

ARTICLE

Using R and lme/lmer to fit different two- and three-level longitudinal models

POSTED BY KRISTOFFER MAGNUSSON ON 2015-04-21 17:30:00+02:00 IN R
([HTTPS://RPSYCHOLOGIST.COM/CATEGORY/R](https://rpsychologist.com/category/r))

TAGGED AS  STATISTICS ([HTTPS://RPSYCHOLOGIST.COM/TAG/STATISTICS](https://rpsychologist.com/tag/statistics))  LONGITUDINAL
([HTTPS://RPSYCHOLOGIST.COM/TAG/LONGITUDINAL](https://rpsychologist.com/tag/longitudinal))  MULTILEVEL ([HTTPS://RPSYCHOLOGIST.COM/TAG/MULTILEVEL](https://rpsychologist.com/tag/multilevel))
 LINEAR MIXED-EFFECTS MODELS ([HTTPS://RPSYCHOLOGIST.COM/TAG/LINEAR-MIXED-EFFECTS-MODELS](https://rpsychologist.com/tag/linear-mixed-effects-models))  GROWTH
CURVE ([HTTPS://RPSYCHOLOGIST.COM/TAG/GROWTH-CURVE](https://rpsychologist.com/tag/growth-curve))  LME4 ([HTTPS://RPSYCHOLOGIST.COM/TAG/LME4](https://rpsychologist.com/tag/lme4)) 
NLME ([HTTPS://RPSYCHOLOGIST.COM/TAG/NLME](https://rpsychologist.com/tag/nlme))

Contents

- Data format
- Power analysis, and simulating these models
- Longitudinal two-level model
 - Unconditional model
 - Unconditional growth model
 - Conditional growth model
 - Conditional growth model: dropping random slope
 - Conditional growth model: dropping random intercept
 - Conditional growth model: dropping intercept-slope covariance
- Three-level models
 - Conditional three-level growth model
 - Subject level randomization (therapist crossed effect)
 - Different level 3 variance-covariance matrix
 - Partially nested models
- More on level 1 specification
 - Heteroscedasticity at Level 1
 - More grouping
 - First-order Autoregressive AR(1) residuals
 - Heterogenous AR(1)
 - More level 1 variance-covariances matrices
 - Changing the functional form of time
 - Quadratic trend
 - Orthogonal polynomials
 - Piecewise growth curve
 - Coding scheme 1: separate slopes
 - Coding scheme 2: incremental/decremental slope
 - Drop the correlation between time piece 1 and 2
 - Adding a quadratic effect
- Hypothesis tests
 - Wald test
 - Likelihood ratio test
 - Profile confidence intervals
 - Parametric bootstrap
 - Kenward-Roger degrees of freedom approximation
 - Shattertwaite degrees of freedom approximation
- Book recommendations
- Suggestions, errors or typos

Share on:

🐦 (<https://twitter.com/intent/tweet?text=Using R and lme/lmer to fit different two- and three-level longitudinal models&url=https%3A//rpsychologist.com%2Fr-guide-longitudinal-lme-lmer&via=krstoffr>)

f (<http://www.facebook.com/sharer.php?u=https://rpsychologist.com/r-guide-longitudinal-lme-lmer&t=Using R and lme/lmer to fit different two- and three-level longitudinal models>)

g+ (<https://plus.google.com/share?url=https%3A//rpsychologist.com%2Fr-guide-longitudinal-lme-lmer>)

I often get asked how to fit different multilevel models (or individual growth models, hierarchical linear models or linear mixed-models, etc.) in R. In this guide I have compiled some of the more common and/or useful models (at least common in clinical psychology), and how to fit them using `nlme::lme()` and `lme4::lmer()`. I will cover the common two-level random intercept-slope model, and three-level models when subjects are clustered due to some higher level grouping (such as therapists), partially nested models where there are clustering in one group but not the other, and different level 1 residual covariances (such as AR(1)). The point of this post is to show how to fit these longitudinal models in R, not to cover the statistical theory behind them, or how to interpret them.

Data format

In all examples I assume this data structure.

| subjects | tx | therapist | time | y |
|----------|----|-----------|------|----|
| 1 | 0 | 1 | 0 | 10 |
| 1 | 0 | 1 | 1 | 12 |
| 1 | 0 | 1 | 2 | 14 |
| 2 | 0 | 1 | 0 | 4 |
| 2 | 0 | 1 | 1 | 14 |
| 2 | 0 | 1 | 2 | 13 |
| 3 | 0 | 2 | 0 | 12 |
| 3 | 0 | 2 | 1 | 15 |
| 3 | 0 | 2 | 2 | 16 |

| subjects | tx | therapist | time | y |
|----------|-----|-----------|------|-----|
| 4 | 0 | 2 | 0 | 17 |
| 4 | 0 | 2 | 1 | 13 |
| 4 | 0 | 2 | 2 | 12 |
| 5 | 0 | 3 | 0 | 15 |
| 5 | 0 | 3 | 1 | 13 |
| ... | ... | ... | ... | ... |

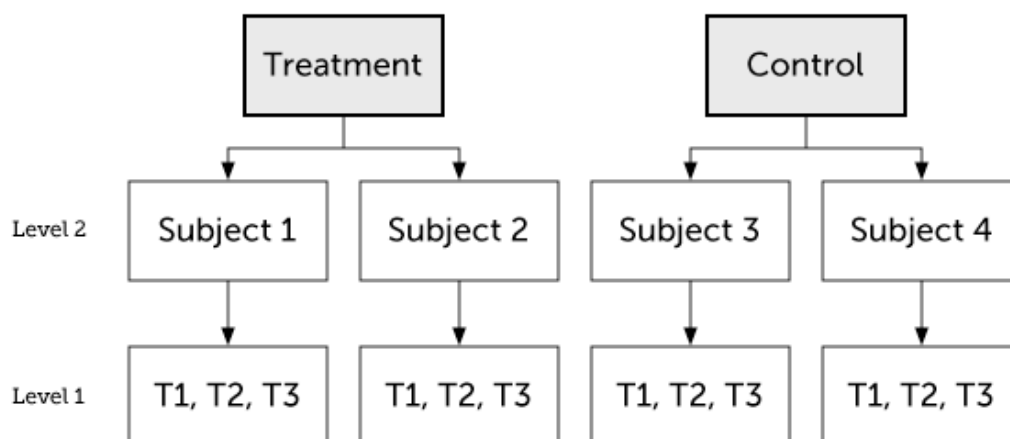
Where `subjects` is each subject's id, `tx` represent treatment allocation and is coded 0 or 1, `therapist` is the refers to either clustering due to therapists, or for instance a participant's group in group therapies. `y` is the outcome variable.

Power analysis, and simulating these models

Most of the designs covered in this post are supported by my R package `powerlmm`, (<http://cran.r-project.org/package=powerlmm>) (<http://cran.r-project.org/package=powerlmm>). It can be used to calculate power for these models, or to simulate them to investigate model misspecification. I will soon integrate the package into this post, in order to create example data sets. For now, see the package's vignettes for tutorials.

Longitudinal two-level model

We will begin with the two-level model, where we have repeated measures on individuals in different treatment groups.



