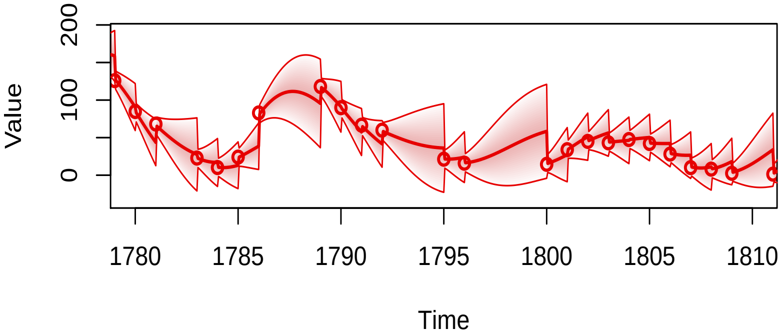
Hierarchical Continuous Time Dynamic Models with ctsem

July, 2018

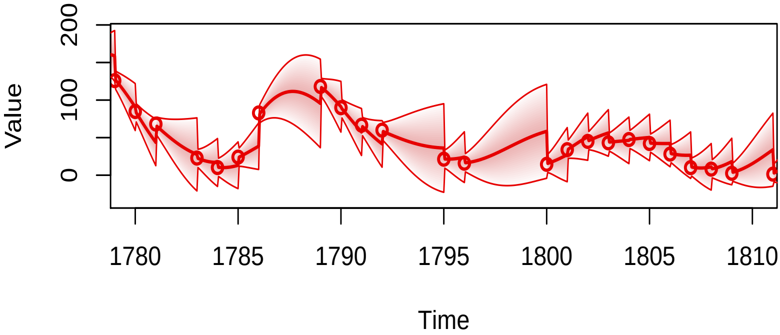
Charles Driver

Max Planck Institute for Human Development

- Common approaches to dynamic modelling (latent change, autoregressive) assume that the relation between observations at different occasions are equal.

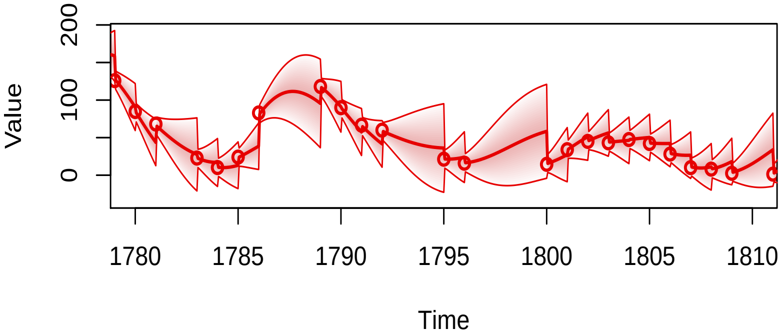


- Common approaches to dynamic modelling (latent change, autoregressive) assume that the relation between observations at different occasions are equal.
- This is fine in some cases, however problems can arise when time intervals between measurements are not equal.



- Common approaches to dynamic modelling (latent change, autoregressive) assume that the relation between observations at different occasions are equal.
- This is fine in some cases, however problems can arise when time intervals between measurements are not equal.
- Continuous time models determine the relations between each measurement occasion via a deterministic function of the continuous time parameters and the time interval between measurements.

$\eta_t = A^* \eta_{t-1} + b^* + \zeta^*(t)$



- More accuracy by incorporating time explicitly.

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- More accurate than embedding approaches to differential equations, and no need to set embedding dimension.

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- More accurate than embedding approaches to differential equations, and no need to set embedding dimension.
- Tangible benefits:

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- More accurate than embedding approaches to differential equations, and no need to set embedding dimension.
- Tangible benefits:
  - More flexible data collection - no longer a need for consistent measurement intervals within or between individuals.

