State Space Models for Longitudinal Neuropsychological Outcomes

Zach Baucom

Introduction

- Zach Baucom, 5th year Biostatistics PhD student at Boston University.
- Initially selected to be funded through an NIH training grant.
- Have been given the opportunity to work in many different areas:
  - Genomics
  - Genetics
  - Observational studies
  - Clinical trials
- Course work focused on computational statistics and machine learning

Introduction

- Currently working with Yorghos Tripodis.
  - Data and Biostatistics Director of the Boston University Alzheimer's Disease Center
- Interested in modeling subject level cognitive decline over time/cognitive trajectories.

Dementia is a Problem

- According to the World Health Organization dementia effects around 50 million people in the world today
  - 60-70% of those due to Alzheimer's disease (AD)
- We want to create a model that can be used to,
  - Illuminate how and why dementia progresses.
  - Assist in early disease diagnosis.
  - Determine intervention effectiveness.

Motivating Data

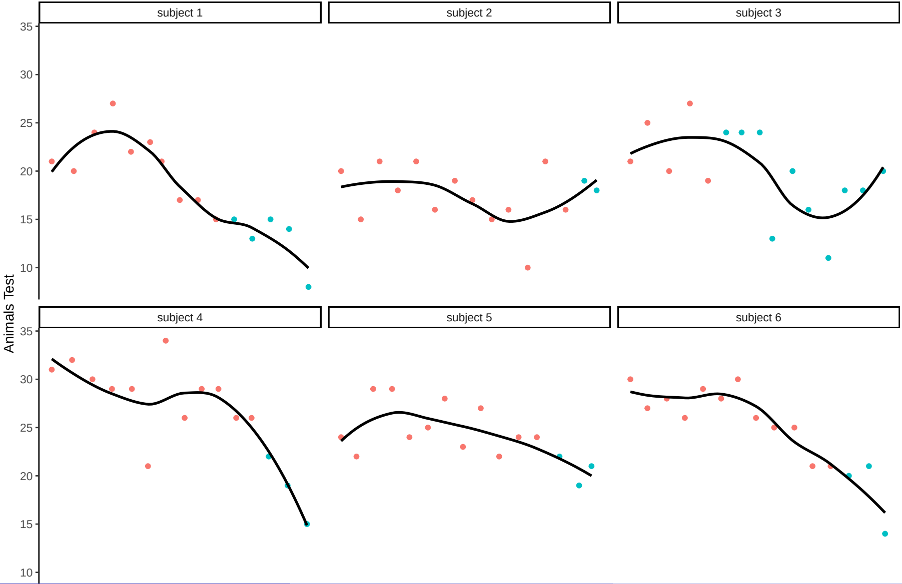
- Data collected by the National Alzheimer's Coordinating Center (NACC).
  - Established by the National Institute of Aging in 1990.
  - Centralizes neuropsychological data from 34 different research facilities.
  - Neuropsychological data include a number of cognitive tests repeated over time.

Model of Interest

- Studying the cognitive trajectory among those who transitioned from cognitively normal to MCI or Dementia during follow-up.
  - Interested in the effect of the APOE e4 allele on cognition.
  - 1,643 subjects in the analysis with a median of 6 visits.

$$\text{Animals} \sim (1 + \mathrm{I}\{\text{Transitioned to MCI or Dementia}\} + \text{APOE} + \text{Sex} + \text{APOE*Sex} + \text{Race} + \text{Age} + \text{Education}) * \text{Time}$$

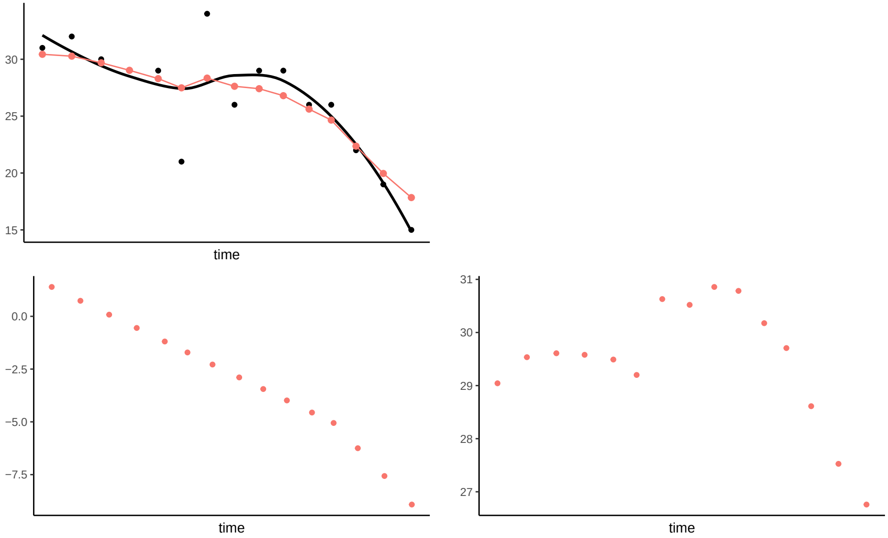
Data Characteristics



We Want a Model That. . .

- Captures general trajectory.
- Has the subject specific heterogeneity shown by the non-parametric method.
- We are able to interpret and make inference on effects of interest.

Welcome to the State Space Model



State Space Model Introduction

- State Space Models have been primarily used for time series data with a large number of time points and only a small number of chains observed.
- We are working to apply these models to a small number of time points and a large number of subjects.
  - Small t and large n are typically what we see in observational data.
- We wish to show that the State Space Model can be more accommodating than the commonly used linear mixed effect models (LMEM)(Laird and Ware, 1983; Diggle, Liang and Zeger, 1994).

State Space Model

A general linear state space model can be denoted as:

$$y_t = F_t \mu_t + v_t$$
$$\mu_t = G_t \mu_{t-1} + w_t$$

where at time t,

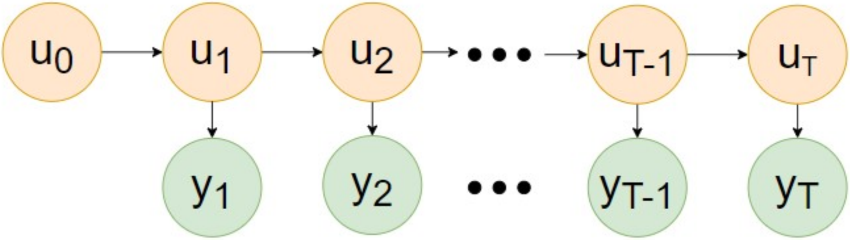
- y_t is the an $n \times 1$ observation vector.
- μ_t is the $q \times 1$ latent state vector, where q is the number of latent states.
- F_t is the $n \times q$ observation matrix.
- G_t is the $q \times q$ state transition matrix.

We assume v_t and w_t are independent identically distributed with distributions $v_t \sim N(0, V)$ and $w_t \sim N(0, W)$ respectively (Harvey, 1990; Durbin and Koopman, 2012) .

State Space Model Illustration

General Model:

$$y_t = F_t \mu_t + v_t$$
$$\mu_t = G_t \mu_{t-1} + w_t$$



Proposed Model

We wish to model the data according to a specific SSM, the Local Linear Trend Model (LLT),

$$y_{it} = \alpha_{it} + x_{it}^T \beta_t + \varepsilon_t$$

$$\mu_{it} = \begin{bmatrix} \alpha_{it} \\ \beta_t \end{bmatrix} = \begin{bmatrix} \alpha_{i(t-1)} \\ \beta_{(t-1)} \end{bmatrix} + \begin{bmatrix} \eta_{it} \\ 0_{p\times 1} \end{bmatrix}$$

Where $\alpha_0 \sim N(a_0, P_0)$, $\beta_0 \sim N(\beta, 0)$, $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$, and $\eta_{it} \sim N(0, \sigma_\eta^2)$.

- y_t is an $n \times 1$ observation vector where n indicates the number of subjects.
- α_t is an $n \times 1$ latent state vector.
  - Variation in α_t over time creates a dynamic moving average auto-correlation between observations y_t.
- X_t is an $n \times p$ matrix of time varying covariates (can be $X_t = t * X$ where X are baseline covarties).

What is α_t

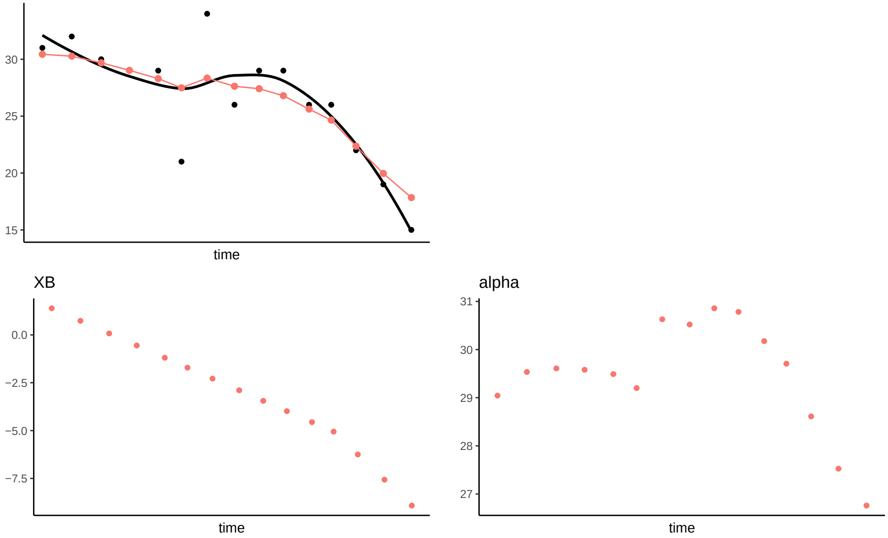
Consider the model,

$$y_{it} = \alpha_{it} + x_{it}^T \beta_t + \varepsilon_t$$
$$\mu_{it} = \begin{bmatrix} \alpha_{it} \\ \beta_t \end{bmatrix} = \begin{bmatrix} \alpha_{i(t-1)} \\ \beta_{(t-1)} \end{bmatrix} + \begin{bmatrix} \eta_{it} \\ 0_{p\times 1} \end{bmatrix}$$

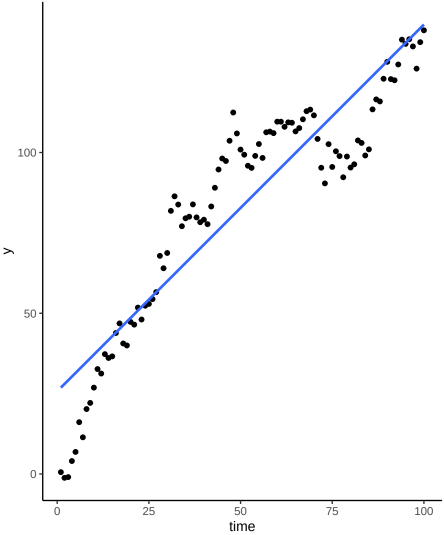
We can think of α_t as the underlying cognitive state not accounted for by covariates X_t. The α_t is there to capture unobserved effects on the outcome.

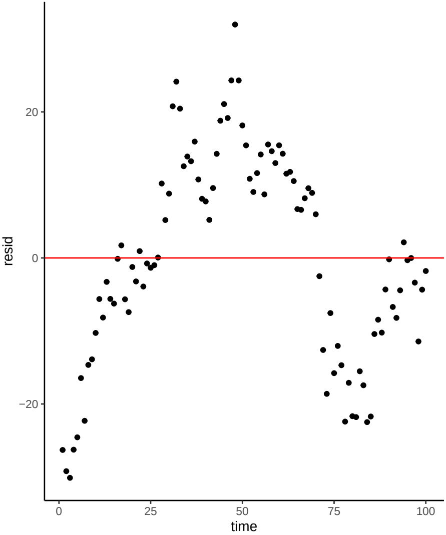
Notice $\alpha_t|\alpha_{t-1} \sim N(\alpha_{t-1}, \sigma_\eta^2)$. This means our next underlying cognitive state will be centered at the previous underlying cognitive state.

Revisited Plot



Single subject from an LLT





Auto-correlation

The correlation between observations at any two time points is called the auto-correlation.

Our proposed SSM model has the following auto correlation structure.

$$corr(y_{it}, y_{i(t+\tau)}) = \frac{t\sigma_\eta^2}{\sqrt{\sigma_\varepsilon^2 + t\sigma_\eta^2}\sqrt{\sigma_\varepsilon^2 + (t+\tau)\sigma_\eta^2}}$$

This is equivalent to a dynamic moving average covariance structure. If $\sigma_\eta^2 = 0$ then auto-correlation is 0 and our proposed model boils down to a LMEM.

$$y_t = \alpha_0 + X_t\beta + \varepsilon_t$$

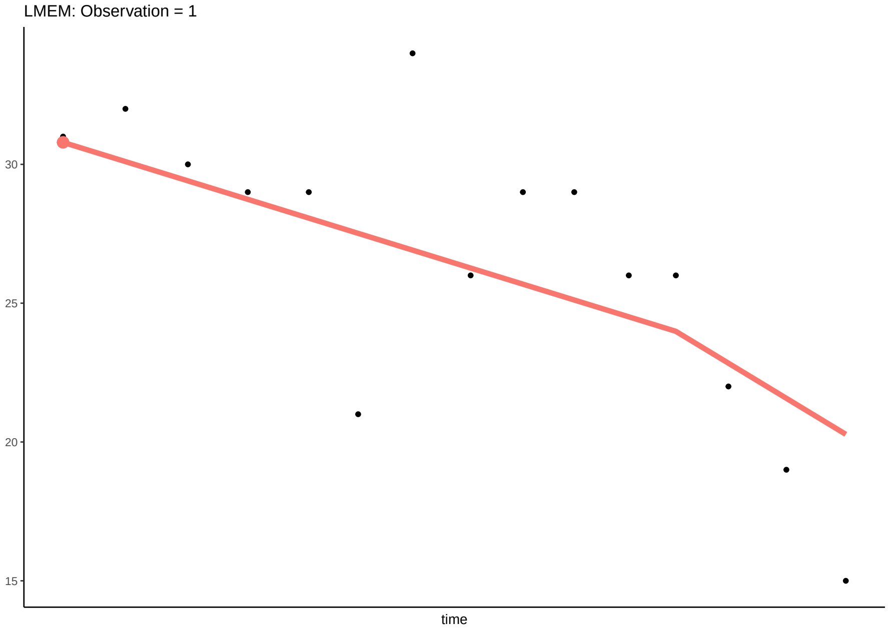
Accounting for Autocorrelation in LMEM Framework

- Many different autocorrelation techniques have been used.
- A common practice has been to model an AR(1) covariance structure on the errors.