Unconditional model

Model formulation

Level 1

$$Y_{ij} = \beta_{0j} + R_{ij}$$

Level 2

$$\beta_{0j} = \gamma_{00} + U_{0j}$$

with,

$$U_{0j} \sim \mathcal{N}(0, \tau_{00}^2),$$

and

$$R_{ij} \sim \mathcal{N}(0, \sigma^2)$$

To fit this model we run

```
# lme4
lmer(y ~ 1 + (1 | subjects), data=data)

# nlme
lme(y ~ 1, random = ~ 1 | subjects, data=data)
```

Unconditional growth model

Model formulation

Level 1

$$Y_{ij} = \beta_{0j} + \beta_{1j}t_{ij} + R_{ij}$$

Level 2

$$\beta_{0j} = \gamma_{00} + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + U_{1j}$$

with,

$$\begin{pmatrix} U_{0j} \\ U_{1j} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00}^2 & \tau_{01} \\ \tau_{01} & \tau_{10}^2 \end{pmatrix} \right),$$

and

$$R_{ij} \sim \mathcal{N}(0, \sigma^2)$$

To fit this model we run

```
# lme4
lmer(y ~ time + (time | subjects), data=data)

# nlme
lme(y ~ time, random = ~ time | subjects, data=data)
```

Conditional growth model

Model formulation

Level 1

$$Y_{ij} = \beta_{0j} + \beta_{1j}t_{ij} + R_{ij}$$

Level 2

$$\beta_{0j} = \gamma_{00} + \gamma_{01}TX_j + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}TX_j + U_{1j}$$

with,

$$\begin{pmatrix} U_{0j} \\ U_{1j} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00}^2 & \tau_{01} \\ \tau_{01} & \tau_{10}^2 \end{pmatrix} \right),$$

and

$$R_{ij} \sim \mathcal{N}(0, \sigma^2)$$

To fit this model we run

```
# lme4
lmer(y ~ time * tx + (time | subjects), data=data)

# nlme
lme(y ~ time * tx, random = ~ time | subjects, data=data)
```

Conditional growth model: dropping random slope

Model formulation

Level 1

$$Y_{ij} = \beta_{0j} + \beta_{1j}t_{ij} + R_{ij}$$

Level 2

$$\beta_{0j} = \gamma_{00} + \gamma_{01}TX_j + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}TX_j$$

with,

$$U_{0j} \sim \mathcal{N}(0, \tau_{00}^2)$$

and

$$R_{ij} \sim \mathcal{N}(0, \sigma^2)$$

To fit this model we run

```
# lme4
lmer(y ~ time * tx + (1 | subjects), data=data)

# nlme
lme(y ~ time * tx, random = ~ 1 | subjects, data=data)
```

Conditional growth model: dropping random intercept

Model formulation

Level 1

$$Y_{ij} = \beta_{0j} + \beta_{1j}t + R_{ij}$$

Level 2

$$\beta_{0j} = \gamma_{00} + \gamma_{01}TX_j$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}TX_j + U_{1j}$$

with,

$$U_{0j} \sim \mathcal{N}(0, \tau_{10}^2)$$

and

$$R_{ij} \sim \mathcal{N}(0, \sigma^2)$$

To fit this model we run

```
# lme4
lmer(y ~ time * tx + ( 0 + time | subjects), data=data)

# nlme
lme(y ~ time * tx, random = ~ 0 + time | subjects, data=data)
```

Conditional growth model: dropping intercept-slope covariance

Model formulation

Level 1

$$Y_{ij} = \beta_{0j} + \beta_{1j}t + R_{ij}$$

Level 2

$$\beta_{0j} = \gamma_{00} + \gamma_{01}T X_j + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}T X_j + U_{1j}$$

with,

$$\begin{pmatrix} U_{0j} \\ U_{1j} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00}^2 & 0 \\ 0 & \tau_{10}^2 \end{pmatrix} \right),$$

and

$$R_{ij} \sim \mathcal{N}(0, \sigma^2)$$

To fit this model we run

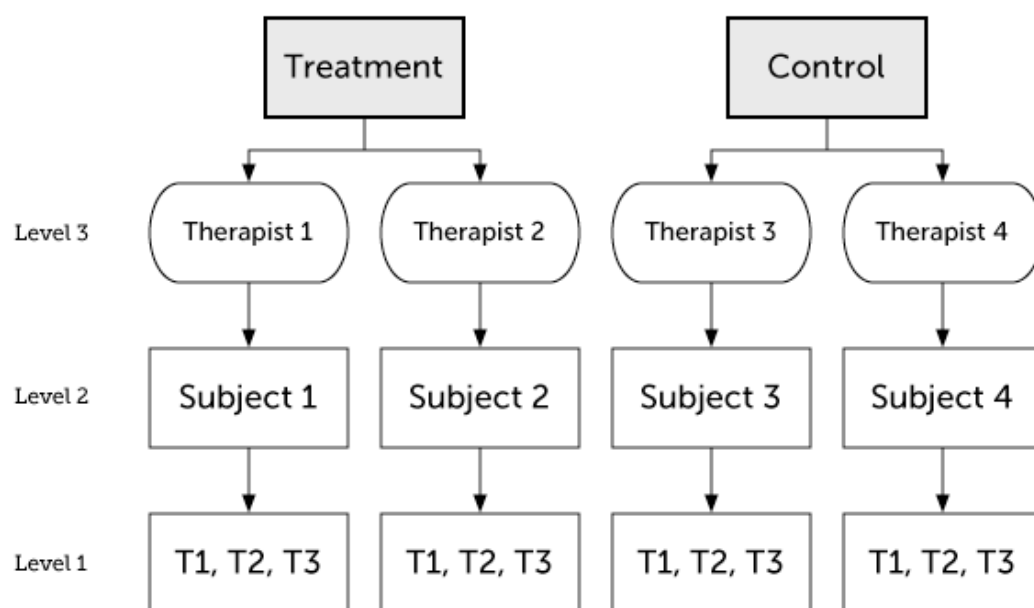
```
# lme4
lmer(y ~ time * tx +
      (time || subjects) , data=data)

# same as above
lmer(y ~ time * tx +
      (1 | subjects) +
      (0 + time | subjects) ,
      data=data)

# nlme
lme(y ~ time * tx,
     random = list(subjects = pdDiag(~time)),
     data=data)
```

Three-level models

Here I will cover some different three-level models. In my examples clustering at the highest level is due to therapists. But the examples generalize to other forms of clustering as well, such as group therapy or clustering due to health-care provider.



Conditional three-level growth model

We will jump straight to the conditional three-level growth model, with the following model formulation:

Level 1

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk}t_{ijk} + R_{ijk}$$

Level 2

$$\beta_{0jk} = \gamma_{00k} + U_{0jk}$$

$$\beta_{1jk} = \gamma_{10k} + U_{1jk}$$

Level 3

$$\gamma_{00k} = \delta_{000} + \delta_{001}TX_k + V_{0k}$$

$$\gamma_{10k} = \delta_{100} + \delta_{101}TX_k + V_{1k}$$

with,

$$\begin{pmatrix} U_{0j} \\ U_{1j} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00}^2 & \tau_{01} \\ \tau_{01} & \tau_{10}^2 \end{pmatrix} \right),$$

and,

$$\begin{pmatrix} V_{0k} \\ V_{1k} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \varphi_{00}^2 & \varphi_{01} \\ \varphi_{01} & \varphi_{10}^2 \end{pmatrix} \right),$$

and

$$R_{ijk} \sim \mathcal{N}(0, \sigma^2)$$

To fit this model we use `therapist/subjects`, which specifies nesting. This formula expands to a main effect of therapist and a interaction between therapist and subjects (which is the subject level effect).

```
# lme
lmer(y ~ time * tx +
      (time | therapist/subjects),
      data=df)

## expands to
lmer(y ~ time * tx +
      (time | therapist:subjects) +
      (time | therapist),
      data=df)

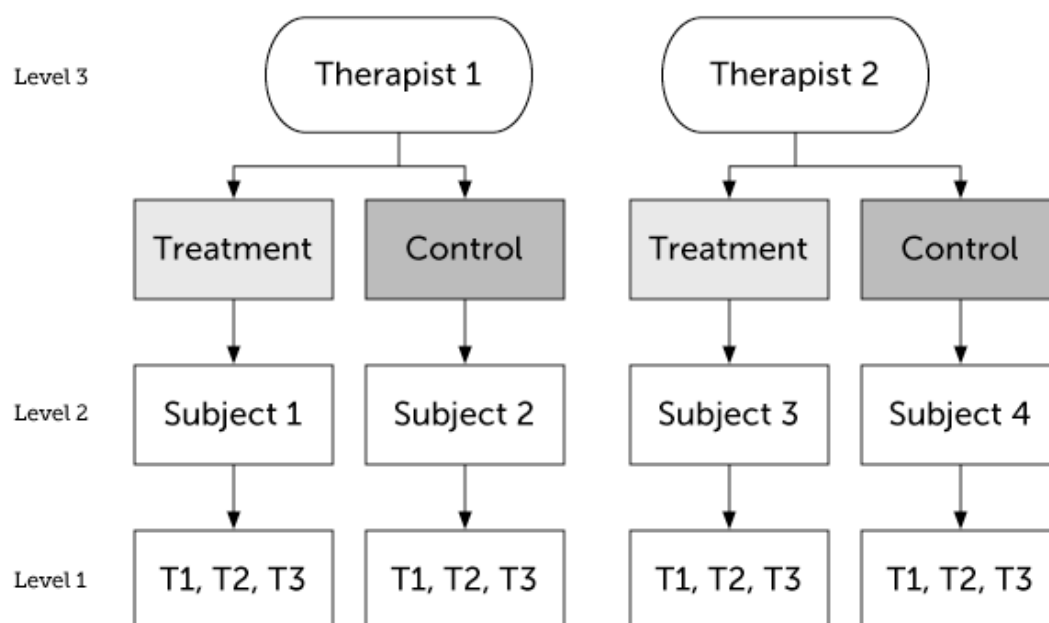
# nlme
lme(y ~ time * tx,
     random = ~time | therapist/subjects,
     data=df)

## expands to
lme(y ~ time * tx,
     random = list(therapist = ~time,
                   subjects = ~time),
     data=df)
```