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- More accurate than embedding approaches to differential equations, and no need to set embedding dimension.
- Tangible benefits:
  - More flexible data collection - no longer a need for consistent measurement intervals within or between individuals.
  - Consistent parameter comparisons between different data collection schedules - between waves, studies, or both.

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- More accurate than embedding approaches to differential equations, and no need to set embedding dimension.
- Tangible benefits:
    - More flexible data collection - no longer a need for consistent measurement intervals within or between individuals.
    - Consistent parameter comparisons between different data collection schedules - between waves, studies, or both.
    - Differential equations:

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- More accurate than embedding approaches to differential equations, and no need to set embedding dimension.
- Tangible benefits:
  - More flexible data collection - no longer a need for consistent measurement intervals within or between individuals.
  - Consistent parameter comparisons between different data collection schedules - between waves, studies, or both.
  - Differential equations:
    - Parsimonious specification of higher order dynamics such as oscillations and slow changing trends.

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- More accurate than embedding approaches to differential equations, and no need to set embedding dimension.
- Tangible benefits:
  - More flexible data collection - no longer a need for consistent measurement intervals within or between individuals.
  - Consistent parameter comparisons between different data collection schedules - between waves, studies, or both.
  - Differential equations:
    - Parsimonious specification of higher order dynamics such as oscillations and slow changing trends.
    - Direct interpretation re causality – given true model!

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- More accurate than embedding approaches to differential equations, and no need to set embedding dimension.
- Tangible benefits:
  - More flexible data collection - no longer a need for consistent measurement intervals within or between individuals.
  - Consistent parameter comparisons between different data collection schedules - between waves, studies, or both.
  - Differential equations:
    - Parsimonious specification of higher order dynamics such as oscillations and slow changing trends.
    - Direct interpretation re causality – given true model!
  - Helps (forces?) us to think more coherently in terms of time.

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- More accurate than embedding approaches to differential equations, and no need to set embedding dimension.
- Tangible benefits:
  - More flexible data collection - no longer a need for consistent measurement intervals within or between individuals.
  - Consistent parameter comparisons between different data collection schedules - between waves, studies, or both.
  - Differential equations:
    - Parsimonious specification of higher order dynamics such as oscillations and slow changing trends.
    - Direct interpretation re causality – given true model!
  - Helps (forces?) us to think more coherently in terms of time.
- Downsides:

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- More accurate than embedding approaches to differential equations, and no need to set embedding dimension.
- Tangible benefits:
  - More flexible data collection - no longer a need for consistent measurement intervals within or between individuals.
  - Consistent parameter comparisons between different data collection schedules - between waves, studies, or both.
  - Differential equations:
    - Parsimonious specification of higher order dynamics such as oscillations and slow changing trends.
    - Direct interpretation re causality – given true model!
  - Helps (forces?) us to think more coherently in terms of time.
- Downsides:
  - More mathematically and computationally demanding than autoregressive / latent change approaches.

- Latent dynamic process with measurement model, 1+ subjects, each with 1+ obs, varying time intervals. Dynamics are a linear stochastic differential equation:

$$\mathrm{d}\boldsymbol{\eta}(t) = \Big(\mathbf{A}\boldsymbol{\eta}(t) + \mathbf{b} + \mathbf{M}\chi(t)\Big)\mathrm{d}t + \mathbf{G}\mathrm{d}\mathbf{W}(t) \tag{1}$$

Observations for each subject are described by:

$$\mathbf{y}(t) = \boldsymbol{\Lambda}\boldsymbol{\eta}(t) + \boldsymbol{\tau} + \boldsymbol{\epsilon}(t) \quad \text{where } \boldsymbol{\epsilon}(t) \sim \mathrm{N}(\mathbf{0}_c, \boldsymbol{\Theta}) \tag{2}$$