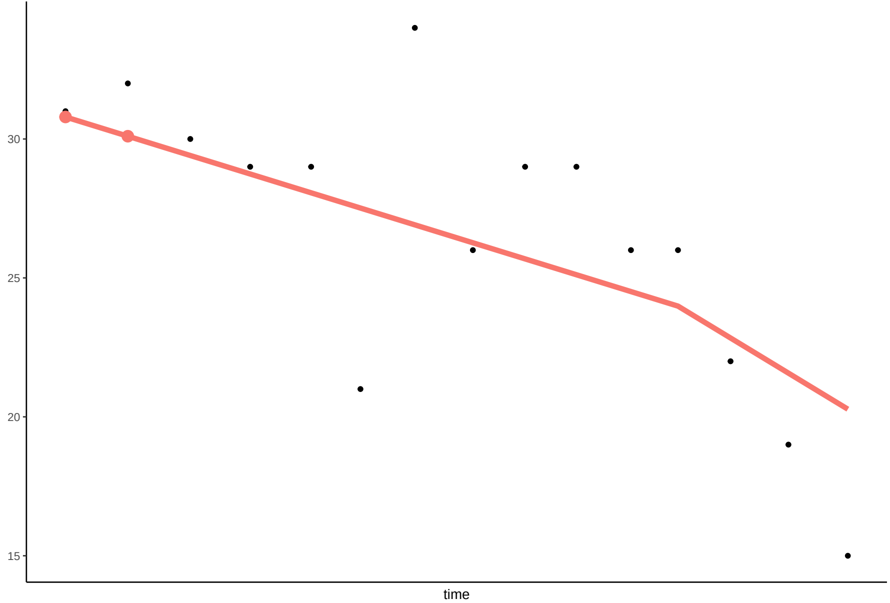
$$y_t = b_0 + X_t\beta + e_t, \quad b_0 \sim N(0, \sigma_b^2)$$
$$e_t = \rho e_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2), \quad -1 < \rho < 1$$

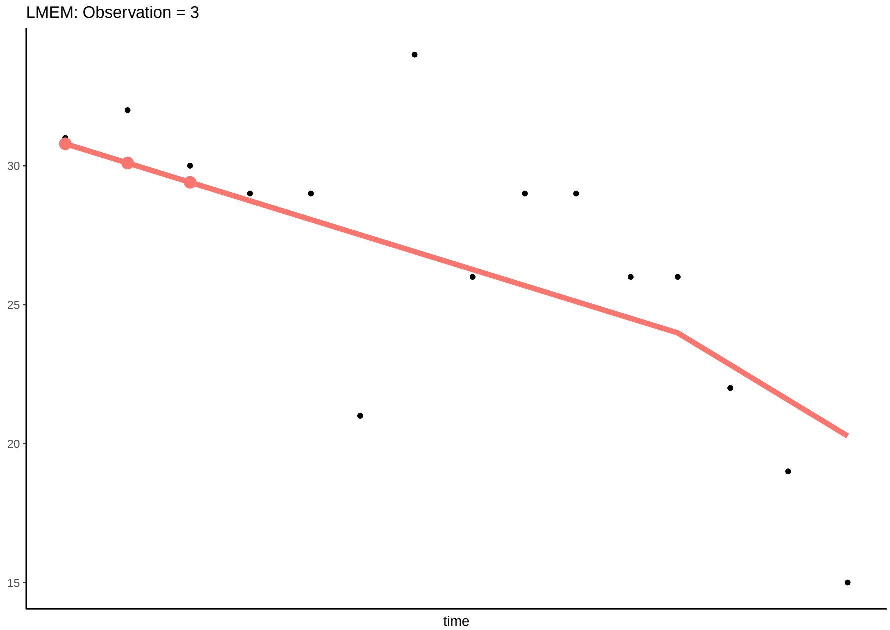
Lots of Plots

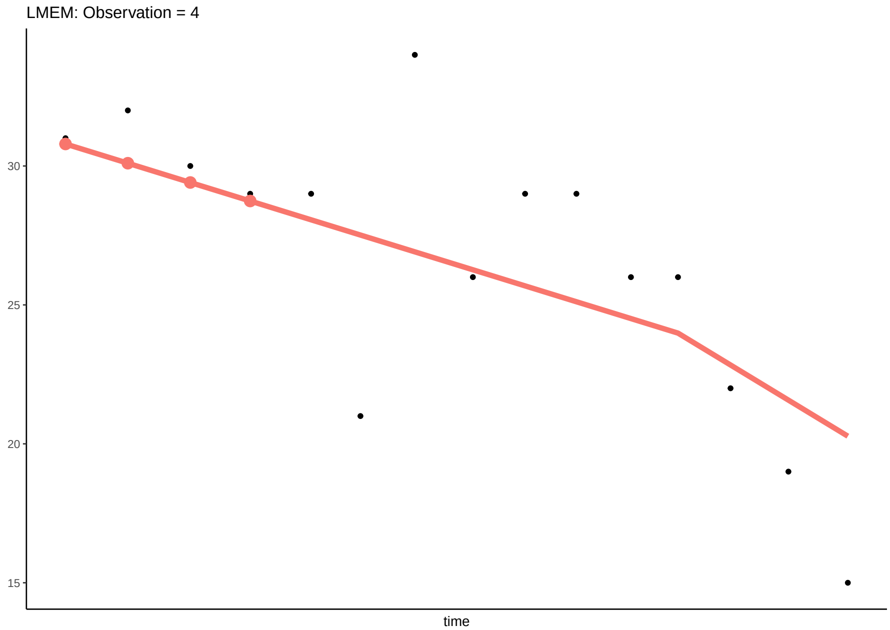
- The next few plots show the $E(Y_{t+\tau}|Y_t, \beta, X)$.
- Illustrates the LLT allows for dynamic changes in predicted trajectory while the LMEM and LMEM with AR(1) are very restrictive.