Subject level randomization (therapist crossed effect)

In the previous example therapists only provided one type of treatment (nested design).

Sometimes therapists will be a crossed effect, i.e. in a parallel group design they will deliver both treatments. If it's a randomized trial then in this design we have subject level randomization, whereas in the previous example randomization was at the therapist level.



Level 1

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk}t_{ijk} + R_{ijk}$$

Level 2

$$\beta_{0jk} = \gamma_{00k} + \gamma_{01k}TX_{jk} + U_{0jk}$$

$$\beta_{1jk} = \gamma_{10k} + \gamma_{11k}TX_{jk} + U_{1jk}$$

Level 3

$$\gamma_{00k} = \delta_{000} + V_{0k}$$

$$\gamma_{10k} = \delta_{100} + V_{1k}$$

$$\gamma_{01k} = \delta_{010} + V_{2k}$$

$$\gamma_{11k} = \delta_{110} + V_{3k}$$

with,

$$\begin{pmatrix} U_{0j} \\ U_{1j} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00}^2 & \tau_{01} \\ \tau_{01} & \tau_{10}^2 \end{pmatrix} \right),$$

and,

$$\begin{pmatrix} V_{0k} \\ V_{1k} \\ V_{2k} \\ V_{3k} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \varphi_{00}^2 & 0 & 0 & 0 \\ 0 & \varphi_{10}^2 & 0 & 0 \\ 0 & 0 & \varphi_{20}^2 & 0 \\ 0 & 0 & 0 & \varphi_{30}^2 \end{pmatrix} \right),$$

and

$$R_{ijk} \sim \mathcal{N}(0, \sigma^2)$$

In this model we estimate no covariances at level 3. However, at the therapist level we have random effects for *time*, *treatment* and *time * treatment*. I fit this saturated model because you can easily delete a random effect in the expanded `lmer` syntax below.

```

# lme4
lmer(y ~ time * tx +
      (time | therapist:subjects) +
      (time * tx || therapist),
      data=df)

## expands to
lmer(y ~ time*tx +
      (time | subjects:therapist) +
      (1 | therapist) +
      (0 + tx | therapist) +
      (0 + time | therapist) +
      (0 + time:tx | therapist), data=df)

# nlme
lme(y ~ time * tx,
     random = list(therapist = pdDiag(~time * tx),
                   subjects = ~time),
     data=df)

```

Different level 3 variance-covariance matrix

We might hypothesize that therapists that are allocated participants that report worse symptoms at treatment start have better outcomes (more room for improvement). To allow for separate covariances in each treatment group we update the variance-covariance matrix at level 3

$$\begin{pmatrix} V_{0k} \\ V_{1k} \\ V_{2k} \\ V_{3k} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 & \varphi_{00}^2 & \varphi_{01} & 0 & 0 \\ 0 & \varphi_{01} & \varphi_{10}^2 & 0 & 0 \\ 0 & 0 & 0 & \varphi_{20}^2 & \varphi_{23} \\ 0 & 0 & 0 & \varphi_{23} & \varphi_{30}^2 \end{pmatrix}, \right)$$

To fit this model we run

```
# lme4
lmer(y ~ time * tx + (time | therapist:subjects) +
      (time | therapist) +
      (0 + tx + time:tx | therapist),
      data=data)

# nlme
lme(y ~time * tx,
    random = list(therapist = pdBlocked(list(~time, ~0 + tx + time:tx)),
                  subjects = ~time ),
    data=data)
```

Of course, we could also estimate all six covariances at level 3. For instance, we could look at if therapists who are more successful with Treatment A are also more successful with Treatment B, i.e. $cov(V_{2k}, V_{3k}) = \varphi_{23}$, and so forth. The full unstructured level 3 variance-covariance matrix we will estimate is thus

$$\begin{pmatrix} V_{0k} \\ V_{1k} \\ V_{2k} \\ V_{3k} \end{pmatrix} \sim \mathcal{N} \begin{pmatrix} 0 & \varphi_{00}^2 & \varphi_{01} & \varphi_{02} & \varphi_{03} \\ 0 & \varphi_{01} & \varphi_{10}^2 & \varphi_{12} & \varphi_{13} \\ 0 & \varphi_{02} & \varphi_{12} & \varphi_{20}^2 & \varphi_{23} \\ 0 & \varphi_{03} & \varphi_{13} & \varphi_{23} & \varphi_{30}^2 \end{pmatrix}$$

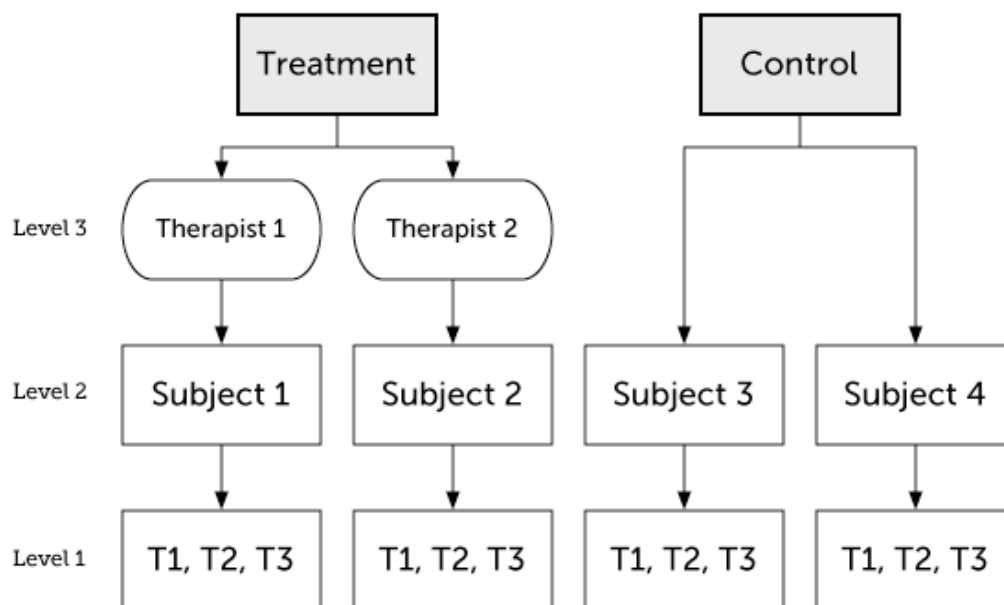
Which we fit by running

```
# lme4
lmer(y ~ time * tx +
      (time | therapist:subjects) +
      (time * tx | therapist),
      data=df)

# nlme
lme(y ~ time * tx,
    random = list(therapist = ~time * tx,
                  subjects = ~time),
    data=df)
```

Partially nested models

Partially nesting occurs when we have nesting in one group but not the other. For instance, we might compare a treatment group to a wait-list condition. Subjects in the wait-list will not be nested, but subjects in treatment group will be nested within therapists.



