

# Mixed models in R using the lme4 package

## Part 3: Longitudinal data

Douglas Bates

University of Wisconsin - Madison  
and R Development Core Team  
<Douglas.Bates@R-project.org>

UseR!2009, Rennes, France  
July 7, 2009

Longitudinal data: sleepstudy

A model with random effects for intercept and slope

Conditional means

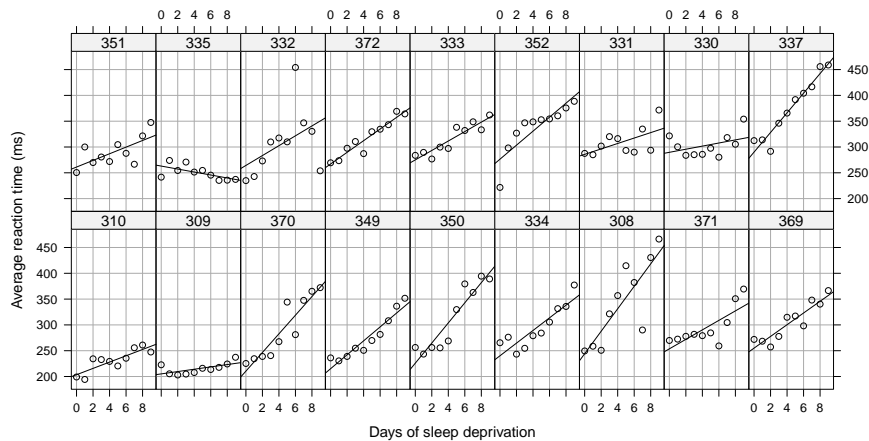
## Simple longitudinal data

- ▶ *Repeated measures* data consist of measurements of a response (and, perhaps, some covariates) on several *experimental* (or observational) *units*.
- ▶ Frequently the experimental (observational) unit is *Subject* and we will refer to these units as “subjects”. However, the methods described here are not restricted to data on human subjects.
- ▶ *Longitudinal* data are repeated measures data in which the observations are taken over time.
- ▶ We wish to characterize the response over time within subjects and the variation in the time trends between subjects.
- ▶ Frequently we are not as interested in comparing the particular subjects in the study as much as we are interested in modeling the variability in the population from which the subjects were chosen.

## Sleep deprivation data

- ▶ This laboratory experiment measured the effect of sleep deprivation on cognitive performance.
- ▶ There were 18 subjects, chosen from the population of interest (long-distance truck drivers), in the 10 day trial. These subjects were restricted to 3 hours sleep per night during the trial.
- ▶ On each day of the trial each subject’s reaction time was measured. The reaction time shown here is the average of several measurements.
- ▶ These data are *balanced* in that each subject is measured the same number of times and on the same occasions.

## Reaction time versus days by subject



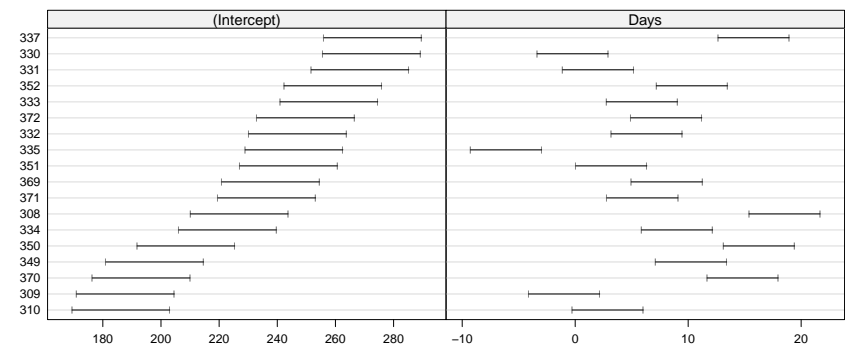
## Assessing the linear fits

- ▶ In most cases a simple linear regression provides an adequate fit to the within-subject data.
- ▶ Patterns for some subjects (e.g. 350, 352 and 371) deviate from linearity but the deviations are neither widespread nor consistent in form.
- ▶ There is considerable variation in the intercept (estimated reaction time without sleep deprivation) across subjects – 200 ms. up to 300 ms. – and in the slope (increase in reaction time per day of sleep deprivation) – 0 ms./day up to 20 ms./day.
- ▶ We can examine this variation further by plotting confidence intervals for these intercepts and slopes. Because we use a pooled variance estimate and have balanced data, the intervals have identical widths.
- ▶ We again order the subjects by increasing intercept so we can check for relationships between slopes and intercepts.