LMEM: Observation = 1
30
25
20
15
time

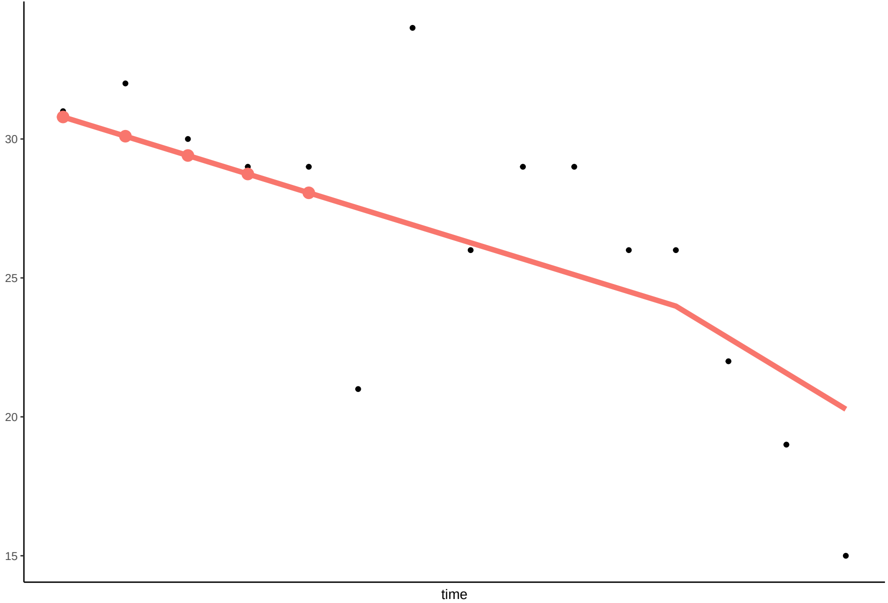
LMEM: Observation = 2



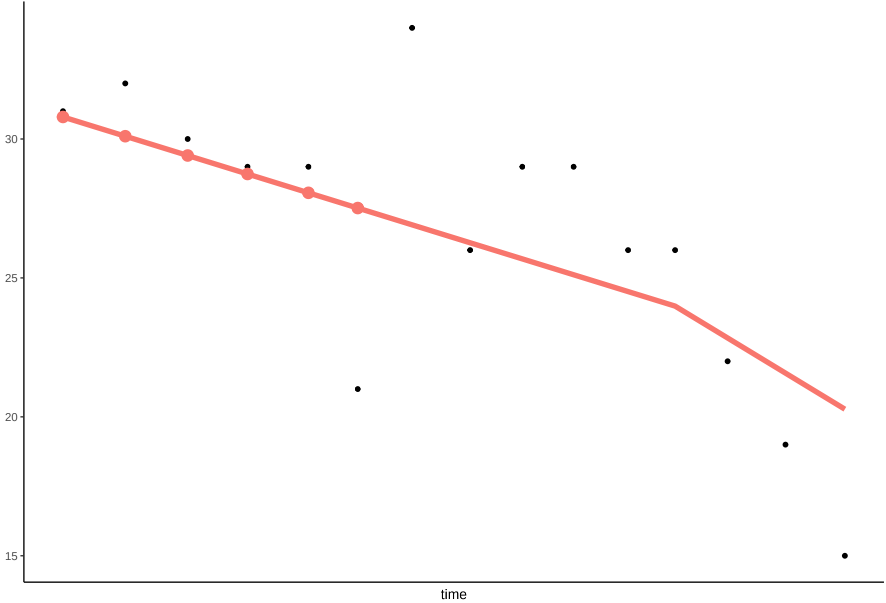

LMEM: Observation = 3
30
25
20
15
time

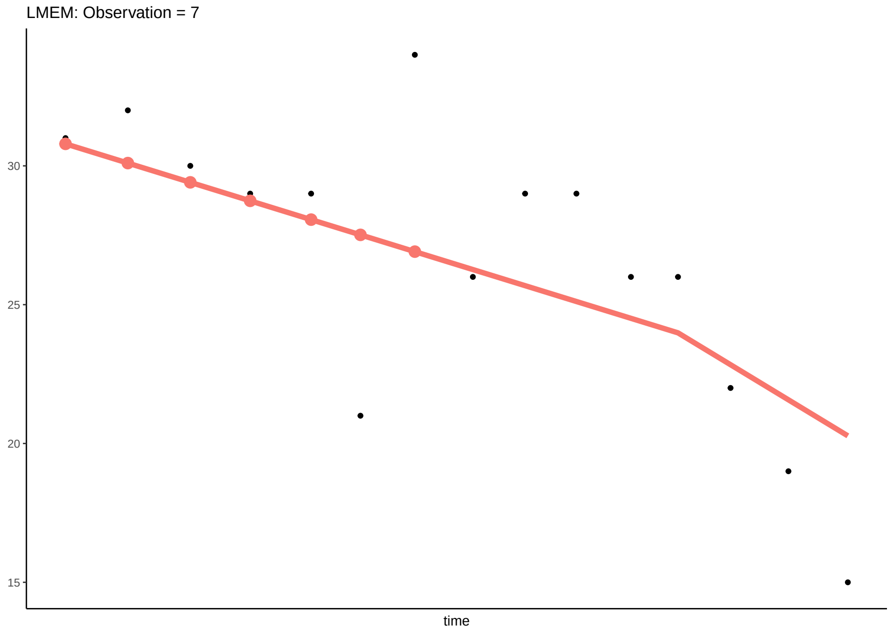

LMEM: Observation = 4
30
25
20
15
time

LMEM: Observation = 5

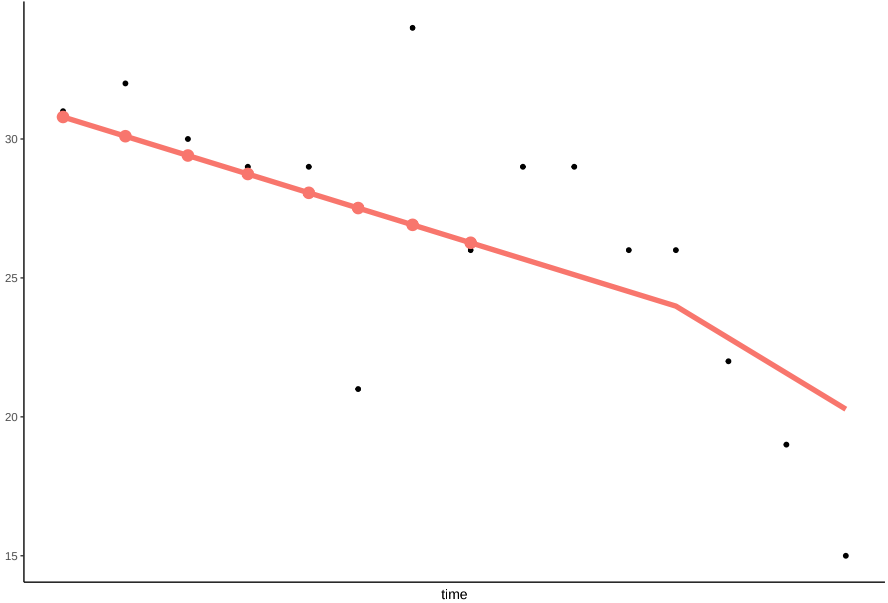


LMEM: Observation = 6

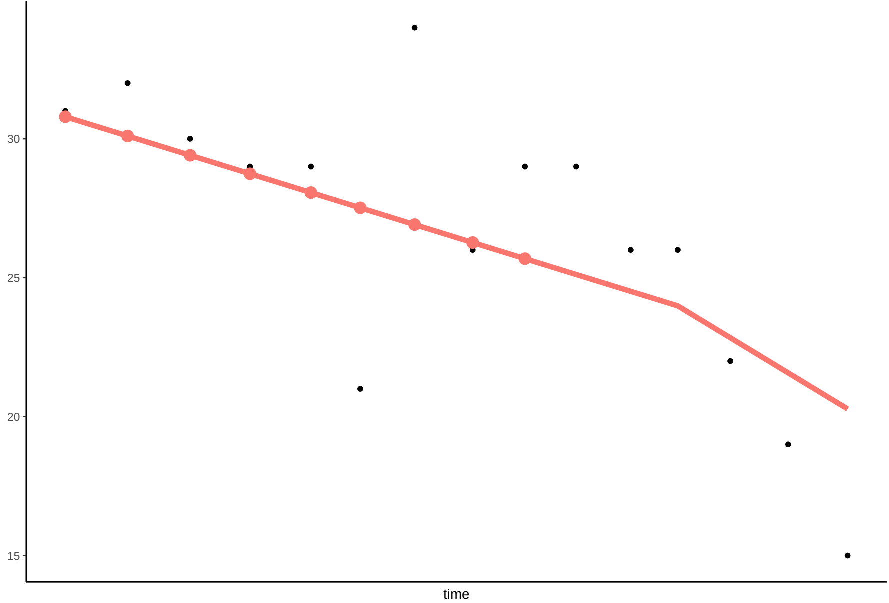


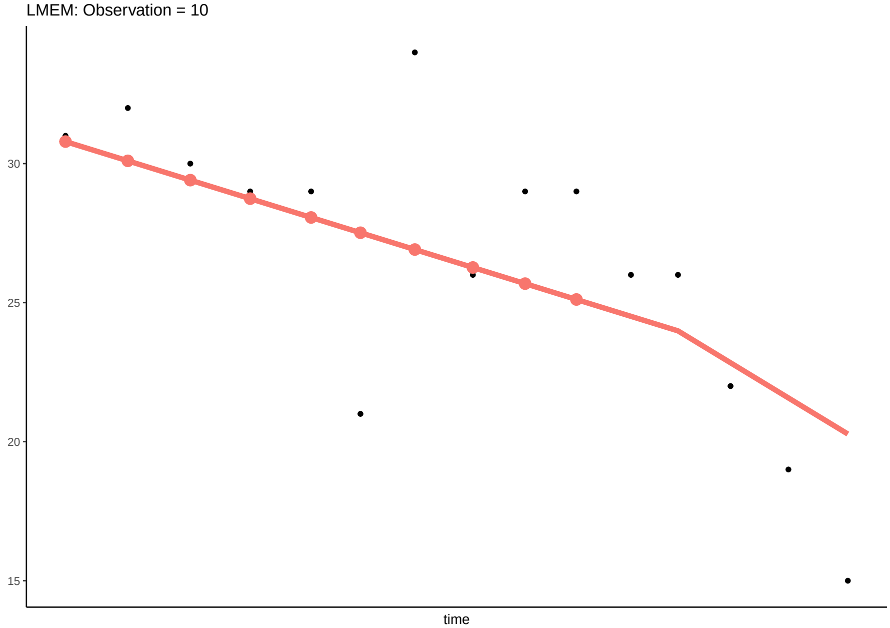

LMEM: Observation = 7
30
25
20
15
time

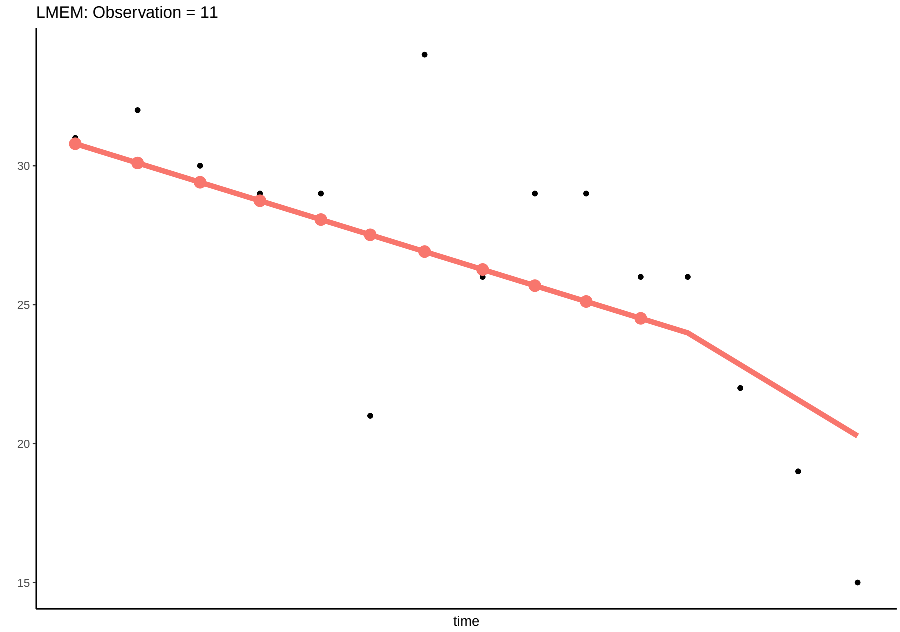
LMEM: Observation = 8



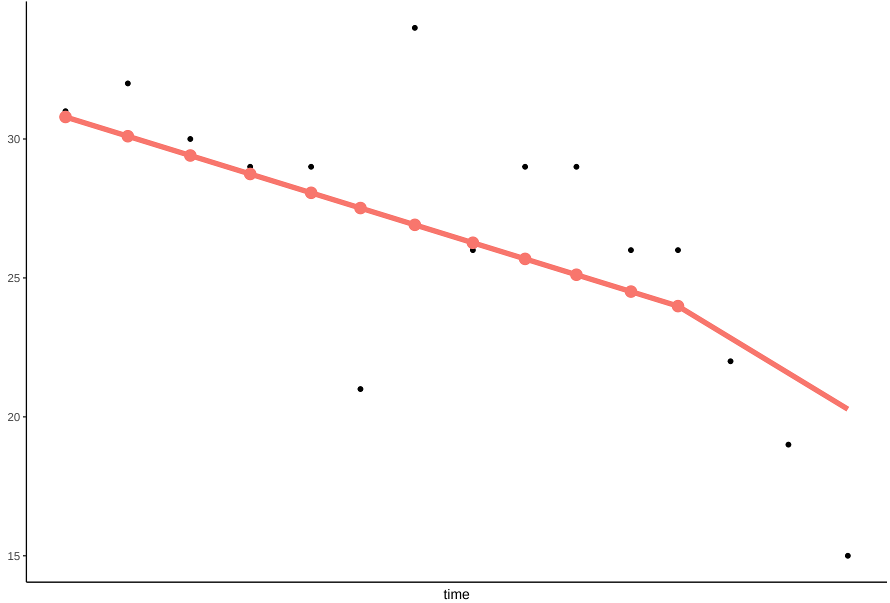