We can write this model like this

Level 1

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk}t_{ijk} + R_{ijk}$$

Level 2

$$\beta_{0jk} = \gamma_{00} + \gamma_{01k}TX_{jk} + U_{0jk}$$

$$\beta_{1jk} = \gamma_{10} + \gamma_{11k}TX_{jk} + U_{1jk}$$

Level 3

$$\gamma_{01k} = \delta_{010} + V_{0k}$$

$$\gamma_{11k} = \delta_{110} + V_{1k}$$

with,

$$\begin{pmatrix} U_{0j} \\ U_{1j} \end{pmatrix} \sim \mathcal{N} \begin{pmatrix} 0 & \tau_{00}^2 & 0 \\ 0 & 0 & \tau_{10}^2 \end{pmatrix},$$

and,

$$\begin{pmatrix} V_{0k} \\ V_{1k} \end{pmatrix} \sim \mathcal{N} \begin{pmatrix} 0 & \varphi_{00}^2 & 0 \\ 0 & 0 & \varphi_{10}^2 \end{pmatrix},$$

and

$$R_{ijk} \sim \mathcal{N}(0, \sigma^2)$$

```
# lme4
lmer(y ~ time * tx +
      (1 | therapist:subjects) +
      (0 + time | therapist:subjects) +
      (0 + time:tx | therapist) +
      (0 + tx | therapist),
      data=data)

# nlme
lme(y ~ time * tx,
     random = list(therapist = pdDiag(~ 0 + time:tx + tx),
                   subjects = pdDiag(~time)),
     data=data)
```

More on level 1 specification

Heteroscedasticity at Level 1

Only `lme` allows modeling heteroscedastic residual variance at level 1. If we wanted to extend our two level model and allow for different level 1 residual variance in the treatment groups, we'd get

$$\begin{aligned}(R_{ij} \mid \text{TX} = 0) &\sim \mathcal{N}(0, \sigma_0^2) \\ (R_{ij} \mid \text{TX} = 1) &\sim \mathcal{N}(0, \sigma_1^2)\end{aligned}$$

If we wanted to extend our two-level model with this level 1 structure we'd run

```
lme(y ~ time * tx,
     random = ~ time | subjects,
     weights = varIdent(form= ~ 1 | tx), data=data)
```

More grouping

We could also add another grouping factor such as time, and fit a model with heteroscedastic level 1 residuals for each time point in each treatment group. For instance, for $i = 0, 1, 2$ we get

$$(R_{ij} \mid \text{TX} = 0) \sim \mathcal{N} \left(\begin{pmatrix} 0 & \sigma_{00}^2 & 0 & 0 \\ 0 & 0 & \sigma_{01}^2 & 0 \\ 0 & 0 & 0 & \sigma_{02}^2 \end{pmatrix} \right)$$

$$(R_{ij} \mid \text{TX} = 1) \sim \mathcal{N} \begin{pmatrix} 0 & \sigma_{10}^2 & 0 & 0 \\ 0 & , & 0 & \sigma_{11}^2 & 0 \\ 0 & 0 & 0 & \sigma_{12}^2 \end{pmatrix}$$

which we'd fit by running

```
lme(y ~ time * tx,
    random = ~ time | subjects,
    weights = varIdent(form= ~ 1 | tx * time),
    data=data)
```

First-order Autoregressive AR(1) residuals

For $T = 1, 2, 3, \dots, N_1$ time points we get the level 1 variance-covariance matrix

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{T-1} \\ \rho & 1 & \rho & \dots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \dots & 1 \end{pmatrix}$$

we leads to

$$R_{ij} \sim \mathcal{N}(0, \Sigma)$$

To fit this level 1 residual structure we use the `correlation` argument.

```
lme(y ~ time * tx,
    random = ~ time | subjects,
    correlation = corAR1(),
    data=data)
```

Heterogenous AR(1)

We can also extend the level 1 variance-covariance matrix from above, to allow for different residuals at each time point.

$$\Sigma = \begin{pmatrix} \sigma_0^2 & \sigma_0\sigma_1\rho & \sigma_0\sigma_2\rho^2 & \cdots & \sigma_0\sigma_i\rho^{T-1} \\ \sigma_1\sigma_0\rho & \sigma_1^2 & \sigma_1\sigma_2\rho & \cdots & \sigma_1\sigma_i\rho^{T-2} \\ \sigma_2\sigma_0\rho^2 & \sigma_2\sigma_1\rho & \sigma_2^2 & \cdots & \sigma_2\sigma_i\rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_T\sigma_0\rho^{T-1} & \sigma_T\sigma_1\rho^{T-2} & \sigma_T\sigma_2\rho^{T-2} & \cdots & \sigma_T^2 \end{pmatrix}$$

and we have that

$$R_{ij} \sim \mathcal{N}(0, \Sigma)$$

To fit this level 1 model we simply use both the `correlation` and the `weights` argument.

```
lme(y ~ time * tx,
    random = ~ time | subjects,
    weights = varIdent(form= ~ 1 | time),
    correlation = corAR1(),
    data=data)
```

More level 1 variance-covariances matrices

See `?corClasses` for the different types of residual variance-covariances matrices `lme` can estimate.

Changing the functional form of time

All of the examples above assume linear change. Here I will cover some examples of how to model nonlinear change at level 1. The examples will be based on the two-level model, but you could easily be combined them with the three-level models outlined above.

Quadratic trend

Level 1

$$Y_{ij} = \beta_{0j} + \beta_{1j}t_{1ij} + \beta_{2j}t_{1ij}^2 + R_{ij}$$

Level 2

$$\beta_{0j} = \gamma_{00} + \gamma_{01}TX_j + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}TX_j + U_{1j}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}TX_j + U_{2j}$$

with,

$$\begin{pmatrix} U_{0j} \\ U_{1j} \\ U_{2j} \end{pmatrix} \sim \mathcal{N} \begin{pmatrix} 0 & \tau_{00}^2 & \tau_{01} & \tau_{02} \\ 0 & \tau_{01} & \tau_{10}^2 & \tau_{12} \\ 0 & \tau_{02} & \tau_{12} & \tau_{20}^2 \end{pmatrix},$$

and

$$R_{ij} \sim \mathcal{N}(0, \sigma^2)$$

```
# lme4
lmer(y ~ (time + I(time^2)) * tx +
      (time + I(time^2) | subjects),
      data=data)

# or use poly()
lmer(y ~ poly(time, 2, raw = TRUE) * tx +
      (poly(time, 2, raw = TRUE) | subjects),
      data=data)

# nlme
lme(y ~ (time + I(time^2)) * tx,
     random = ~time + I(time^2) | subjects,
     data=data)
```

Orthogonal polynomials

If you'd like to fit orthogonal polynomials you can use the `poly()` function with `raw = FALSE` (which is the default).

```
# lme4
lmer(y ~ poly(time, 2) * tx +
      (poly(time, 2) | subjects),
      data=data)

# nlme
lme(y ~ poly(time, 2) * tx,
     random = ~poly(time, 2) | subjects,
     data=data)
```

Piecewise growth curve

Segmenting the time trend into different pieces has got more to do with simple dummy coding of regression variables, than any specifics of `lme` or `lmer`. However, I will cover some common scenarios anyway.

To fit a piecewise growth model we simply replace `time` with two dummy variables `time1` and `time2`, that represent the different time periods. A common scenario is that the first piece represents the acute treatment phase, and piece 2 represent the follow-up phase.

