\log \int \prod_{j=1}^{N} \tilde{p}(Y_{ij}|Y_{i,j-1}, a_i, b_i, c)p(a_i, b_i|\alpha, \beta, \tau)\, d(a_i, b_i)\right] \\
&= \tilde{L}_\infty(\theta)
\end{aligned}
$$

Assume that $\tilde{L}_\infty$ has well-seperated maximum at $\tilde{\theta}$. Under suitable tightness conditions: $\hat{\theta}_M \xrightarrow{P} \tilde{\theta}$.

What happens for $M \to \infty$ but fixed Δ, fixed N?

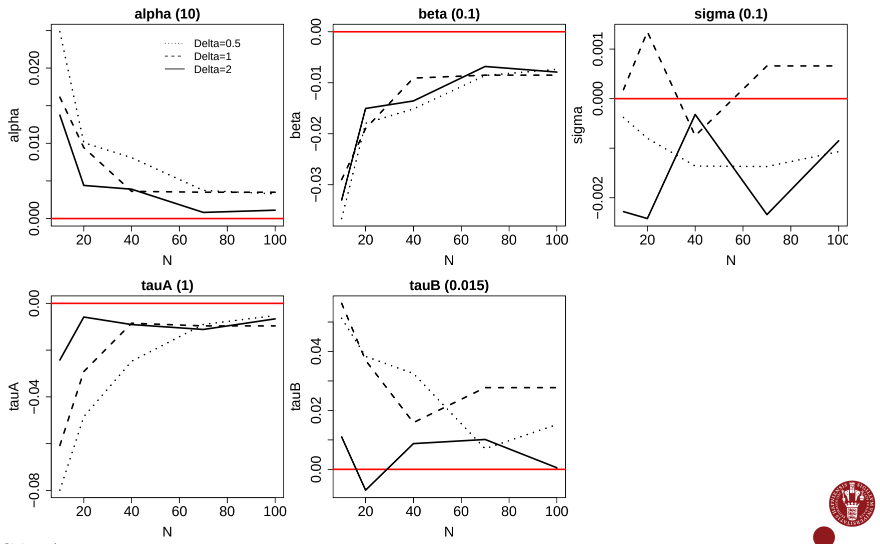
$$
\begin{aligned}
\frac{1}{M}\log \tilde{L}_M &= \frac{1}{M}\sum_{i=1}^{M} \log \int \prod_{j=1}^{N} \tilde{p}(Y_{ij}|Y_{i,j-1}, a_i, b_i, c)p(a_i, b_i|\alpha, \beta, \tau)\, d(a_i, b_i) \\
&\to \mathrm{E}\left[\log \int \prod_{j=1}^{N} \tilde{p}(Y_{ij}|Y_{i,j-1}, a_i, b_i, c)p(a_i, b_i|\alpha, \beta, \tau)\, d(a_i, b_i)\right] \\
&= \tilde{L}_\infty(\theta)
\end{aligned}
$$

Assume that $\tilde{L}_\infty$ has well-seperated maximum at $\tilde{\theta}$. Under suitable tightness conditions: $\hat{\theta}_M \xrightarrow{P} \tilde{\theta}$.

Numerial experiment:

- Square-root process with correct conditional variance
- Compute maximum $\tilde{\theta}$ for various values of Δ and N
- Cumbersome: Integration, expectation, 5D-optimization

Relative error of $\tilde{\theta}$ compared to θ_0. "$M = +\infty$"



Conclusions on asymptotic properties

- Small asymptotic bias, even for large N
- Good results for Δ 'large'. This is because we use the true conditional moments rather than small-δ approximations.
- Concistency seems to require $M \to \infty$ as well as $N \to \infty$, but presumably not $\Delta \to 0$