LMEM: Observation = 9

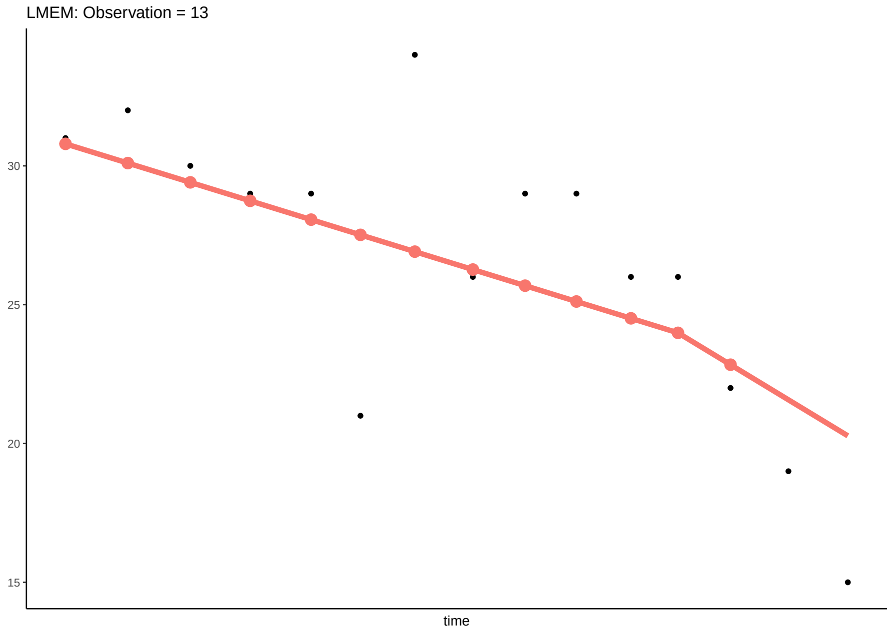



LMEM: Observation = 10
30
25
20
15
time


LMEM: Observation = 11
30
25
20
15
time

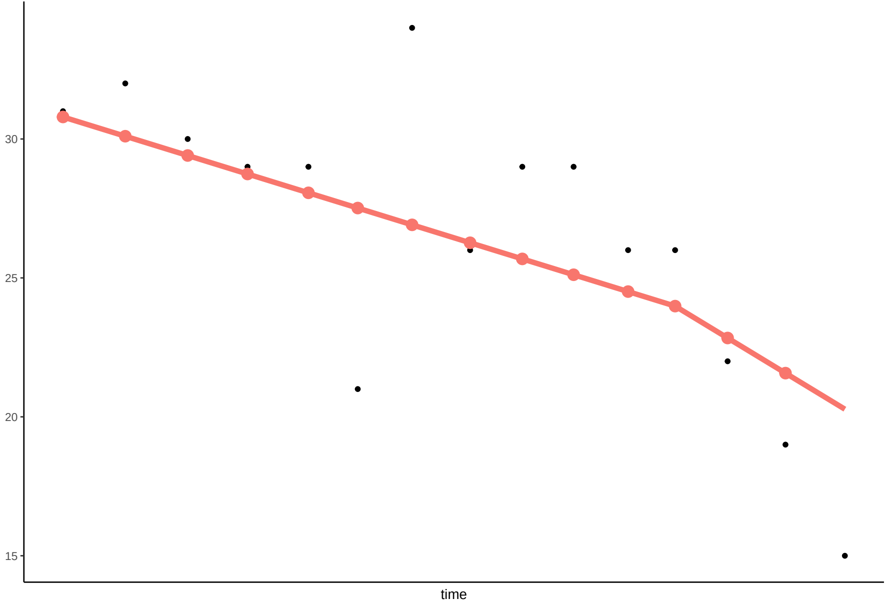
LMEM: Observation = 12



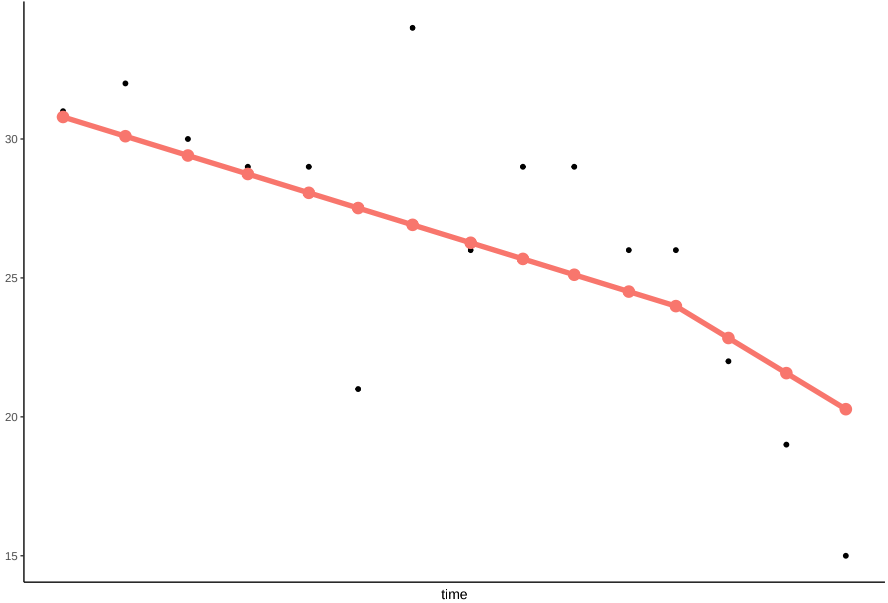


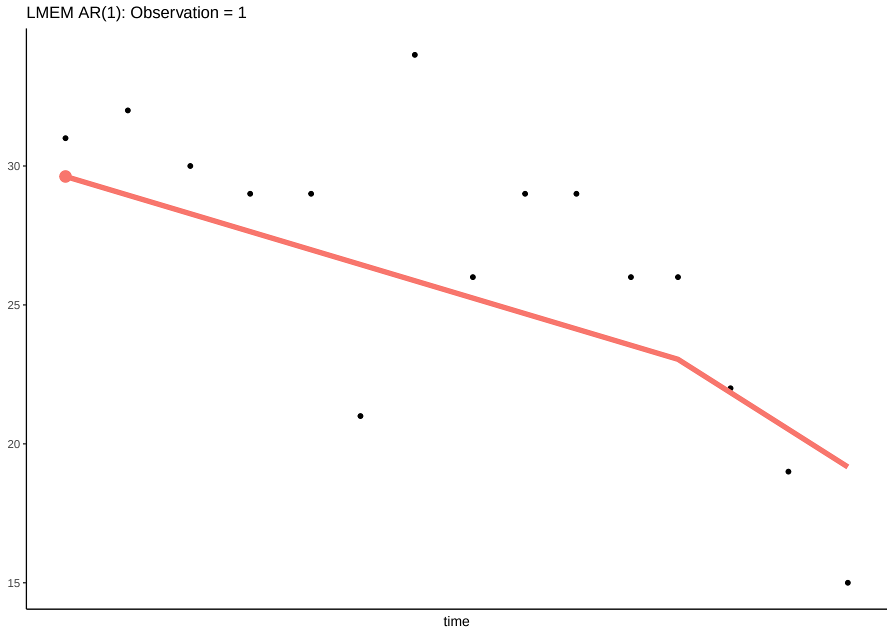
LMEM: Observation = 13
30
25
20
15
time

LMEM: Observation = 14



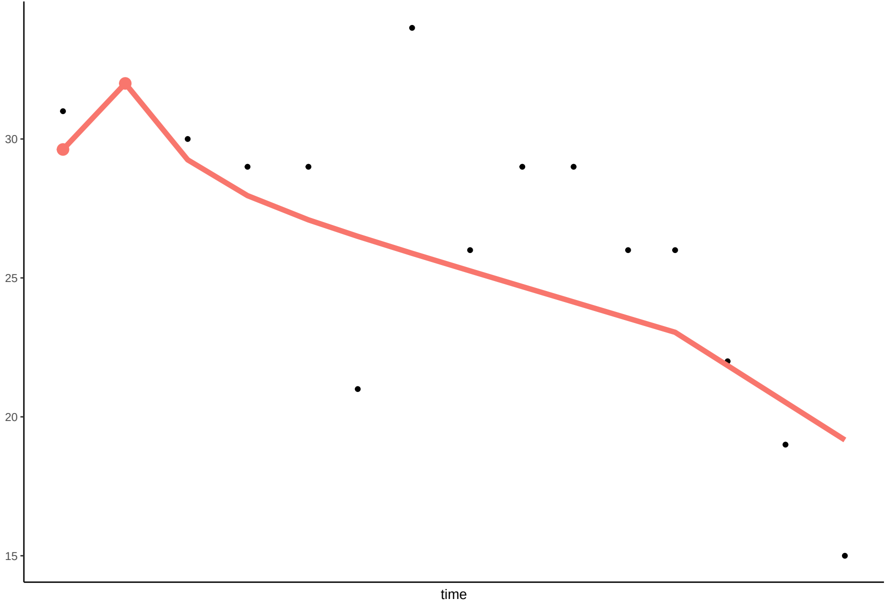
LMEM: Observation = 15

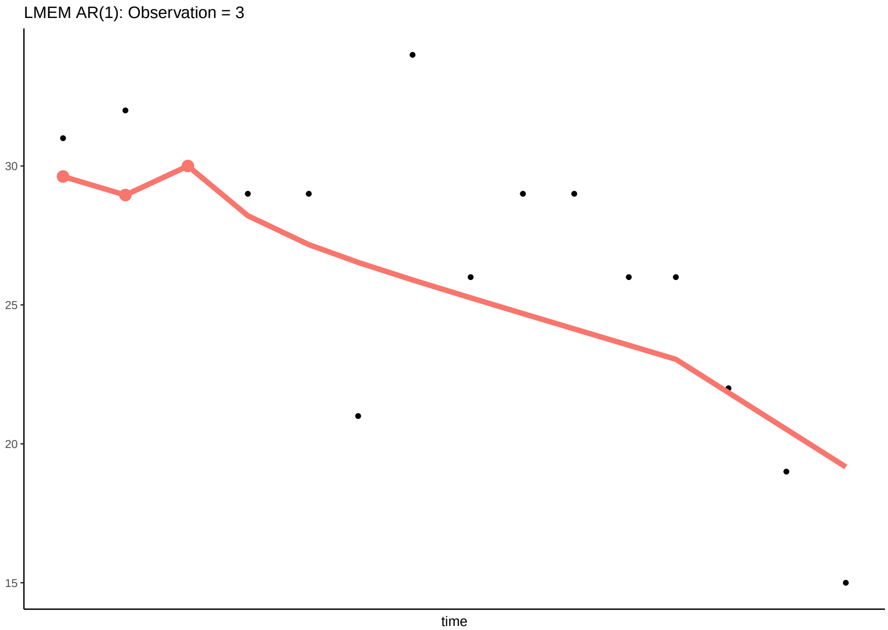


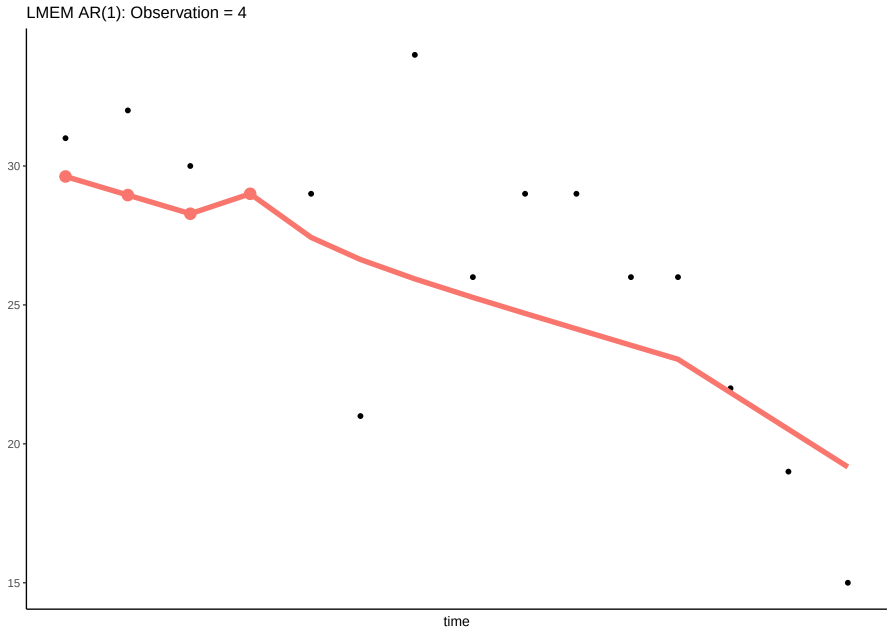


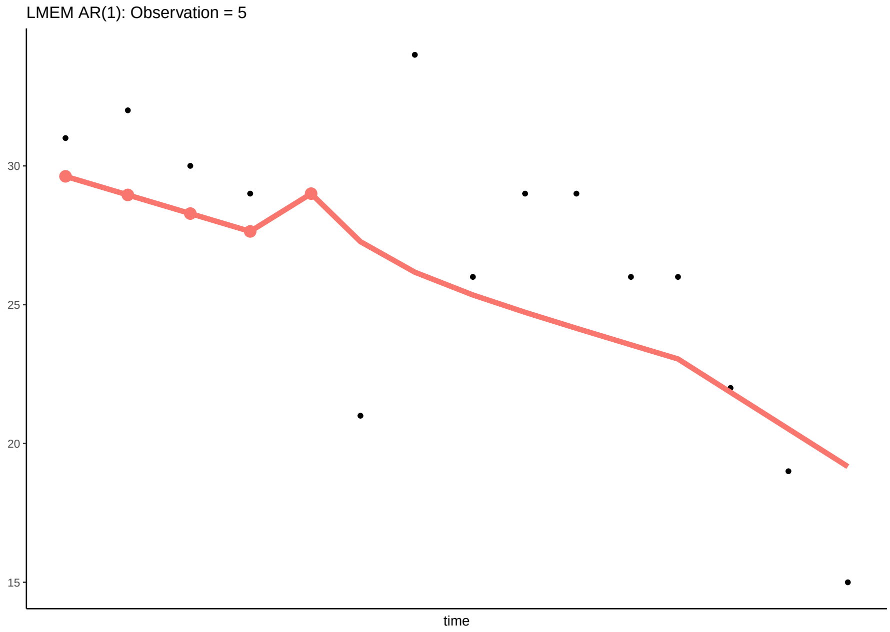
LMEM AR(1): Observation = 1
30
25
20
15
time