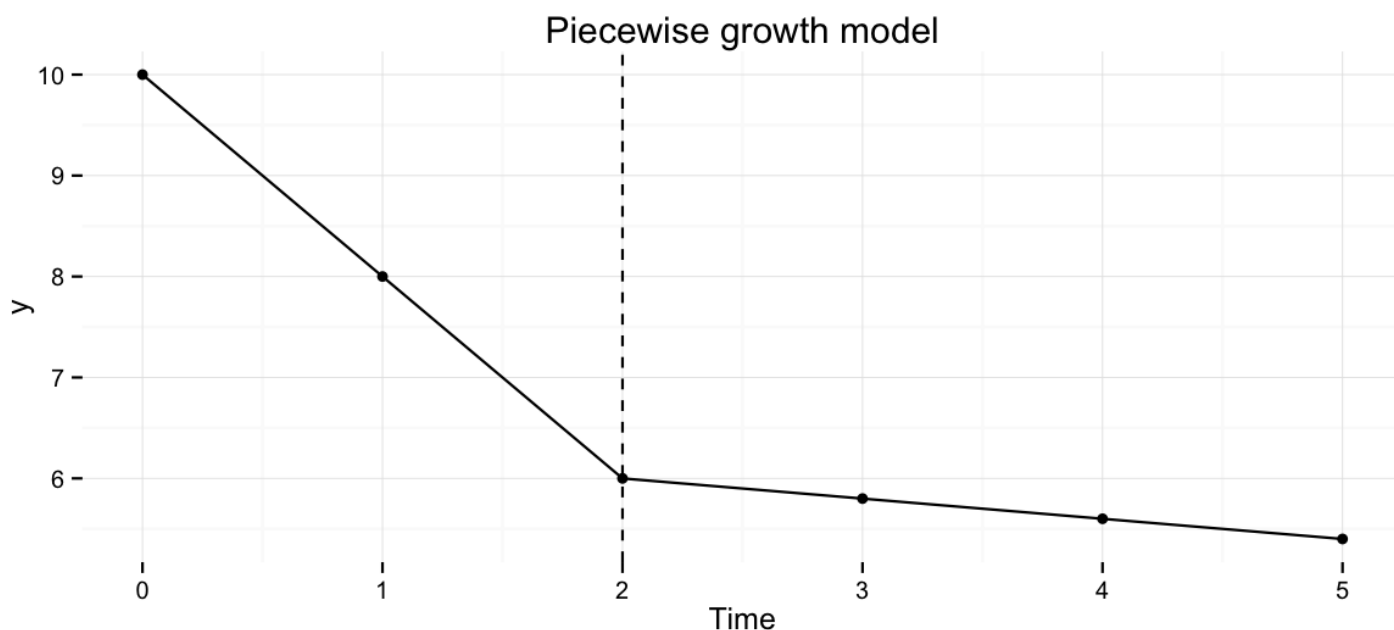
Coding scheme 1: separate slopes

| Time | 0 | 1 | 2 | 3 | 4 | 5 |
|--------|---|---|---|---|---|---|
| Time 1 | 0 | 1 | 2 | 2 | 2 | 2 |
| Time 2 | 0 | 0 | 0 | 1 | 2 | 3 |

Coding scheme 2: incremental/decremental slope

| Time | 0 | 1 | 2 | 3 | 4 | 5 |
|--------|---|---|---|---|---|---|
| Time 1 | 0 | 1 | 2 | 3 | 4 | 5 |
| Time 2 | 0 | 0 | 0 | 1 | 2 | 3 |

These two coding schemes only differ in the interpretation of the regression coefficients. In scheme 1 the two slope coefficients represent the actual slope in the respective time period. Whereas in scheme 2 the coefficient for time 2 represents the deviation from the slope in period 1, i.e. if the estimate is 0 then the rate of change is the same in both periods.



We could specify this model like this

Level 1

$$Y_{ij} = \beta_{0j} + \beta_{1j}t_{1ij} + \beta_{2j}t_{2ij} + R_{ij}$$

Level 2

$$\beta_{0j} = \gamma_{00} + \gamma_{01}TX_j + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}TX_j + U_{1j}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}TX_j + U_{2j}$$

with,

$$\begin{pmatrix} U_{0j} \\ U_{1j} \\ U_{2j} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 & \tau_{00}^2 & \tau_{01} & \tau_{02} \\ 0 & \tau_{01} & \tau_{10}^2 & \tau_{12} \\ 0 & \tau_{02} & \tau_{12} & \tau_{20}^2 \end{pmatrix} \right),$$

and

$$R_{ij} \sim \mathcal{N}(0, \sigma^2)$$

In this model I've fit the full level 2 variance-covariance matrix. If we wanted to fit this model we'd do it like this

```

J <- 20 # J participants
data$time1 <- rep(0:5, J)
data$time2 <- rep(c(0,0,0,1,2,3), J)

# lme4
lmer(y ~ (time1 + time2) * tx +
      (time1 + time2 | subjects),
      data=data)

# nlme
lme(y ~ (time1 + time2) * tx,
     random = ~time1 + time2 | subjects,
     data=data)

```

Drop the correlation between time piece 1 and 2

Sometimes you might want to fit a model with a correlation between the random intercept and time piece 1, but no correlation between time piece 2 and the other effects. This would change the level 2 variance-covariance from above to this

$$\begin{pmatrix} U_{0j} \\ U_{1j} \\ U_{2j} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 & \tau_{00}^2 & \tau_{01} & 0 \\ 0 & \tau_{01} & \tau_{10}^2 & 0 \\ 0 & 0 & 0 & \tau_{20}^2 \end{pmatrix} \right)$$

Fitting this model is straight-forward in `lmer` and more complicated in `lme`.

```

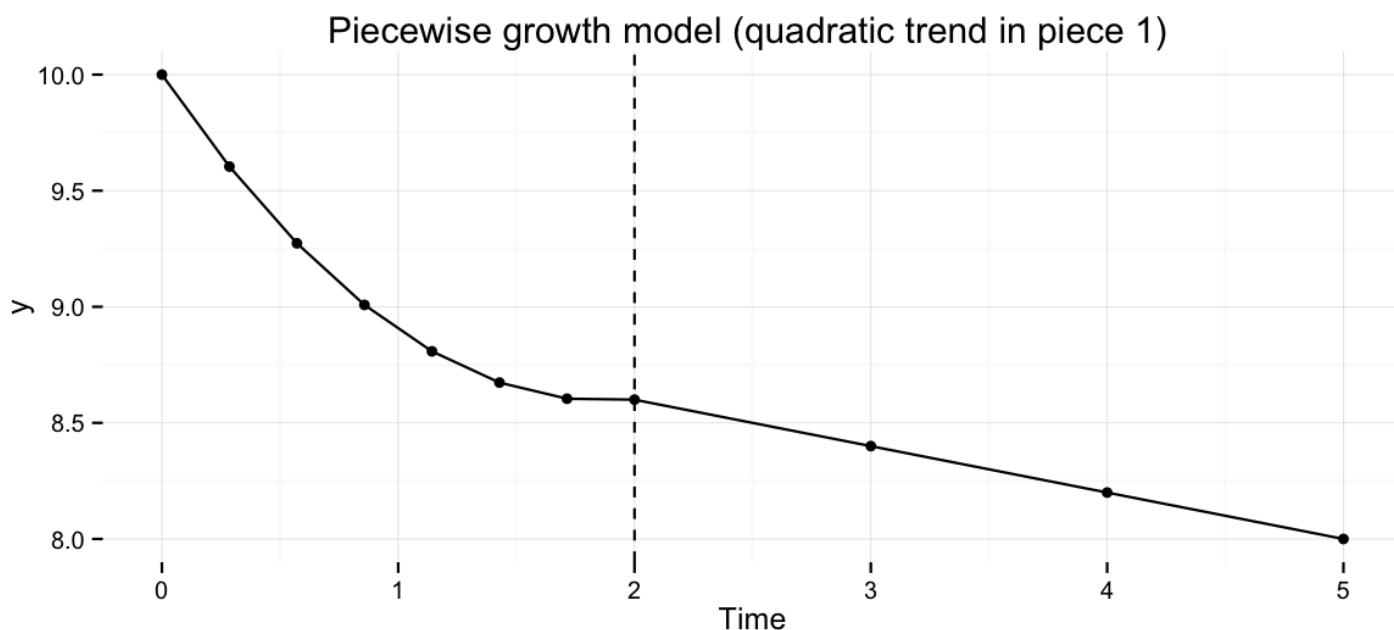
# lme4
lmer(y ~ (time1 + time2) * tx +
      (time1 | subjects) +
      (0 + time2 | subjects),
      df)

# nlme
lme(y ~ (time1 + time2) * tx,
     random = list(subjects = pdBlocked(list(~time1, ~0 + time2))),
     data=df)

```

Adding a quadratic effect

We could extend the two-part piecewise growth model to allow for non-linear change during one or both of the pieces. As an example, I'll cover extending the model to allow for quadratic change during piece 1.