- Latent dynamic process with measurement model, 1+ subjects, each with 1+ obs, varying time intervals. Dynamics are a linear stochastic differential equation:

$$\mathrm{d}\boldsymbol{\eta}(t) = \Big(\mathbf{A}\boldsymbol{\eta}(t) + \mathbf{b} + \mathbf{M}\boldsymbol{\chi}(t)\Big)\mathrm{d}t + \mathbf{G}\mathrm{d}\mathbf{W}(t) \tag{1}$$

  Observations for each subject are described by:

$$\mathbf{y}(t) = \boldsymbol{\Lambda}\boldsymbol{\eta}(t) + \boldsymbol{\tau} + \boldsymbol{\epsilon}(t) \quad \text{where } \boldsymbol{\epsilon}(t) \sim \mathrm{N}(\mathbf{0}_c, \boldsymbol{\Theta}) \tag{2}$$

- The SDE may be solved, for any observation $u \in \mathbf{U}$:

$$\boldsymbol{\eta}_u = \mathbf{A}^*_u \boldsymbol{\eta}_{u-1} + \mathbf{b}^*_u + \mathbf{M}\mathbf{x}_u + \boldsymbol{\zeta}^*_u \qquad \boldsymbol{\zeta}^*_u \sim \mathrm{N}(\mathbf{0}_v, \mathbf{Q}^*_u) \tag{3}$$

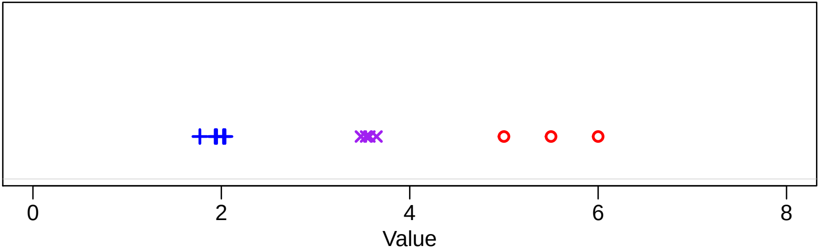
$$\mathbf{A}^*_u = e^{\mathbf{A}(t_u - t_{u-1})} \tag{4}$$

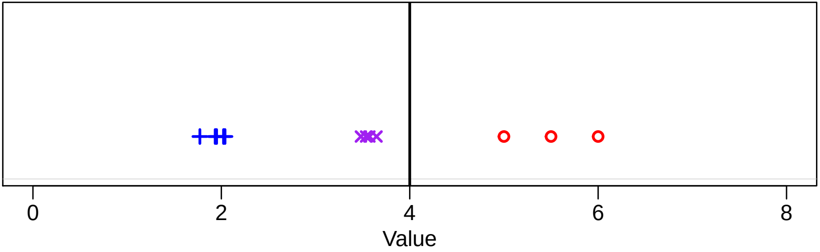
$$\mathbf{b}^*_u = \mathbf{A}^{-1}(\mathbf{A}^*_u - \mathbf{I})\mathbf{b} \tag{5}$$

$$\mathbf{Q}^*_u = \mathbf{Q}_\infty - \mathbf{A}^*_u \mathbf{Q}_\infty \mathbf{A}^{*\top}_u \tag{6}$$

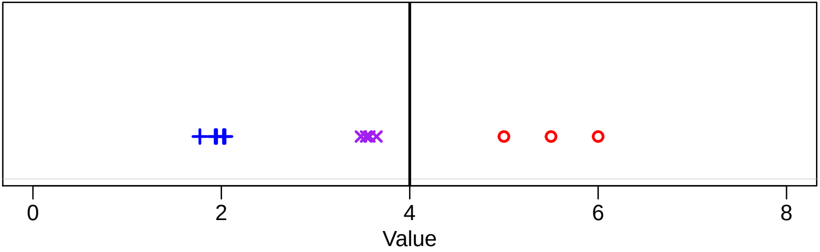
$$\mathbf{Q}_\infty = \mathrm{irow}\big(-\mathbf{A}_\#^{-1}\,\mathrm{row}(\mathbf{Q})\big) \tag{7}$$

Individual differences in dynamics?