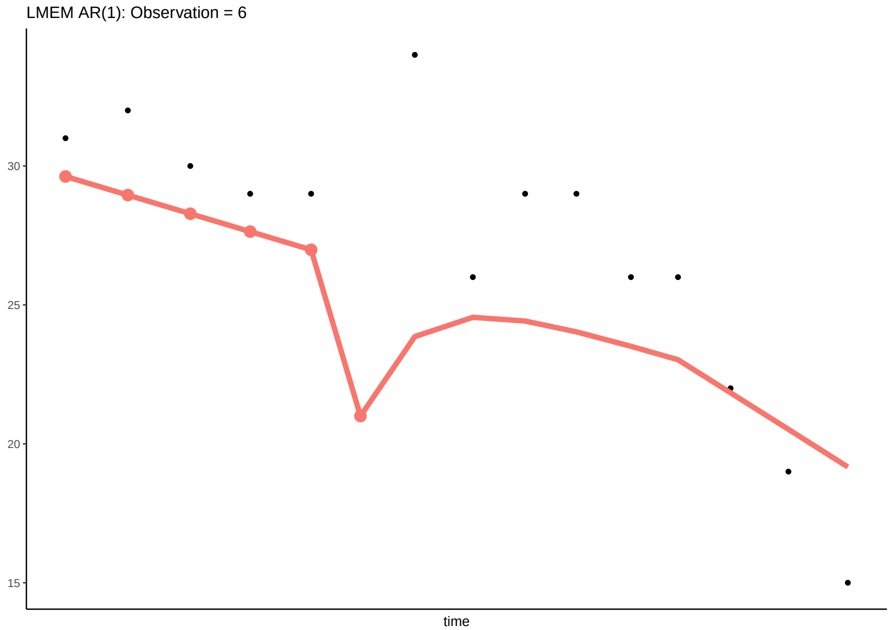
LMEM AR(1): Observation = 2

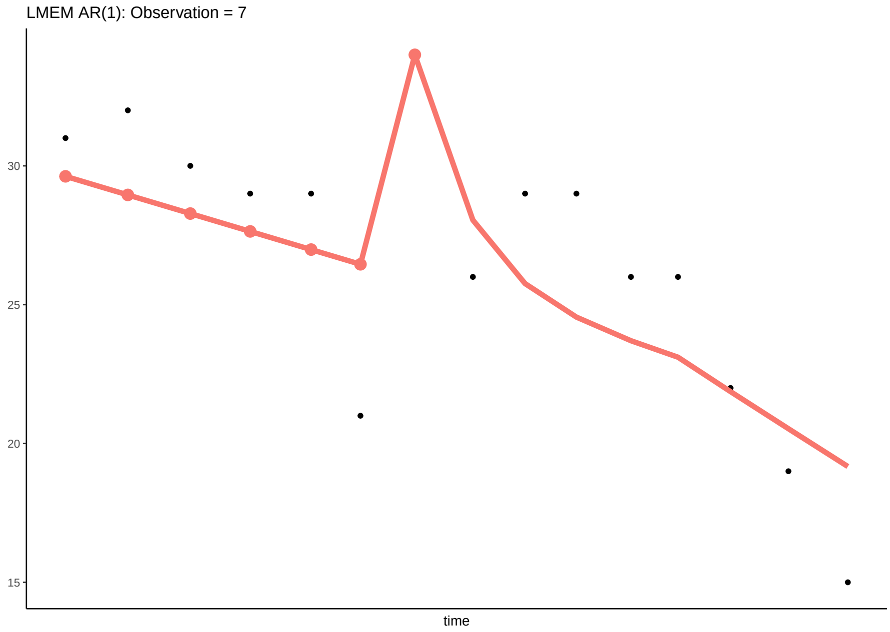


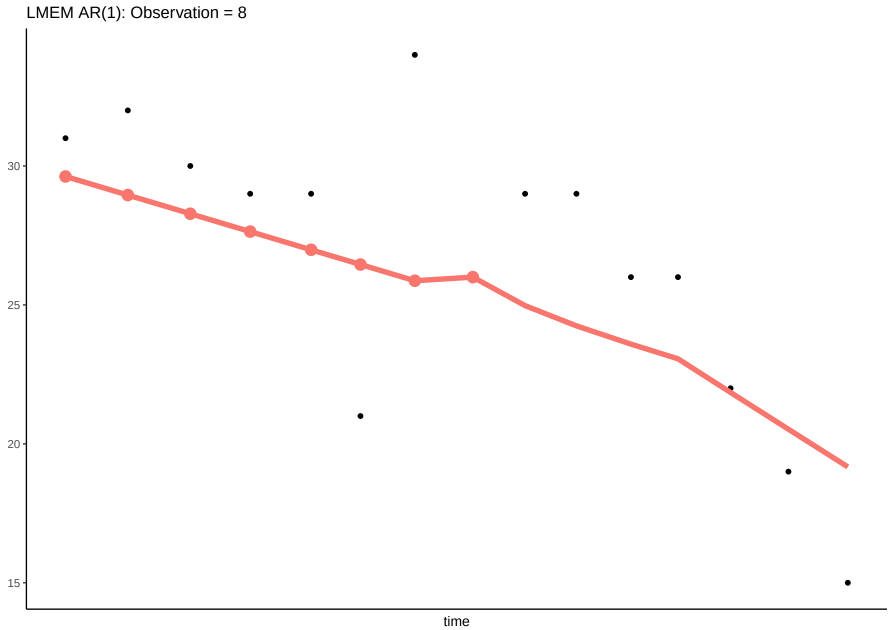

LMEM AR(1): Observation = 3
30
25
20
15
time


LMEM AR(1): Observation = 4
30
25
20
15
time

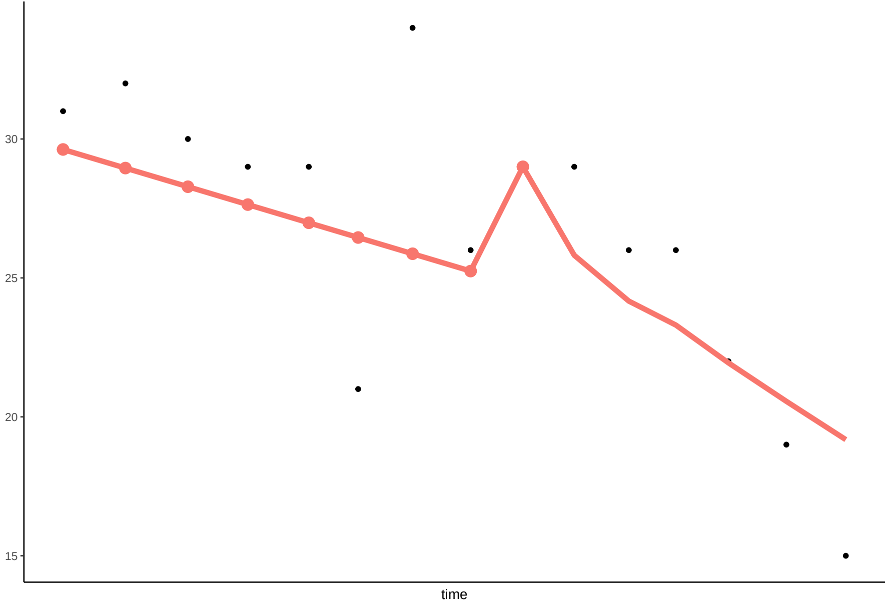

LMEM AR(1): Observation = 5
30
25
20
15
time

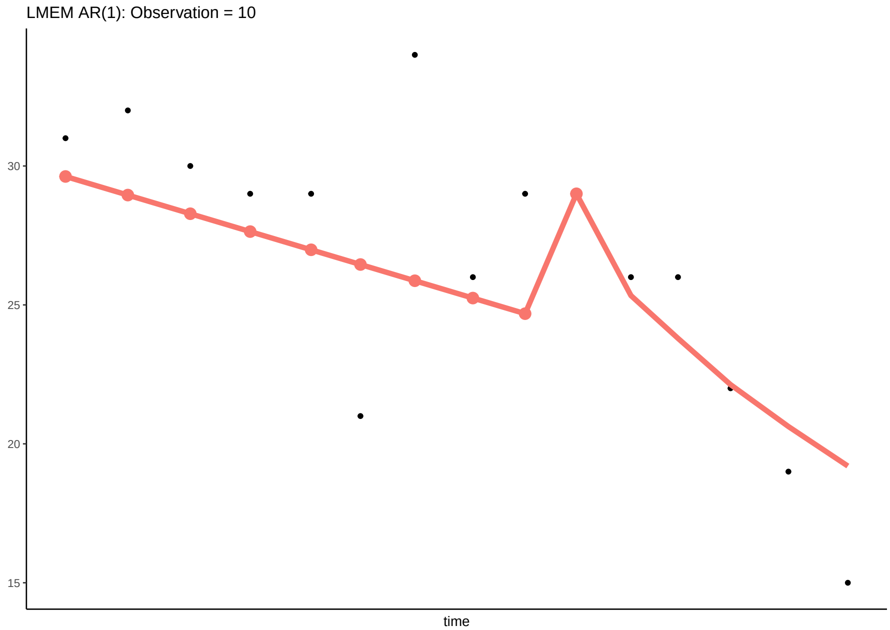

LMEM AR(1): Observation = 6
30
25
20
15
time


LMEM AR(1): Observation = 7
30
25
20
15
time

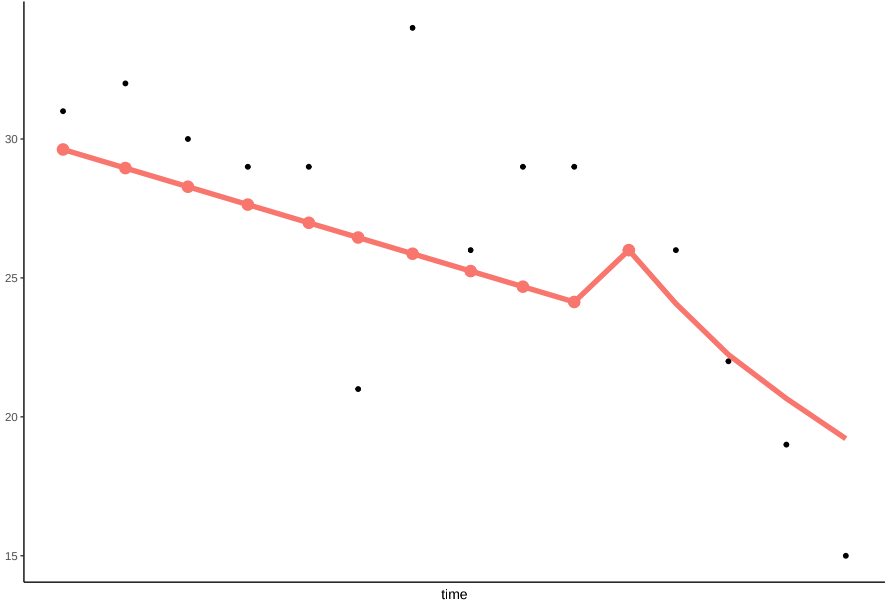

LMEM AR(1): Observation = 8
30
25
20
15
time

LMEM AR(1): Observation = 9

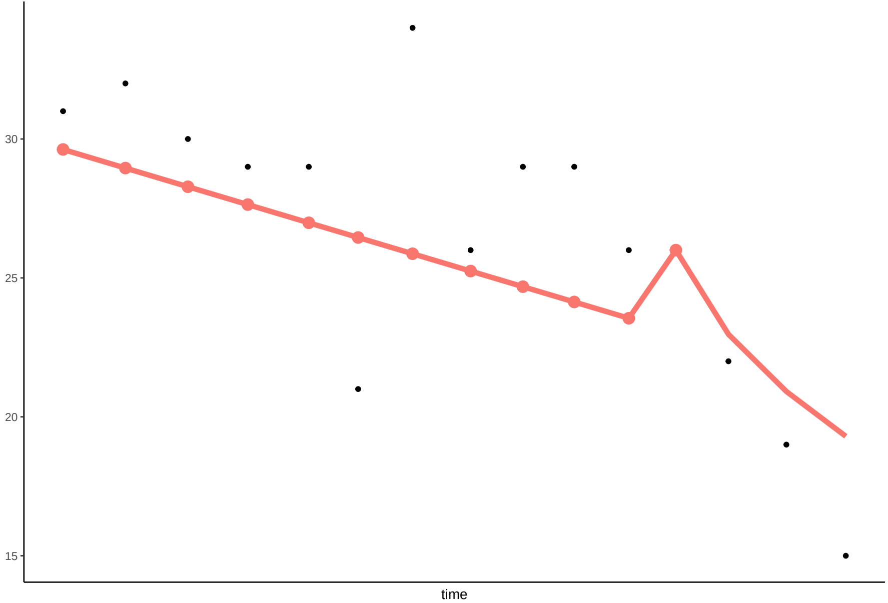



LMEM AR(1): Observation = 10
30
25
20
15
time

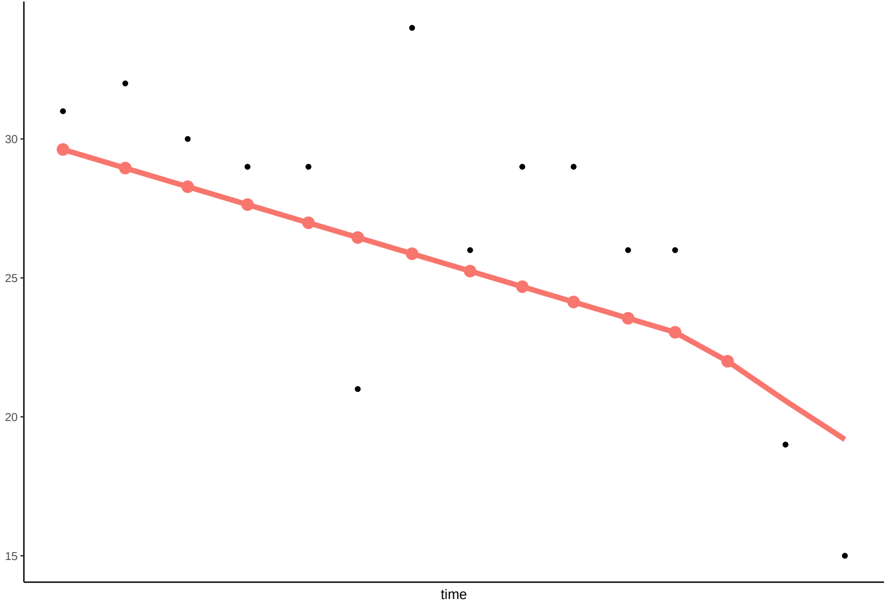
LMEM AR(1): Observation = 11

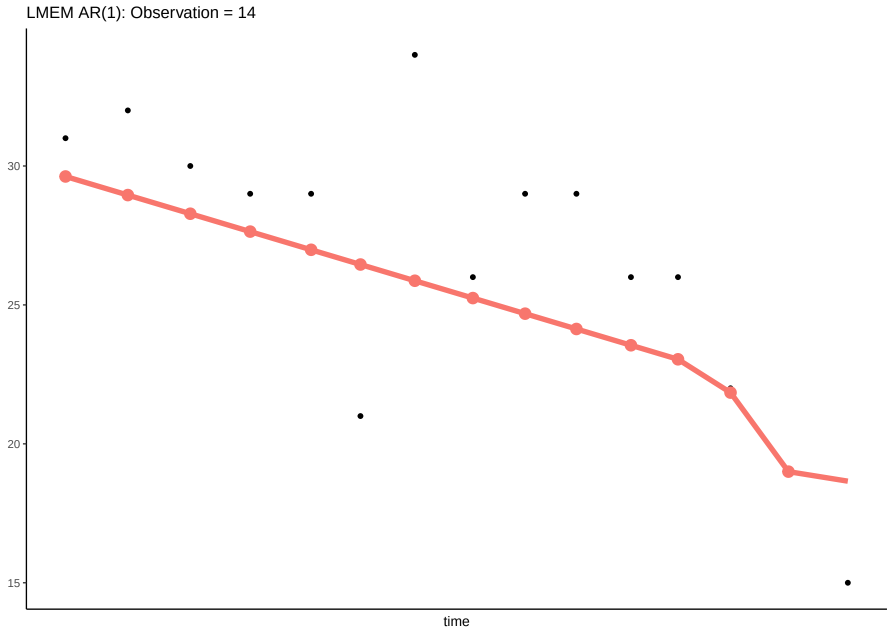


LMEM AR(1): Observation = 12

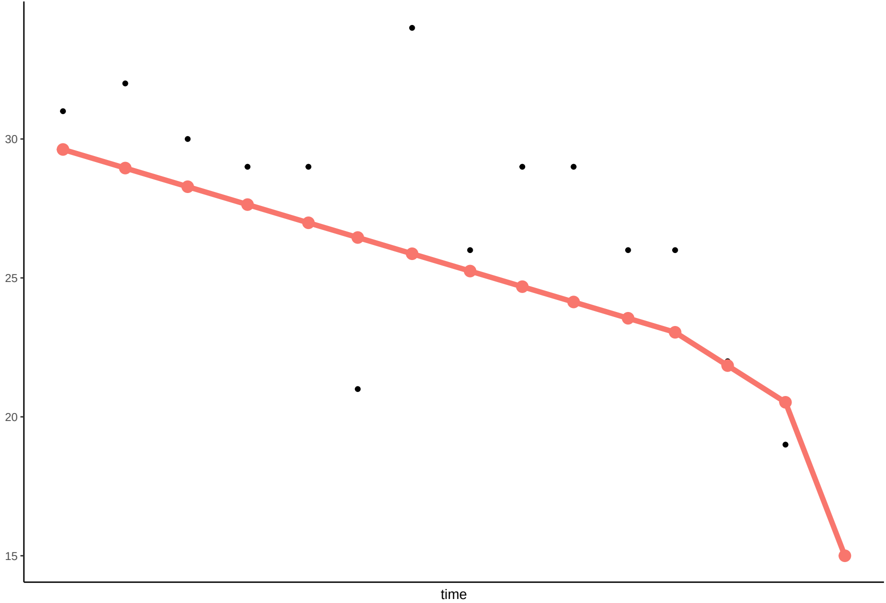


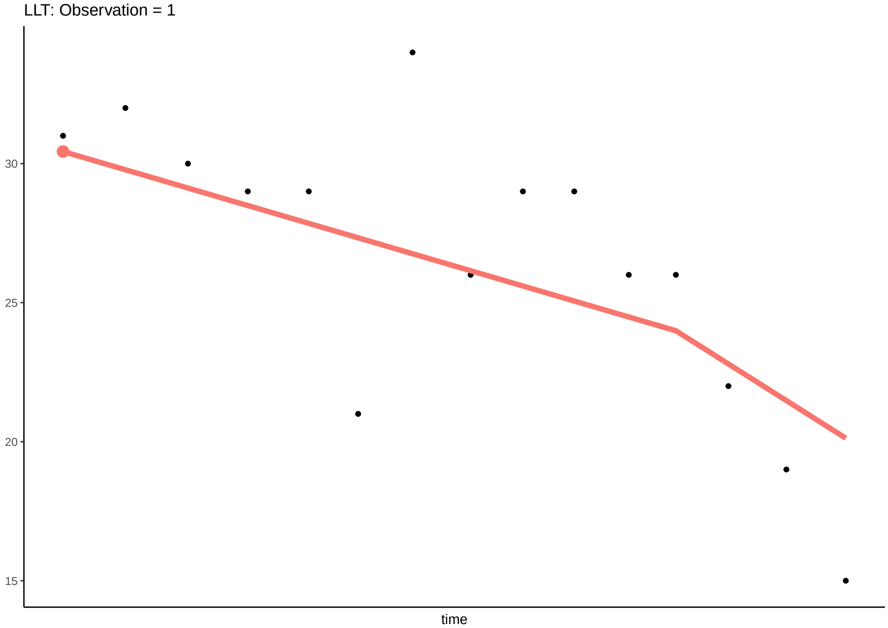
LMEM AR(1): Observation = 13



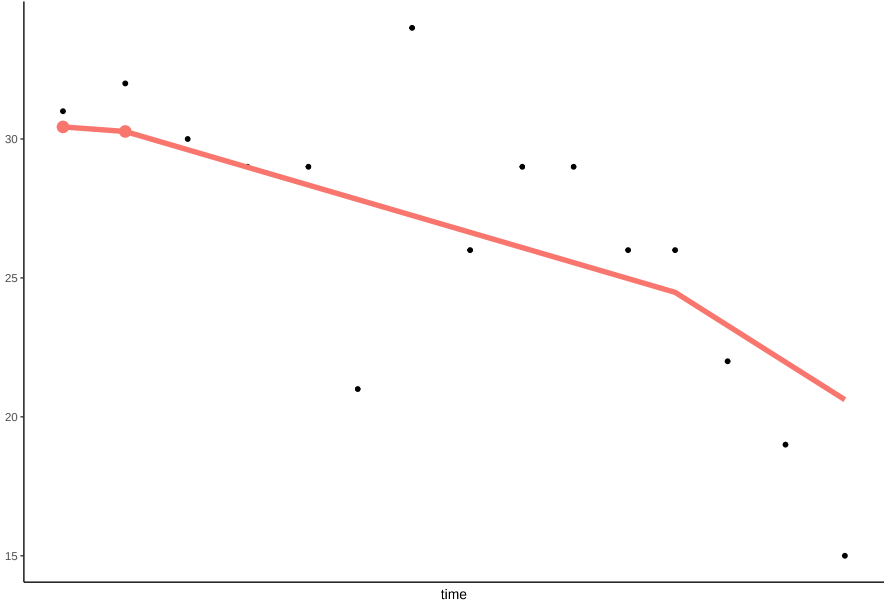

LMEM AR(1): Observation = 14
30
25
20
15
time

LMEM AR(1): Observation = 15

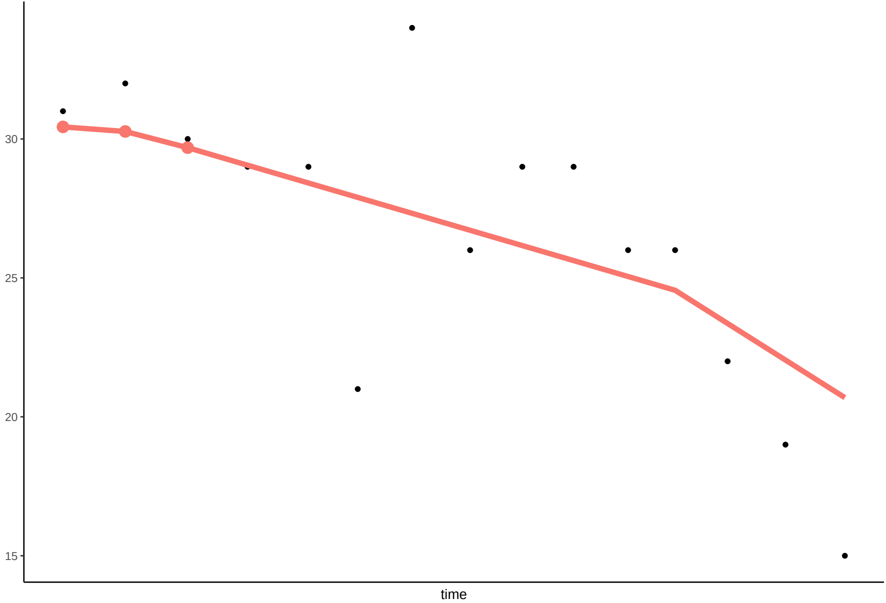



LLT: Observation = 1
30
25
20
15
time

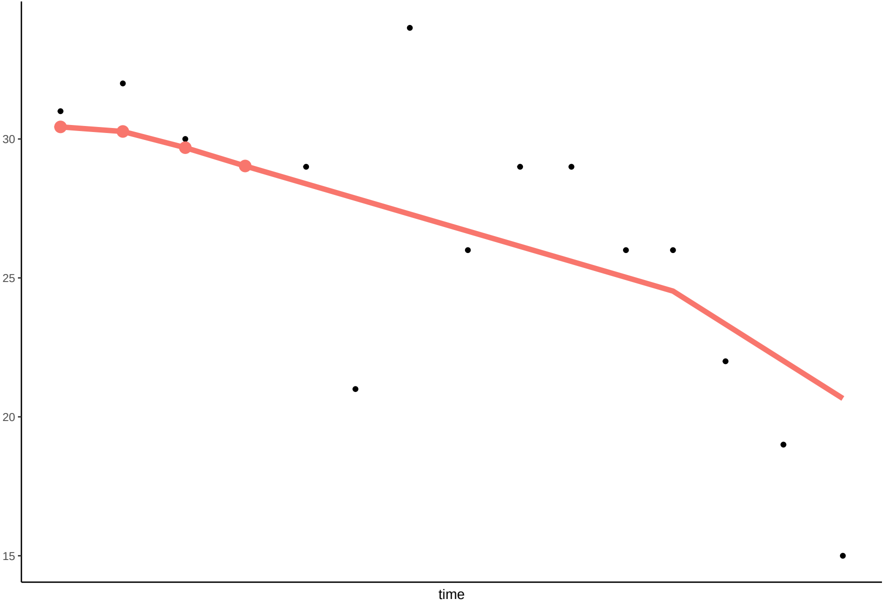
LLT: Observation = 2



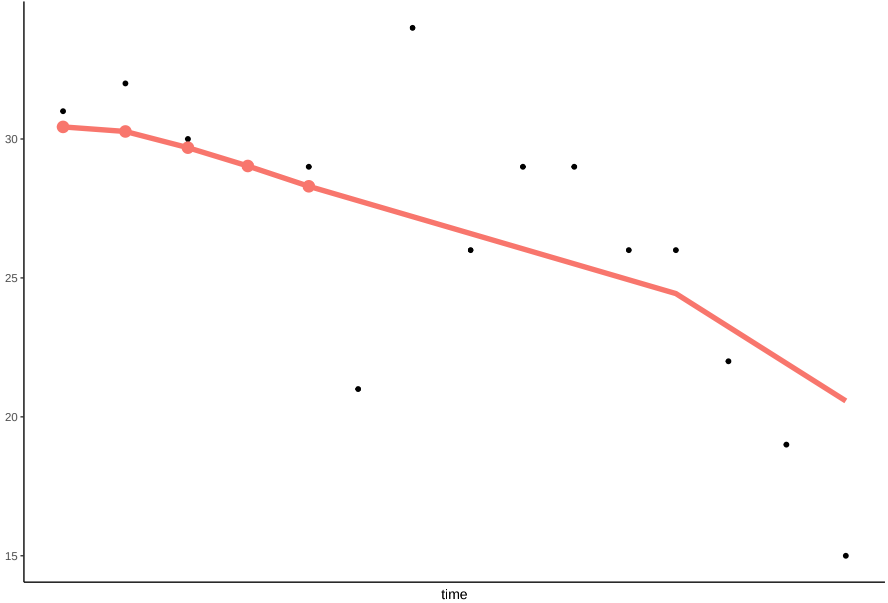
LLT: Observation = 3



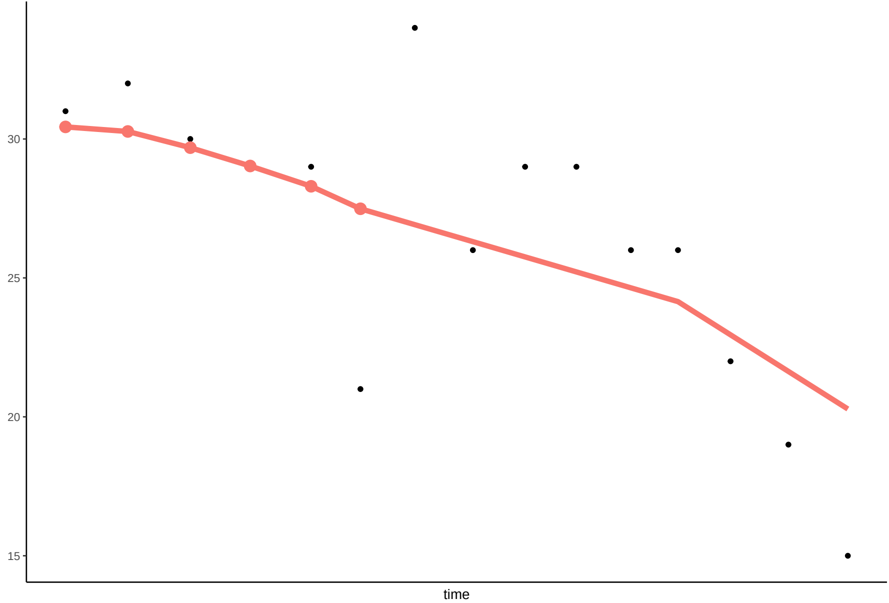
LLT: Observation = 4



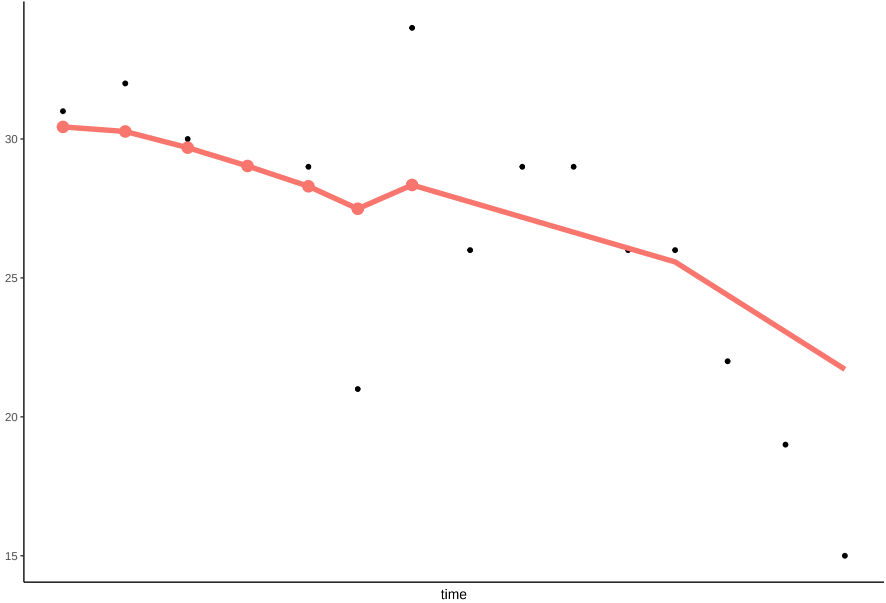
LLT: Observation = 5

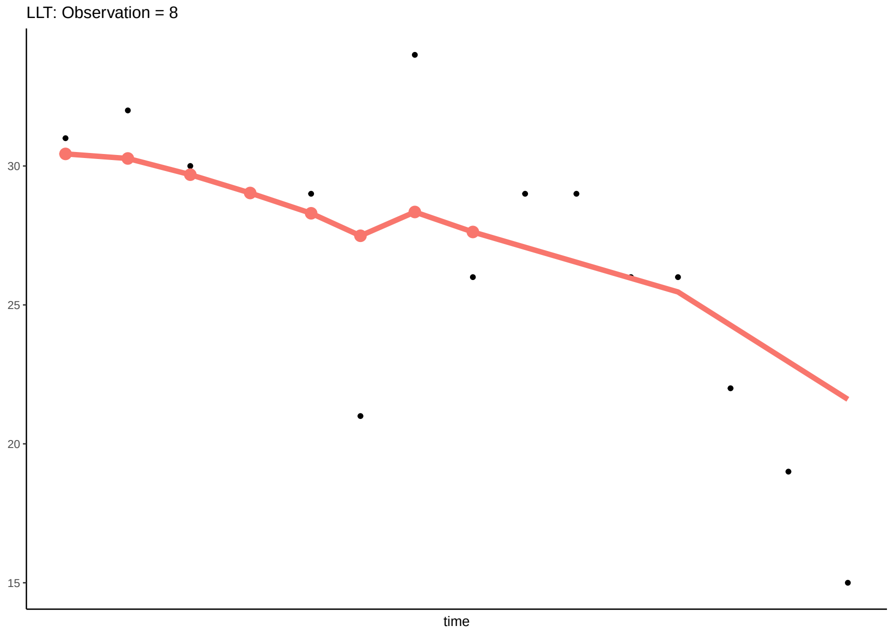


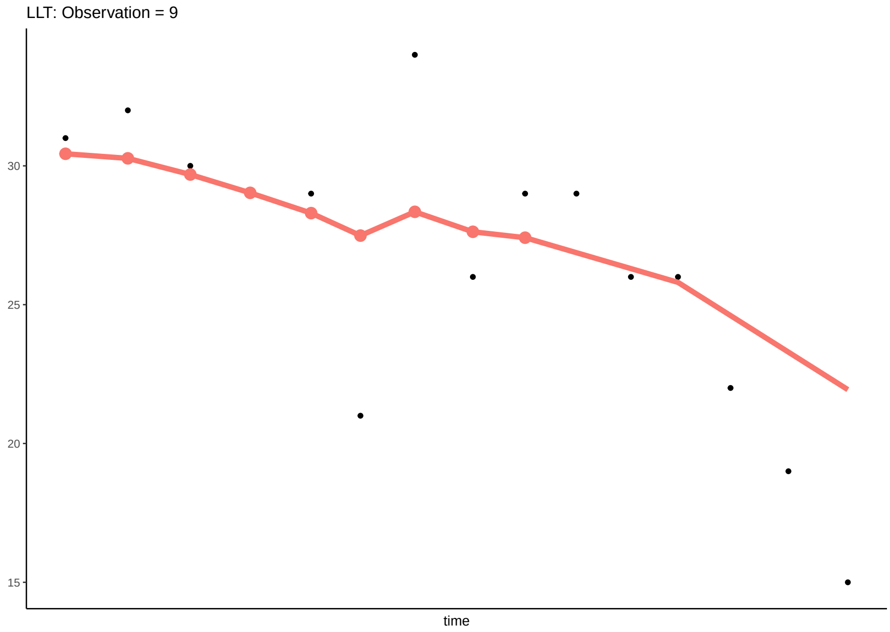
LLT: Observation = 6



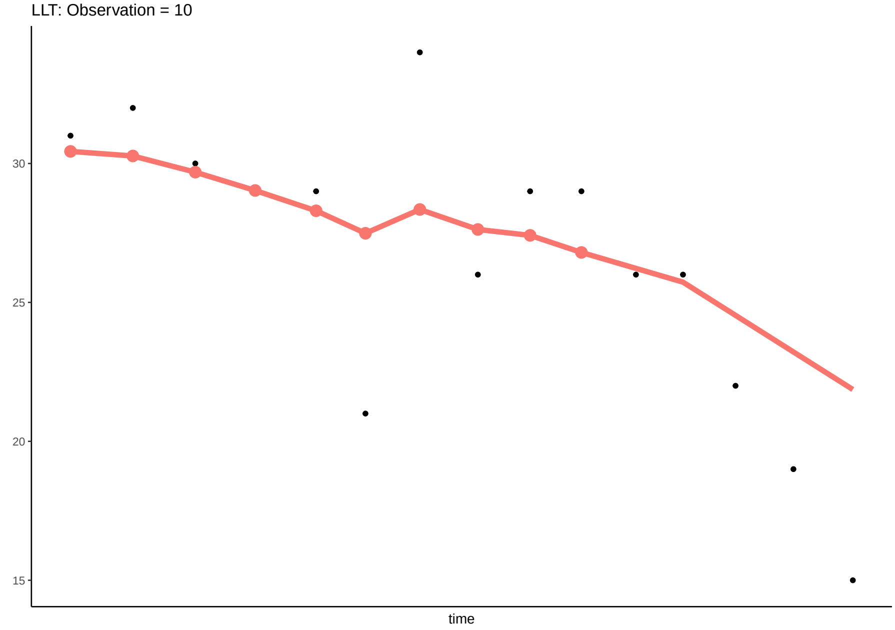
LLT: Observation = 7



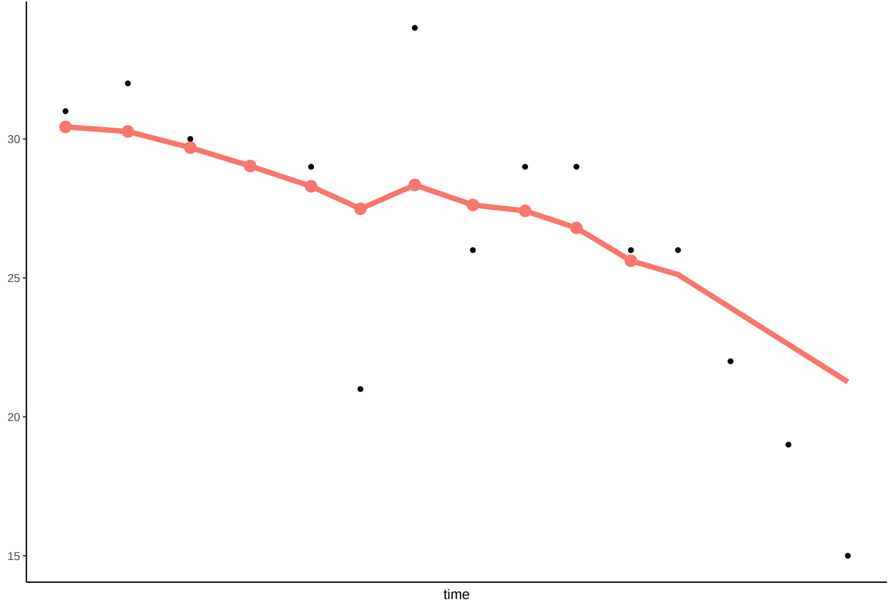

LLT: Observation = 8
30
25
20
15
time



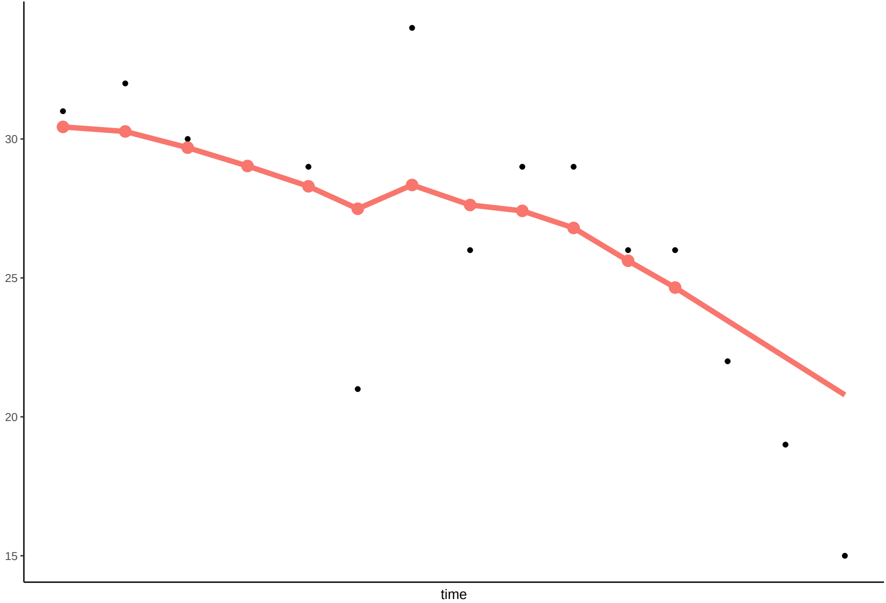
LLT: Observation = 9
30
25
20
15
time

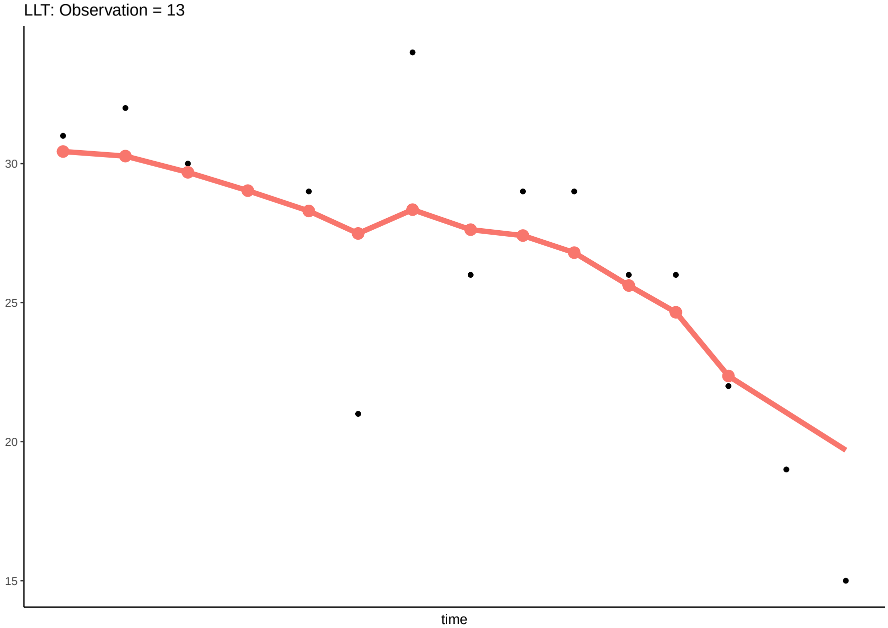

LLT: Observation = 10
30
25
20
15
time

LLT: Observation = 11

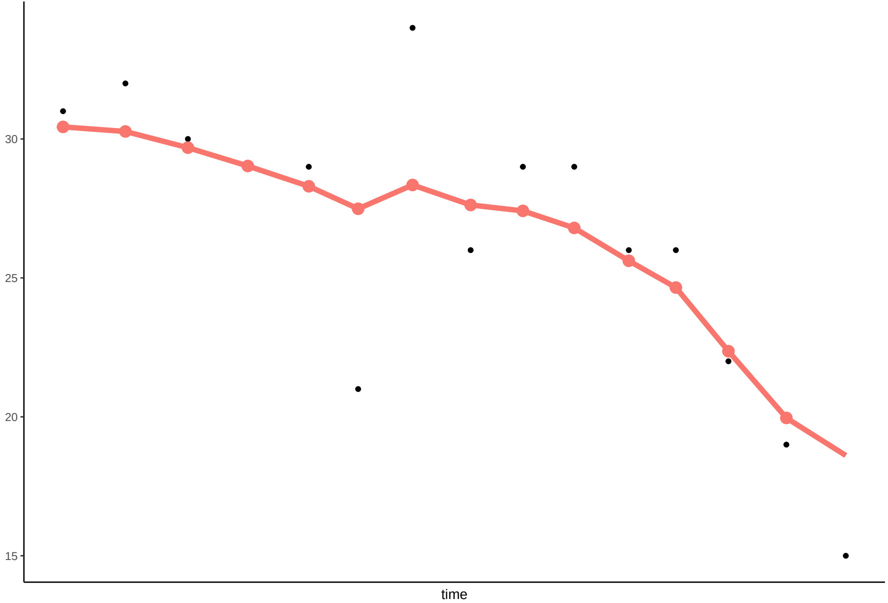


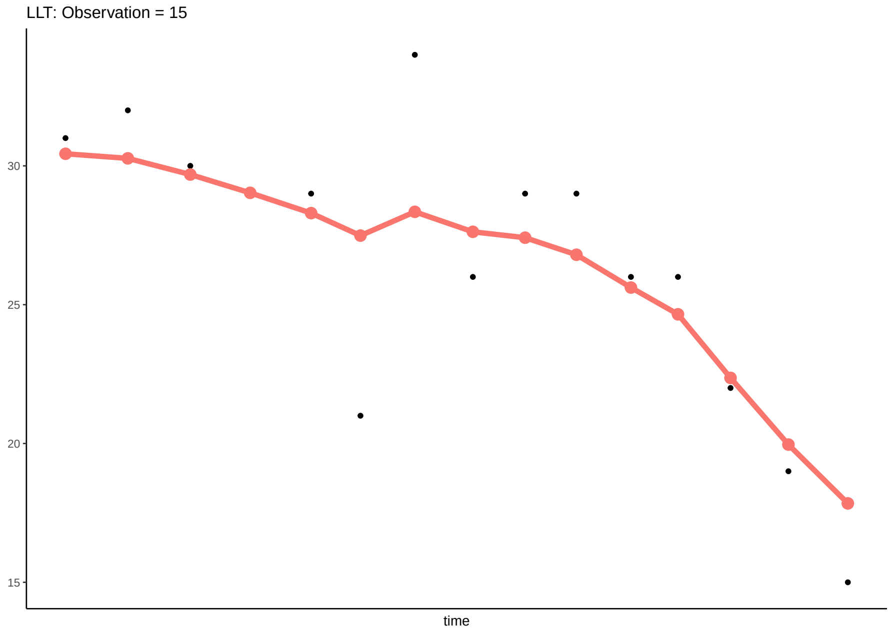
LLT: Observation = 12




LLT: Observation = 13
30
25
20
15
time

LLT: Observation = 14




LLT: Observation = 15
30
25
20
15
time

LLT Estimation