We could write this model like this

Level 1

$$Y_{ij} = \beta_{0j} + \beta_{1j}t_1 + \beta_{2j}t_{1ij}^2 + \beta_{3j}t_{2ij} + R_{ij}$$

Level 2

$$\beta_{0j} = \gamma_{00} + \gamma_{01}TX_j + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}TX_j + U_{1j}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}TX_j + U_{2j}$$

$$\beta_{3j} = \gamma_{30} + \gamma_{31}TX_j + U_{3j}$$

with,

$$\begin{pmatrix} U_{0j} \\ U_{1j} \\ U_{2j} \\ U_{3j} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 & \tau_{00}^2 & \tau_{01} & \tau_{02} & \tau_{03} \\ 0 & \tau_{01} & \tau_{10}^2 & \tau_{12} & \tau_{13} \\ 0 & \tau_{02} & \tau_{12} & \tau_{20}^2 & \tau_{23} \\ 0 & \tau_{03} & \tau_{13} & \tau_{23} & \tau_{30}^2 \end{pmatrix} \right),$$

and

$$R_{ij} \sim \mathcal{N}(0, \sigma^2)$$

This model could be fit like this

```
# lme4
lmer(y ~ (time1 + I(time1^2) + time2) * tx +
      (time1 + I(time1^2) + time2 | subjects),
      df)

# nlme
lme(y ~ (time1 + I(time1^2) + time2) * tx,
    random = ~time1 + I(time1^2) + time2 | subjects,
    data=df)
```

If you wanted to fit a reduced random effects structure you could use the method outlined in *"Drop the correlation between time piece 1 and 2"*.

Hypothesis tests

`lmer` does not report p -values or degrees of freedoms, see `?pvalues` and `r-sig-mixed-models` FAQ (<http://glmm.wikidot.com/faq>) for why not. However, there are other packages that will calculate p -values for you. I will cover some of them here.

Wald test

```
summary(lme.mod)
```

Likelihood ratio test

```
fm1 <- lmer(y ~ 1 + (1 | subjects), data=data)
fm2 <- lmer(y ~ 1 + (time | subjects), data=data)
# also works with lme objects
anova(fm1, fm2)
```

Profile confidence intervals

```
confint(lmer.mod)
```

Parametric bootstrap

```
confint(lmer.mod, method="boot", nsim=1000)
```

Kenward-Roger degrees of freedom approximation

```
library(lmerTest)
anova(lmer.mod, ddf = "Kenward-Roger")
```

Shattertwaite degrees of freedom approximation

```
library(lmerTest)
anova(lmer.mod)
#or
summary(lmer.mod)
```

Book recommendations

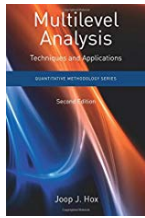
Not all of these books are specific to R and longitudinal data analysis. However, I've found them all useful over the years when working with multilevel/linear mixed models.

- Singer & Willett (2003). Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence (http://www.amazon.com/gp/product/0195152964/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0195152964&linkCode=as2&tag=rpsyc-20&linkId=53JJ3JAUSFZ64VHP)
- Pinheiro & Bates (2009). Mixed-Effects Models in S and S-PLUS (Statistics and Computing) (http://www.amazon.com/gp/product/1441903178/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=1441903178&linkCode=as2&tag=rpsyc-20&linkId=2Z4KDF5IE73GMZP5)
- Hedeker & Gibsson (2006). Longitudinal Data Analysis (http://www.amazon.com/gp/product/0471420271/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0471420271&linkCode=as2&tag=rpsyc-20&linkId=RPVTFDG4UHVXX5FY)
- Fitzmaurice, Laird, & Ware (2011). Applied Longitudinal Analysis (http://www.amazon.com/gp/product/0470380276/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0470380276&linkCode=as2&tag=rpsyc-20&linkId=CSBEDQ3ABDZ3COSP)
- Diggle et al (2013). Analysis of Longitudinal Data (Oxford Statistical Science Series) (http://www.amazon.com/gp/product/0199676755/ref=as_li_tl)

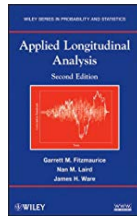
ie=UTF8&camp=1789&creative=9325&creativeASIN=0199676755&linkCode=as2&tag=rpsyc-20&linkId=WITAB5KBDIZPYOBE)

- Gelman & Hill (2006). Data Analysis Using Regression and Multilevel/Hierarchical Models (http://www.amazon.com/gp/product/052168689X/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=052168689X&linkCode=as2&tag=rpsyc-20&linkId=W3EUL34BHUQSWV45)
- Raudenbush & Bryk (2001). Hierarchical Linear Models: Applications and Data Analysis Methods (http://www.amazon.com/gp/product/076191904X/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=076191904X&linkCode=as2&tag=rpsyc-20&linkId=R7H6FE6VV4LTDSBX)
- Snijders & Bosker (2011). Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling (http://www.amazon.com/gp/product/184920201X/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=184920201X&linkCode=as2&tag=rpsyc-20&linkId=JST5JW5FZ2DTUZZD)
- Hox, J (2010). Multilevel Analysis: Techniques and Applications, Second Edition (http://www.amazon.com/gp/product/1848728468/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=1848728468&linkCode=as2&tag=rpsyc-20&linkId=PRHBSG3QY2LPID4V)
- Goldstein, H (2010). Multilevel Statistical Models (http://www.amazon.com/gp/product/0470748656/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0470748656&linkCode=as2&tag=rpsyc-20&linkId=J3IUJFIMN7VNYH72)
- Mirman (2014). Growth Curve Analysis and Visualization Using R (http://www.amazon.com/gp/product/1466584327/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=1466584327&linkCode=as2&tag=rpsyc-20&linkId=UQ6VHSCCLULBTZVO)
- Finch, Bolin, & Kelley (2014). Multilevel Modeling Using R (http://www.amazon.com/gp/product/1466515856/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=1466515856&linkCode=as2&tag=rpsyc-20&linkId=4Q3A3CMXA5DZCSMJ)

Shop Related Products



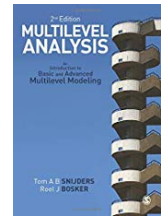
~~£29.99~~ ~~£14.99~~
 Multilevel Analysis: Techniques
 and Applications (2) (https://aax-eu.amazon-



~~£91.99~~ ~~£44.99~~
 Applied Longitudinal Analysis
 (Wiley Series in Probability and Statistics) (2) (https://aax-eu.amazon-



~~£37.45~~ ~~£20.99~~
 Data Analysis Using Regression
 and Multilevel/Hierarchical Models (5) (https://aax-eu.amazon-



~~£29.99~~ ~~£14.99~~
 Multilevel Analysis: An
 Introduction To Basic An... (6) (https://aax-eu.amazon-

Ads by Amazon (https://aax-eu.amazon-adsystem.com/x/c/QgXROQEcV49yCXZ3aDqpdKQAAAFwAPLGGAMAAAE0ATrZ8pQ/https://affiliate-program.amazon.co.uk/home/ads/ref=sm_n_au_dka_GB_logo?adId=logo&creativeASIN=logo&linkId=7f89fd2c0404b8abb13b12aeff53999c&tag=rpsych-21&linkCode=w41&ref-refURL=https%3A%2F%2Frpsychologist.com%2Fguide-longitudinal-lme-lmer&slotNum=0&imprToken=62Ny-oi4M6mXyuhTVqFMVw&adType=smart&adMode=auto&adFormat=grid&impressionTimestamp=1580563876871&ac-ms-src=nsa-ads&cid=nsa-ads) (https://aax-eu.amazon-

adsystem.com/x/c/QgXROQEcV49yCXZ3aDqpdKQAAAFwAPLGGAMAAAE0ATrZ8pQ/https://www.amazon.co.uk/adprefs/ref=sm_n_au_dka_GB_ac?tag=rpsych-21&linkCode=w41)

Suggestions, errors or typos

Please don't hesitate to contact me (<http://rpsychologist.com/about>) if you find some errors in this guide.