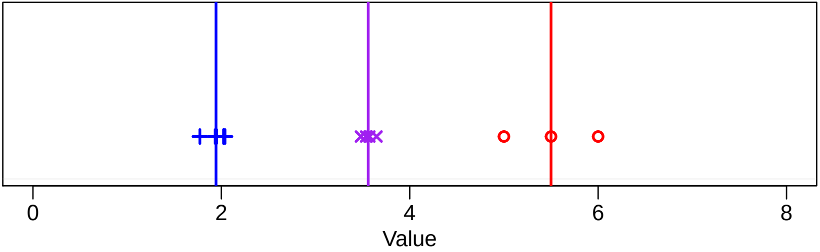

0
2
4
6
8
Value



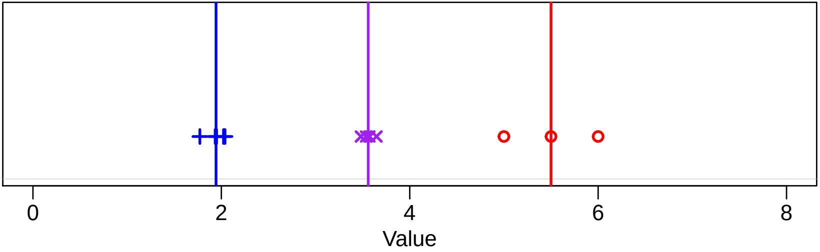
- Complete pooling - estimate single fixed effect parameter for entire sample.



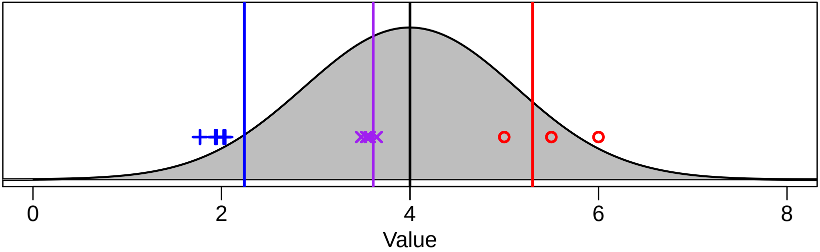
- Complete pooling - estimate single fixed effect parameter for entire sample.
    - Contaminates within subject model with between subjects differences, and not a great way to understand individual differences...



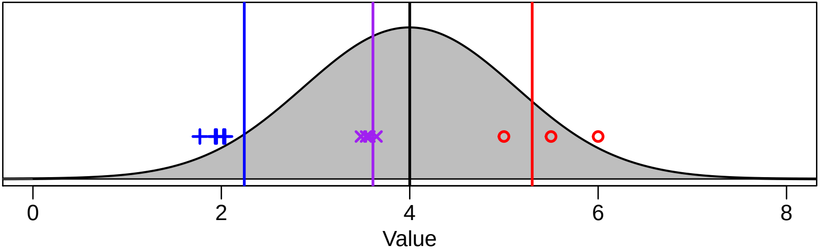
- Complete pooling - estimate single fixed effect parameter for entire sample.
    - Contaminates within subject model with between subjects differences, and not a great way to understand individual differences...
- No pooling - assume individuals have no similarities, estimate seperate model for each.



- Complete pooling - estimate single fixed effect parameter for entire sample.
  - Contaminates within subject model with between subjects differences, and not a great way to understand individual differences...
- No pooling - assume individuals have no similarities, estimate seperate model for each.
  - Simple and perfect if sufficient data exists – 'sufficient' may be extremely large – otherwise prone to finite sample biases and high variance.



- Complete pooling - estimate single fixed effect parameter for entire sample.
    - Contaminates within subject model with between subjects differences, and not a great way to understand individual differences...
- No pooling - assume individuals have no similarities, estimate seperate model for each.
    - Simple and perfect if sufficient data exists – ‘sufficient’ may be extremely large – otherwise prone to finite sample biases and high variance.
- Partial pooling - estimate population distribution for individual models.