We can rewrite the proposed model to fit the state space model as follows,

$$y_t = \begin{bmatrix} I_n & X_t \end{bmatrix} \begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix} + \varepsilon_t$$

$$\begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} I_{(n+p)\times(n+p)} \end{bmatrix} \begin{bmatrix} \alpha_{t-1} \\ \beta_{t-1} \end{bmatrix} + \begin{bmatrix} \eta_t \\ 0_{p\times 1} \end{bmatrix}$$

- $F_t = \begin{bmatrix} I_n & X_t \end{bmatrix}$
- $v_t = \varepsilon_t$
- $w_t = \begin{bmatrix} \eta_t \\ 0_{p\times 1} \end{bmatrix}$
- $\mu_t = \begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix}$
- $G_t = I_{(n+p)\times(n+p)}$

Kalman Filter

The Kalman filter is a recursive algorithm to estimate the unobserved states conditioned on the observed data (Kalman, 1960; Durbin and Koopman, 2012). Let $\hat{\mu}_{i|j} = E(\mu_i|y_{1:j})$ and $P_{i|j} = var(\mu_i|y_{1:j})$.

Predicted state: $\hat{\mu}_{t|t-1} = G_t\hat{\mu}_{t-1|t-1}$

Predicted state covariance: $P_{t|t-1} = G_t P_{t-1|t-1} G_t' + W$

Innovation covariance: $S_t = F_t P_{t|t-1} F_t' + V$

Kalman Gain: $K_t = P_{t|t-1} F_t' S_t^{-1}$

Innovation: $\tilde{f}_t = y_t - F_t\hat{\mu}_{t|t-1}$

Updated state estimate: $\hat{\mu}_{t|t} = \hat{\mu}_{t|t-1} + K_t\tilde{f}_t$

Updated state covariance: $P_{t|t} = (I - K_t F_t)P_{t|t-1}$

Updated innovation: $\tilde{f}_{t|t} = y_t - F_t\hat{\mu}_{t|t}$

Kalman Smoother

Let $J_t = P_{t|t} G'_{t+1} + P^{-1}_{t+1|t}$. We can then calculate $E(\mu_t | y_{1:T})$ and $var(\mu_t | y_{1:T})$ using the following Kalman smoother equations.

$$E(\mu_t | y_{1:T}) = \hat{\mu}_{t|t} + J_t(\hat{\mu}_{t+1|T} - \hat{\mu}_{t+1|t})$$
$$var(\mu_t | y_{1:T}) = P_{t|t} - J_t G_{t+1} P_{t|t}$$

Setting Parameters

We assume $\mu_0 \sim N(u_0, P_0)$, however u_0 and P_0 are unknown.

- By initializing $u_0 = 0$ and $P_0 = \infty$ we are essentially putting a flat prior on μ_0.
- It has been shown $\hat{\mu}_{0|T}$ and $P_{0|T}$ quickly converge to u_0 and P_0 respectively for even small T (Kalman, 1960; Durbin and Koopman, 2012).

In our proposed model, $\mu_t = \begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix}$.