Share on:

🐦 (<https://twitter.com/intent/tweet?text=Using R and lme/lmer to fit different two- and three-level longitudinal models&url=https%3A//rpsychologist.com%2Fguide-longitudinal-lme-lmer&via=krstoffr>)

f (<http://www.facebook.com/sharer.php?u=https://rpsychologist.com/r-guide-longitudinal-lme-lmer&t=Using R and lme/lmer to fit different two- and three-level longitudinal models>)

g+ (<https://plus.google.com/share?url=https%3A//rpsychologist.com%2Fguide-longitudinal-lme-lmer>)



About Kristoffer Magnusson

I'm a PhD-student and a clinical psychologist from Sweden with a passion for research and statistics. This is my personal blog about psychological research and statistical programming with R.

Donate

The content on this blog is shared for free under a CC-BY license. If you like my work and want to support it you can:



Buy me a coffee (<https://www.buymeacoffee.com/krstoffr>)

(or use PayPal

<https://www.paypal.me/krstoffr>)

If you like my interactive visualizations you can also support them:

[Become a patron](#)
[⦿ \(http://twitter.com/krstoffer\)](http://twitter.com/krstoffer)
[in \(https://www.linkedin.com/in/kristoffer/\)](https://www.linkedin.com/in/kristoffer/)
[⦿ \(https://github.com/rpsychologist\)](https://github.com/rpsychologist)
[RSS](#)

 POSTED IN [R \(https://rpsychologist.com/category/r/\)](https://rpsychologist.com/category/r/)

 TAGGED AS [STATISTICS \(https://rpsychologist.com/tag/statistics\)](https://rpsychologist.com/tag/statistics) [LONGITUDINAL \(https://rpsychologist.com/tag/longitudinal\)](https://rpsychologist.com/tag/longitudinal) [MULTILEVEL \(https://rpsychologist.com/tag/multilevel\)](https://rpsychologist.com/tag/multilevel) [LINEAR MIXED-EFFECTS MODELS \(https://rpsychologist.com/tag/linear-mixed-effects-models\)](https://rpsychologist.com/tag/linear-mixed-effects-models) [GROWTH CURVE \(https://rpsychologist.com/tag/growth-curve\)](https://rpsychologist.com/tag/growth-curve) [LME4 \(https://rpsychologist.com/tag/lme4\)](https://rpsychologist.com/tag/lme4) [NLME \(https://rpsychologist.com/tag/nlme\)](https://rpsychologist.com/tag/nlme)

30 Comments

R Psychologist blog

[1 Login](#)
[Recommend](#) 13

[Tweet](#)
[Share](#)
[Sort by Best](#)


Join the discussion...

LOG IN WITH

OR SIGN UP WITH DISQUS [?](#)

Name


Vincent Isoz • 7 months ago

What book or PDF would you recommend to have very extremely mathematical details on longitudinal analysis (all steps of proofs given). Whatever the level, basics or advanced? Thx

[^](#) | [v](#) • [Reply](#) • [Share](#)

Keith Goldfeld • 7 months ago

Kristoffer -

This is a really useful post, which I know you did quite a while ago. I was wondering if you have come across a situation where you needed to impose structure on the correlation of the random effects (of the 2nd level of a 2-level model). In my case, I have measurement periods (and individuals are measured only once. There is a cluster-specific effect in each period (U_t , t in $(0,1,2,3,4)$), and these cluster specific effects are correlated; in particular, the correlations degrade over time (AR1). I don't believe lme4 or nlme can accomodate this, but thought maybe you would correct me if I am wrong.

- Keith

[^](#) | [v](#) • [Reply](#) • [Share](#)

Shalini Kurumathur • 10 months ago

Thank you for such a great website. I understood the three level equations because of this website. I am working on a educational data: three-level discontinuous growth model. I have an issue in the code using control variables. ses(Socio economic status), lcsz(class size) are time variant control variables and demo(demographic such as urban/rural-as 0,1) is time-invariant control variable. After the null model, I used this code. Not sure if this is right. PERF85 is the dependent variable, TIME, TRANS and RECOV are independent variable. I am not getting any error though-

```
step1a <-lme(PERF85~TIME+TRANS+RECOV+ses+lcsizes, random=~1|ENTITY_CD,FINNL85,
control=lmeControl(opt="optim"))
while using (TIME|ses) or (1|ses) I am getting error
```

^ | v • Reply • Share ›



Sketching Sketcher • a year ago

This is very useful! THanks!

^ | v • Reply • Share ›



Marcos Salvino • a year ago

Dear Kristoffer,

Thank you for this website. I tried to follow your commands (in R), but I didn't get your data. Please, can you show how to get your data? I'm trying to understand your thoughts. Thank you in advance, Marcos Salvino

^ | v • Reply • Share ›



radek • a year ago

Hi Kristoffer. How would go about fitting quadratic trend for data without tx? Should that simply be removed from equation?

```
lmer(y ~ (time + I(time^2)) +
(time + I(time^2) | subjects),
data=data)
```

Also, how can one access the results for plotting temporal effect on graphs as you do it?

Many thanks!

^ | v • Reply • Share ›



Mirjam Moerbeek • a year ago

Dear Kristoffer,

Thank you for this great website. I have a question on the partially nested models. What value should you use for the grouping variable "therapist" for those subjects who are in the control condition (i.e. for those subjects who are not nested within therapists). I have seen various suggestions: the control subjects are all in clusters of size 1, the control subjects are all in one large cluster, use the value "none" for the control subjects. What would you recommend?

Thanks in advance,

Mirjam Moerbeek

^ | v • Reply • Share ›



Kristoffer Magnusson Mod → Mirjam Moerbeek • a year ago

Hi Mirjam, that's a great question and something I've been planning on clarifying in this post. As long as the indicator variable (control = 0) is included in the random effects formula all the third-level random effects will be zero in the control group, for all the coding suggestions you mention. Using this parameterization, it is my experience that the choice of coding will make no practical difference on the estimates (at least when using lme4). Using the value "none" (character) will be equivalent to coding it as one large cluster. For practical reasons, I tend to code them as being in 1 large cluster, or sometimes (if it is a simulation) as belonging to the same number of clusters as the treatment group. The model will fit slightly faster compared to

using singletons, and ranet() and coeft() will only include 1 row third level for the control group.

For clarification, here's a small example using powerlmm

```
library(powerlmm)
library(lmerTest)
library(dplyr)

## Partially nested DGP
## random intercept and slope at both lvl 2 and 3
p <- study_parameters(n1 = 11,
                     n2 = 50,
                     n3 = 4,
```

[see more](#)

^ | v • Reply • Share ›



Mirjam Moerbeek → Kristoffer Magnusson • a year ago

Dear Kristoffer,

Thanks for the prompt reply! I have simulated some large data sets for different specifications of the therapist variable and the results are conform your answer.

Best wishes,

Mirjam

^ | v • Reply • Share ›



Johanna TJ • a year ago

Hi Kristoffer,

I have another question regarding centering. I was wondering whether I need to additionally center time (e.g. grand mean centering) if I use {0,1,2,2,2} and {0,0,0,1,2} as the two time periods? Or does this automatically mean that time was centered at the third time point so that this time point served as zero?

J.

^ | v • Reply • Share ›



Johanna TJ • 2 years ago

Hi Kristoffer, I am currently working on a 2-level piecewise growth model with two separate slopes and I was using your coding scheme 1 which was really helpful! Could you provide a reference in which this coding scheme was introduced?

Thanks,

J.

^ | v • Reply • Share ›



Kristoffer Magnusson Mod → Johanna TJ • 2 years ago

Hi Johanna, the coding scheme is from Raudenbush & Bryk (2001) p 179.

^ | v • Reply • Share ›



Johanna TJ → Kristoffer Magnusson • 2 years ago

Great, thanks a lot!

^ | v • Reply • Share ›



Slobodan Ivanovic • 2 years ago

Hi Kristoffer, excellent post. Do you know how to implement multiple membership with lme4?