- Complete pooling - estimate single fixed effect parameter for entire sample.
    - Contaminates within subject model with between subjects differences, and not a great way to understand individual differences...
- No pooling - assume individuals have no similarities, estimate seperate model for each.
    - Simple and perfect if sufficient data exists – ‘sufficient’ may be extremely large – otherwise prone to finite sample biases and high variance.
- Partial pooling - estimate population distribution for individual models.
    - More complex models but most flexible - parameters are not either ‘freely varying’ or ‘not varying at all’ but the extent of allowed variation is estimated.

$$p(\mathbf{\Phi}, \boldsymbol{\mu}, \mathbf{R}, \beta | \mathbf{Y}, \mathbf{Z}) = \frac{p(\mathbf{Y}|\mathbf{\Phi})p(\mathbf{\Phi}|\boldsymbol{\mu}, \mathbf{R}, \boldsymbol{\beta}, \mathbf{Z})p(\boldsymbol{\mu}, \mathbf{R}, \beta)}{p(\mathbf{Y})} \tag{9}$$

Where subject specific parameters $\mathbf{\Phi}_i$ are determined in the following manner:

$$\mathbf{\Phi}_i = \mathrm{tform}\left(\boldsymbol{\mu} + \mathbf{R}\mathbf{h}_i + \boldsymbol{\beta}\mathbf{z}_i\right) \tag{10}$$

$$\mathbf{h}_i \sim \mathrm{N}(\mathbf{0}, \mathbf{1}) \tag{11}$$

$$\boldsymbol{\mu} \sim \mathrm{N}(\mathbf{0}, \mathbf{1}) \tag{12}$$

$$\boldsymbol{\beta} \sim \mathrm{N}(\mathbf{0}, \mathbf{1}) \tag{13}$$

- ctsem - open source R software  

- ctsem - open source R software
  - Originally only mixed effects models, using OpenMx as backend.


- ctsem - open source R software
  - Originally only mixed effects models, using OpenMx as backend.
  - Using Stan as backend, added support for fully random effects with Bayesian estimation.

- ctsem - open source R software  
  - Originally only mixed effects models, using OpenMx as backend.
  - Using Stan as backend, added support for fully random effects with Bayesian estimation.
- ctModel function constructs a ctsem model based on specified matrices of free and fixed parameters.

- ctsem - open source R software
  - Originally only mixed effects models, using OpenMx as backend.
  - Using Stan as backend, added support for fully random effects with Bayesian estimation.
- ctModel function constructs a ctsem model based on specified matrices of free and fixed parameters.
- Modify between subject effects, covariates, priors etc as needed.

- ctsem - open source R software  
  - Originally only mixed effects models, using OpenMx as backend.
  - Using Stan as backend, added support for fully random effects with Bayesian estimation.
- ctModel function constructs a ctsem model based on specified matrices of free and fixed parameters.
- Modify between subject effects, covariates, priors etc as needed.
- ctStanFit function constructs a Stan model and calls rstan for estimation, using either Kalman filter for continuous variables or direct sampling of states for other measurement models.

- Non-linear dynamics and measurement via unscented Kalman filter.

- Non-linear dynamics and measurement via unscented Kalman filter.
- Binary and ordinal measurement models.

- Non-linear dynamics and measurement via unscented Kalman filter.
- Binary and ordinal measurement models.
- Optimization followed by importance sampling for faster results.

- ctsem and vignettes
  `https://cran.r-project.org/web/packages/ctsem/index.html`
- Other articles:
  `https://www.researchgate.net/profile/Charles_Driver`
- Overview of provided R script:
  - Generate some data.
  - Fit a univariate linear growth curve with random effects and a covariate.
  - Add in dynamics.
  - Add in an intervention.
  - And a second latent process.
  - Drop the second latent process and try a state-dependent (non-linear) intervention.