- $\hat{\beta}_{0|T}$ is then our estimate for β and has variance covariance $P_{\hat{\beta}} = [P_{0|T}]_{(n+1):(n+p),(n+1):(n+p)}$.
- We can then use $\hat{\beta}_{0|T}$ and $P_{\hat{\beta}}$ for inference on β.
  - $\hat{\beta} \overset{\text{asym}}{\sim} N(\beta, P_{\hat{\beta}})$.

Computational Challenges

For each iteration of the Kalman filter we must invert $var(Y_t|y_{1:(t-1)}) = S_t$.

- S_t is non-sparse as calculating $var(Y_t|y_{1:(t-1)})$ is a function of β_{t-1} which is shared between all observations.
- S_t is an $n \times n$, so as n increases there is an exponential increase in computation time.

Solution 1: Partitioning

A solution to solving inversion computational inefficiencies is to partition:

- Partition the subjects into k groups.
- Run the Kalman filter and smoother on each group independently to extract $\hat{\beta}_{0|T}^{(i)}$ and $P_{\beta}^{(i)}$ for i in $1, ..., k$.
- Use the estimate $\bar{\beta} = \frac{\sum_{i=1}^{k} \hat{\beta}_{0|T}^{(i)}}{k}$.
  - $\bar{\beta} \sim N(\beta, \frac{\sum_{i=1}^{k} P_{\hat{\beta}^{(i)}}}{k^2})$

Solution 2: Bayesian Gibb's Sampling Approach

- For the Bayesian approach we use a Gibb's sampler.
- Instead of calculating β in the Kalman filter, we can estimate it separately.
- The model,

$$
\begin{aligned}
y_t &= \alpha_t + X_t\beta + \varepsilon_t \\
\alpha_t &= \alpha_{t-1} + \eta_t
\end{aligned}
$$

Gibb's Sampling

- Gibb's sampling is a method to gain an approximate sample from a posterior distribution for a given variable (Gelfand-Smith, 1990).
- It works by:
  - calculating the distribution of a variable conditioned on all other unknown variables, known as the posterior distribution.
  - sampling from the posterior distribution and assigning the new sample to the variable.
  - calculate the posterior of the next variable and continue to sample, update, and recalculate the other posteriors.
  - The process is commonly repeated thousands of times.
- We need to calculate the posterior for $\alpha_{1:T}, \beta, \sigma_\varepsilon^2, \sigma_\eta^2$.