Greetings

^ | v • Reply • Share ›



Kristoffer Magnusson Mod → Slobodan Ivanovic • 2 years ago

Thanks! Last time I checked fitting a multiple membership model in lme4 required some hacking, however that was several years ago. brms or MCMCglmm are probably better options, see `brms::mm` or `MCMCglmm::mult.memb`

1 ^ | v • Reply • Share ›



Slobodan Ivanovic → Kristoffer Magnusson • 2 years ago

Thank you!

^ | v • Reply • Share ›



saddas • 2 years ago

Hi Kristoffer,

I am wondering if R can handle a 3-level multilevel model where the DV is a count variable (i.e., a GLMM model). I know that this is not possible in SAS in Proc GLIMMIX.

Thanks.

Shamel

^ | v • Reply • Share ›



Kristoffer Magnusson Mod → saddas • 2 years ago

Hi, you can use lme4 to fit 3-level GLMMs just replace "lmer" with "glmer", and add the "family" argument.

^ | v • Reply • Share ›



saddas → Kristoffer Magnusson • 2 years ago

Thanks, Kristoffer. As far as I know, glmer (lme4 package) does not allow us to specify a structure for the residual (R) matrix. If this is correct, is there any way to model 3 levels of nesting using NLME, which has flexibility in specifying the R matrix?

^ | v • Reply • Share ›



Kristoffer Magnusson Mod → saddas • 2 years ago

That is correct, lme4 does not support R-side effects. You could try MASS::glmPQL which allows fitting GLMMs with R-side effects using nlme under the hood.

^ | v • Reply • Share ›



saddas → Kristoffer Magnusson • 2 years ago

Thanks, 2 quick follow up questions:

(1) When you say lme4 does not support R-side effects, does this mean it can only model a residual matrix with constant variances and zero covariances (i.e., a variance components matrix)?

(2) Does glmPQL allow for fitting models with 3 levels of nesting? ?I couldn't find anything in the documentation on this.

^ | v • Reply • Share ›



Vijay • 2 years ago



Hi Kristoffer,

First of all thank you for this guide.

This is very useful guide.

I have one question regarding Two-level model.

Actually i have data with two group of treatment having repeated measurements. Total sample size is 50, 25 each group and repeated measurements taken 14 times each individual. I want to apply two level model on this data. I want to check which treatment is better. i want to keep repeated measurements of an individual at level one. Could you please help me out for this issue.

Thanks in advance.

Best,

Vijay

^ | v • Reply • Share ›



JayZ • 2 years ago

Hi Kristoffer,

I was wondering about the piecewise fitting of time variable... I have a dataset where the time points are 1:60 days and every 2 or 3 days(so not 1,2,3 etc but 1,3,5,8 etc), and unequal between subjects. I tried to use the second format as you present it here, and then fit the model, but i get a rank-deficient error in lmer, and a singularity error in lme... Do you have an idea of solving that ?

Thanks a lot,

John

^ | v • Reply • Share ›



Pablo Beltramone • 2 years ago

Hi Kristoffer

Thanks a lot for this guide. I really really need your help with an issue: I need to fit a model with time series data, like a mixed model. I have the hourly arrivals of customers that enter to a company. I'm trying to fit a mixed model with FE such the day of the week and the hour of the day. As ME I'm trying to estimate an ar(1) covar matrix for within day (5 hours from open to close time, so 5x5 matrix) and a ar(1) covar matrix for between days (5 days a week, so 5x5 matrix), so my data is:

date day hour Customers

1/1/2014 1 10 125

1/1/2014 1 11 110

1/1/2014 1 12 180

1/1/2014 1 13 173

1/1/2014 1 14 68

2/1/2014 2 10 114

2/1/2014 2 11 92

...

My model in R is actually

[see more](#)

^ | v • Reply • Share ›



Sarah Landmann • 4 years ago

Hi Kristoffer,

I am so glad that I found your guide; it is extremely helpful to have a tutorial for fitting multilevel models for a two-group repeated measurements design! Thanks a lot!

I have only two questions and would be grateful if you could help me with that:

If I get it right, the conditional growth model includes an interaction term `group*time` and a random part allowing subjects to have individual intercepts and individual slopes for the effect of time on this dependent variable, right?

In some tutorials I have seen that they add a "1+" (`1+time|subjects`) instead of only (`time|subjects`). What is the difference here?

My second question is: Against which model would I have to test this model In a likelihood ratio test in order to test that the interaction "`group*time`" is significant (which is actually what I am interested in, compared to the effect of either group or time)? Should I compare it against a model that only includes an additive term for group and time as predictors? Or against the "intercept-only model" `lmer(y ~ 1 + (1 | subjects))`?

Thanks a lot!

Sarah

^ | v • Reply • Share ›



BERNARSLAN UZUNDAG • 4 years ago

Hi Kristoffer, I found this page very useful. I'm trying to fit a piecewise growth curve model to my data using lme. I coded the time variable as you suggested (option 2). But still I'm not sure how to interpret the findings. If the slope of time2 is very close to 0 but not equal to 0, does it still mean that the slope in time2 is not different from the slope in time1? Is there a test to compare the slopes? Thank you.

^ | v • Reply • Share ›



Krishna Karthik • 5 years ago

Hi, First of all, thank you very much for posting this guide. I cannot express how much this has helped in me in my work! Second, towards the very end of the post, you gave out several suggestions for hypothesis testing, and mentioned that in the 'lme4' package, performing a 'summary' of the LMER model would give p-values of Wald tests. Or did you mean the test statistic only?

Also, can you point me to a source that explains in a simplified manner why the lmer model does not give p-values? Douglas Bates' book explains it, but it is more math heavy and I could barely follow it. Thanks again!!

^ | v • Reply • Share ›



Kristoffer Magnusson Mod ➔ Krishna Karthik • 5 years ago

Glad you found this post helpful! To get p-values by using `summary()` on lme4-objects you need to install and load the `lmerTest`-package. Once its loaded you can use `lmer()` and `summary()` as you normally do.

Regarding why p-values are not reported there's an explanation in in the r-sig-mixed-models FAQ <http://glmm.wikidot.com/faq...>

^ | v • Reply • Share ›



Krishna Karthik • 5 years ago • edited

Thank you so much for this!! I cannot express how much easier you have made my life by posting these. However, I have one question. The `summary` command on the lmer model does not seem to have Wald tests in the output, any idea why? Or how I can get them?

^ | v • Reply • Share ›



Gabbie • a year ago

Hi Kristoffer

This article was very timely as I am in the process of analyzing the data for my first PhD study. Could you explain further what you mean by "In scheme 1 the two slope coefficients represent the actual slope in the respective time period. Whereas in scheme 2 the coefficient for time 2 represents the deviation from the slope in period 1, i.e. if the estimate is 0 then the rate of change is the same in both periods."

I understand in scheme 1, you compare slopes created by time points 0:2 to 3:5. Scheme 2 is a little confusing. Are you saying that here you comparing a hypothetical slope 3:5 (based on what would be expected given slope 0:2) to 3:5?

I was also wondering whether you could describe how to code three growth curves (i.e. my data has two discontinuities) using both schemes you described above. I have data with 15 time points (0:14). At time points 5:7 there was an experimental manipulation which created a downward shift from time point 4 to 5, and then an upward shift from point 7 to 8.

Most of the studies I've looked at for examples only have a single discontinuity or do not adequately describe how they coded their variables.

About Site

Designed and built by Kristoffer Magnusson. Powered by Pelican (<https://getpelican.com>), which takes great advantage of Python (<https://www.python.org>), and Bootstrap (<https://getbootstrap.com>).

License

Except where otherwise noted, content on this site is licensed under a Creative Commons Attribution 4.0 International license. (<https://creativecommons.org/licenses/by/4.0/>)



Categories

D3.js (<https://rpsychologist.com/category/d3js>)

Psychology (<https://rpsychologist.com/category/psychology>)

R (<https://rpsychologist.com/category/r>)

Connect

🐦 Twitter (<https://twitter.com/krstoffer>)

🌐 LinkedIn (<https://www.linkedin.com/in/kristofferm>)

🔗 Github (<https://github.com/rpsychologist>)

📡 RSS (<https://feeds.feedburner.com/RPsychologist>)

Links

R-Bloggers (<https://r-bloggers.com>)