Posterior of α

- Notice, if we are conditioning on β for the posterior $\alpha_{1:T}|...$ then each y_{it} is independent and we can run the Kalman filter chains independently.
- Let $y_t^* = y_t - X_t\beta$, then the model becomes

$$
\begin{aligned}
y_t^* &= \alpha_t + \varepsilon_t \\
\alpha_t &= \alpha_{t-1} + \eta_t
\end{aligned}
$$

- We can then run a forward Kalman filter with a backward sampler to sample from the posterior of $\alpha_{1:T}$ (Fruhwirth-Schnatter, 1994)

Posterior of β

- We let $\beta \sim N(\theta, \sigma_\beta^2)$
- The posterior is $\beta | ... \sim N(\Sigma^{-1}B, \sigma_\varepsilon^2 \sigma_\beta^2 \Sigma^{-1})$ where,
- $B = \sigma_\beta^2 (\sum_{t=1}^T y_t - \alpha_t)' X_t - \sigma_\varepsilon^2 \theta$
- $\Sigma = (\sigma_\beta^2 \sum_{t=1}^T X_t' X_t) + \sigma_\varepsilon^2 I_p$

The Gibbs Sampling Algorithm

1. Select prior parameters for $\theta, \sigma_\beta^2, a_0, b_0, c_0, d_0$.
2. Let $\beta^{(0)} = \theta$, $\sigma_\eta^{2(0)} = \frac{d_0/2}{1+c_0/2}$, and $\sigma_\varepsilon^{2(0)} = \frac{b_0/2}{1+a_0/2}$.
3. Run a forward-filtering backward sampling procedure as described above conditioning on $\beta^{i-1}, \sigma_\eta^{2(i-1)}, \sigma_\varepsilon^{2(i-1)}$ and set the samples equal to $\alpha^{(i)}$ for the i^{th} iteration.
4. Sample σ_η^{2*} from $IG(\frac{nT+a_0}{2}, \frac{\sum_{t=1}^{T}(\alpha_t^{(i)}-\alpha_{t-1}^{(i)})^2+b_0}{2})$ and set $\sigma_\eta^{2(i)} = \sigma_\eta^{2*}$.
5. Sample σ_ε^{2*} from $IG(\frac{nT+c_0}{2}, \frac{d_0+\sum_{t=1}^{T}(y_t-X_t\beta^{(i-1)}-\alpha_t^{(i)})^2}{2})$ and set $\sigma_\varepsilon^{2(i)} = \sigma_\varepsilon^{2*}$.
6. Sample β^* from $N(\Sigma^{-1}B, \sigma_\varepsilon^2\sigma_\beta^2\Sigma^{-1})$ where $\alpha = \alpha^{(i)}, \sigma_\eta^2 = \sigma_\eta^{2(i)}, \sigma_\varepsilon^2 = \sigma_\varepsilon^{2(i)}$ and set $\beta^{(i)} = \beta^*$.
7. Repeat steps 3-6 for i in 1, 2, . . . , M.

Inference on β

- After throwing out a number of initial samples from the Gibb's sampler we can estimate β by taking the mean of the posterior samples.
- We create a 95 credibility interval (as a pseudo-confidence interval) by calculating the 2.5^{th} and 97.5^{th} percentiles of the posterior draws.

Simulation Analyses

- Conducted two separate simulation analyses.
- The most desirable model is one that,
  - Maintains 95% coverage of true parameter.
  - Is unbiased.
  - Has small parameter variance (small 95% confidence intervals)

Simulation Study 1

We sampled from the models,

$$\begin{aligned} y_t &= b_0 + X_t\beta + e_t, \quad b_0 \sim N(0, \sigma_b^2) \\ e_t &= \rho e_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2) \end{aligned} \tag{1}$$

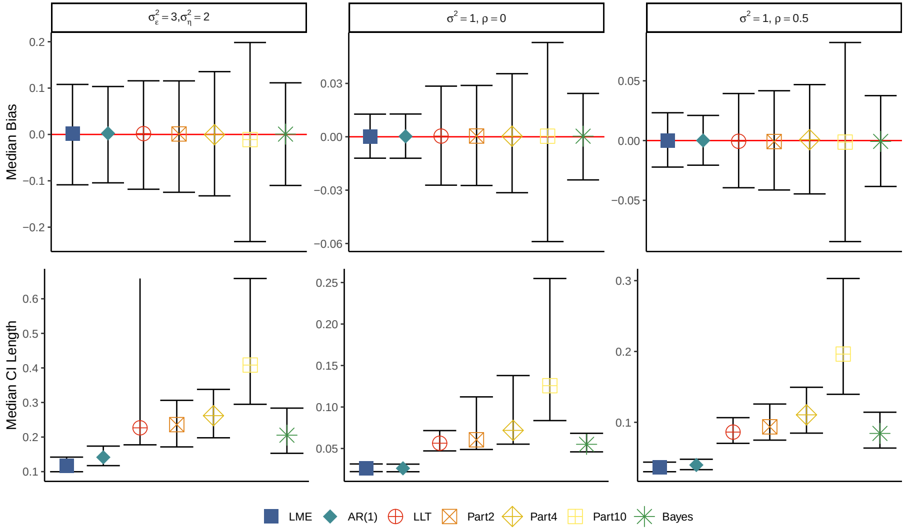
$$\begin{aligned} y_t &= \alpha_t + X_t\beta + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2 I_n) \\ \alpha_t &= \alpha_{t-1} + \eta_t, \qquad \eta_t \sim N(0, \sigma_\eta^2 I_n) \end{aligned} \tag{2}$$

We simulated 100 subjects to have between 3-10 observations. X was simulated to mirror our initial model of interest from the NACC. The variables σ_ε^2, σ_η^2, and ρ varied between simulations. We compared 95% CI coverage, bias, and estimate variance between 1. LMEM with a random intercept, 2. LMEM with a random intercept and AR(1) error correlation structure, the liklihood state space model, the Bayesian state space model, then a state space model partitioned into 2, 4, and 10 groups.

95% Coverage

Variance Parameters		Traditional Methods		State Space Methods				
σ_ε^2	σ_η^2	LME	AR(1)	LLT	Part2	Part4	Part10	Bayes
$\sigma^1 = 1$	$\rho = 0$	0.954	0.953	0.947	0.952	0.963	0.983	0.964
$\sigma^1 = 1$	$\rho = 0.1$	0.938	0.947	0.945	0.953	0.963	0.981	0.955
$\sigma^1 = 1$	$\rho = 0.5$	0.889	0.944	0.965	0.967	0.977	0.987	0.962
$\sigma_\varepsilon^2 = 3$	$\sigma_\eta^2 = 0$	0.941	0.940	0.944	0.965	0.969	0.978	0.958
$\sigma_\varepsilon^2 = 3$	$\sigma_\eta^2 = 1$	0.790	0.847	0.947	0.948	0.956	0.972	0.940
$\sigma_\varepsilon^2 = 3$	$\sigma_\eta^2 = 2$	0.736	0.836	0.948	0.944	0.950	0.970	0.932
$\sigma_\varepsilon^2 = 3$	$\sigma_\eta^2 = 3$	0.715	0.838	0.946	0.940	0.945	0.969	0.928
$\sigma_\varepsilon^2 = 30$	$\sigma_\eta^2 = 10$	0.781	0.840	0.947	0.941	0.945	0.970	0.946
$\sigma_\varepsilon^2 = 60$	$\sigma_\eta^2 = 20$	0.780	0.840	0.946	0.944	0.945	0.967	0.943

Bias and CI length



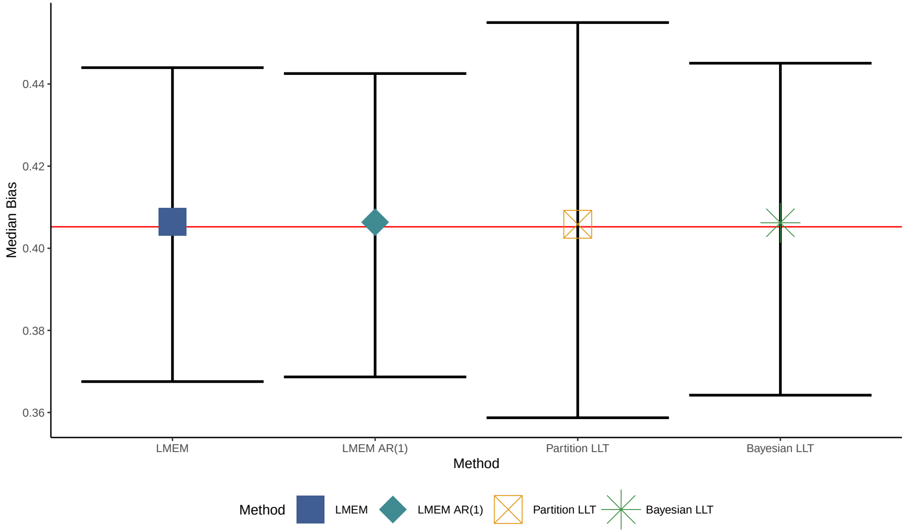
Real Data Simulation

- Add a linear effect on the Animals outcome for half the subjects.
- Estimate the model:

$$\begin{aligned}\textbf{Updated Animals} \sim &(1 + \mathrm{I}\{\text{Transitioned to MCI or Dementia}\} + \text{APOE} \\ &+ \text{Sex} + \text{APOE*Sex} + \text{Race} + \text{Age} + \text{Education} \\ &+ \textbf{Randomized Group}) * \text{Time}\end{aligned}$$

- Estimate the linear effect using the different models,
- LMEM, LMEM AR(1), Partitioned LLT with group size 100, and Bayesian LLT.

Bias



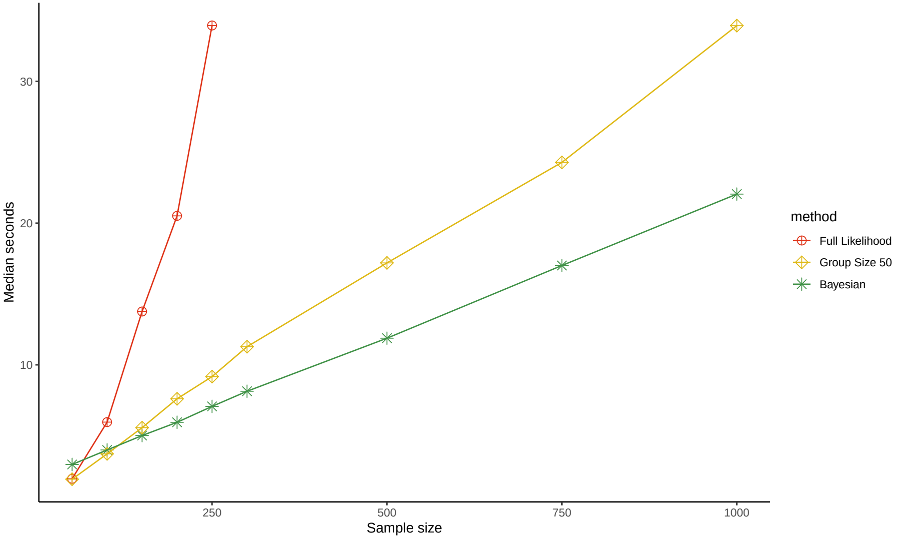
Median Bias
0.44
0.42
0.40
0.38
0.36
LMEM
LMEM AR(1)
Partition LLT
Bayesian LLT
Method
Method
LMEM
LMEM AR(1)
Partition LLT
Bayesian LLT

Coverage

LMEM	LMEM AR(1)	Partition LLT	Bayesian LLT
0.795	0.889	0.935	0.94

- When the underlying data generation process is unknown, the LLT models do a much better estimating the effect of interest.

Computation Time



Analysis

On the NACC data set we fit the model:

$$\text{Animals} \sim (1 + \text{I}\{\text{Transitioned to MCI or Dementia}\} + \text{APOE} + \text{Sex} \\ + \text{APOE*Sex} + \text{Race} + \text{Age} + \text{Education}) * \text{Time}$$

Using the LMEM, LMEM AR(1), and the Bayesian LLT Model.

Results

	APOE	APOE x Sex
LMEM	-0.143 (-0.229, -0.058)	-0.023 (-0.128, 0.082)
LMEM AR(1)	-0.135 (-0.244, -0.025)	-0.049 (-0.182, 0.085)
Partition LLT	-0.18 (-0.36, 0)	-0.047 (-0.261, 0.167)
Bayesian LLT	-0.136 (-0.269, 0.011)	-0.054 (-0.23, 0.122)

Summary

- The LLT shows proper 95% coverage for the fully simulated, even under model misspecification, and for the real NACC data.
- When compared to the full data LLT, the partitioned LLT shows very similar results as long as the number of parameters estimated is reasonable for the group size.
- The Bayesian LLT is the most desirable of the fitted models as it maintains 95% coverage, is unbiased, and has the smallest parameter variance.

Future Projects

- Joint modeling of longitudinal outcomes using SSM framework.
- Cluster analysis of longitudinal outcomes using SSM framework.

Joint Modeling

$$\begin{bmatrix} y_{1ti} \\ y_{2ti} \\ y_{3ti} \end{bmatrix} = \begin{bmatrix} \mu_{1ti} \\ \mu_{2ti} \\ \mu_{3ti} \end{bmatrix} + x_{ti}\beta \begin{bmatrix} 1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_{1ti} \\ \varepsilon_{2ti} \\ \varepsilon_{3ti} \end{bmatrix}, \quad \varepsilon \sim N(0, \sigma_\varepsilon^2)$$

$$\begin{bmatrix} \mu_{1ti} \\ \mu_{2ti} \\ \mu_{3ti} \end{bmatrix} = \begin{bmatrix} \mu_{1(t-1)i} \\ \mu_{2(t-1)i} \\ \mu_{3(t-1)i} \end{bmatrix} + \begin{bmatrix} \eta_{1ti} \\ \eta_{2ti} \\ \eta_{3ti} \end{bmatrix}, \quad \eta_{.ti} \sim N(0, \begin{bmatrix} \sigma_{\eta 11} & \sigma_{\eta 12} & \sigma_{\eta 13} \\ \sigma_{\eta 12} & \sigma_{\eta 22} & \sigma_{\eta 23} \\ \sigma_{\eta 13} & \sigma_{\eta 23} & \sigma_{\eta 33} \end{bmatrix})$$

Cluster Analysis

$$\begin{bmatrix} y_{1ti} \\ y_{2ti} \\ y_{3ti} \\ y_{4ti} \end{bmatrix} = \begin{bmatrix} \mu_{1ti} \\ \mu_{1ti} \\ \mu_{2ti} \\ \mu_{2ti} \end{bmatrix} + x_{it}\beta \begin{bmatrix} 1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{bmatrix} + \begin{bmatrix} \varepsilon_{1ti} \\ \varepsilon_{2ti} \\ \varepsilon_{3ti} \\ \varepsilon_{4ti} \end{bmatrix}, \quad \varepsilon \sim N(0, \sigma_\varepsilon^2)$$

$$\begin{bmatrix} \mu_{1ti} \\ \mu_{2ti} \end{bmatrix} = \begin{bmatrix} \mu_{1(t-1)i} \\ \mu_{2(t-1)i} \end{bmatrix} + \begin{bmatrix} \eta_{1ti} \\ \eta_{2ti} \end{bmatrix}, \quad \eta_{.ti} \sim N(0, \begin{bmatrix} \sigma_{\eta 11} & \sigma_{\eta 12} \\ \sigma_{\eta 12} & \sigma_{\eta 22} \end{bmatrix})$$

Projected Timeline

- Project 1 is under co-author review for submission.
- Project 2 to be completed during fall 2021.
- Project 3 to be completed during spring 2022.
- Defense in the summer of 2022.

Thank you!

- Recommendations?
- Questions?