## Comments on the sleep data plot

- ▶ The plot is a “trellis” or “lattice” plot where the data for each subject are presented in a separate panel. The axes are consistent across panels so we may compare patterns across subjects.
- ▶ A reference line fit by simple linear regression to the panel’s data has been added to each panel.
- ▶ The aspect ratio of the panels has been adjusted so that a typical reference line lies about  $45^\circ$  on the page. We have the greatest sensitivity in checking for differences in slopes when the lines are near  $\pm 45^\circ$  on the page.
- ▶ The panels have been ordered not by subject number (which is essentially a random order) but according to increasing intercept for the simple linear regression. If the slopes and the intercepts are highly correlated we should see a pattern across the panels in the slopes.

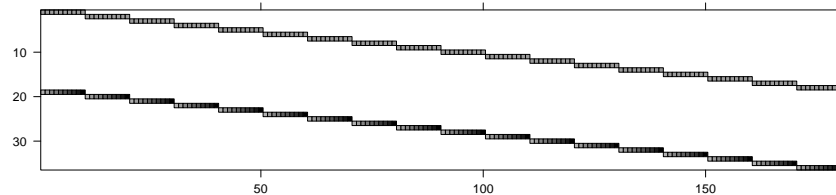
## 95% conf int on within-subject intercept and slope



These intervals reinforce our earlier impressions of considerable variability between subjects in both intercept and slope but little evidence of a relationship between intercept and slope.

## A preliminary mixed-effects model

- ▶ We begin with a linear mixed model in which the fixed effects  $[\beta_1, \beta_2]'$  are the representative intercept and slope for the population and the random effects  $\mathbf{b}_i = [b_{i1}, b_{i2}]'$ ,  $i = 1, \dots, 18$  are the deviations in intercept and slope associated with subject  $i$ .
- ▶ The random effects vector,  $\mathbf{b}$ , consists of the 18 intercept effects followed by the 18 slope effects.



## Terms and matrices

- ▶ The term `Days` in the formula generates a model matrix  $\mathbf{X}$  with two columns, the intercept column and the numeric `Days` column. (The intercept is included unless suppressed.)
- ▶ The term `(Days|Subject)` generates a vector-valued random effect (intercept and slope) for each of the 18 levels of the `Subject` factor.

## Fitting the model

```
> (fm1 <- lmer(Reaction ~ Days + (Days | Subject),  
+   sleepstudy))
```

```
Linear mixed model fit by REML  
Formula: Reaction ~ Days + (Days | Subject)  
Data: sleepstudy  
AIC BIC logLik deviance REMLdev  
1756 1775 -871.8 1752 1744  
Random effects:  
Groups Name Variance Std.Dev. Corr  
Subject (Intercept) 612.092 24.7405  
Days 35.072 5.9221 0.066  
Residual 654.941 25.5918  
Number of obs: 180, groups: Subject, 18  
Fixed effects:  
Estimate Std. Error t value  
(Intercept) 251.405 6.825 36.84  
Days 10.467 1.546 6.77  
  
Correlation of Fixed Effects:  
(Intr)  
Days -0.138
```

## A model with uncorrelated random effects

- ▶ The data plots gave little indication of a systematic relationship between a subject's random effect for slope and his/her random effect for the intercept. Also, the estimated correlation is quite small.
- ▶ We should consider a model with uncorrelated random effects. To express this we use two random-effects terms with the same grouping factor and different left-hand sides. In the formula for an `lmer` model, distinct random effects terms are modeled as being independent. Thus we specify the model with two distinct random effects terms, each of which has `Subject` as the grouping factor. The model matrix for one term is intercept only (`1`) and for the other term is the column for `Days` only, which can be written `0+Days`. (The expression `Days` generates a column for `Days` and an intercept. To suppress the intercept we add `0+` to the expression; `-1` also works.)

## A mixed-effects model with independent random effects

```
Linear mixed model fit by REML
Formula: Reaction ~ Days + (1 | Subject) + (0 + Days | Subject)
Data: sleepstudy
   AIC   BIC logLik deviance REMLdev
1754 1770 -871.8   1752   1744
Random effects:
Groups   Name      Variance Std.Dev.
Subject (Intercept) 627.568  25.0513
Subject Days         35.858   5.9882
Residual                    653.584  25.5653
Number of obs: 180, groups: Subject, 18
Fixed effects:
              Estimate Std. Error t value
(Intercept)  251.405     6.885    36.51
Days          10.467     1.559     6.71

Correlation of Fixed Effects:
(Intr)
Days -0.184
```

## Conclusions from the likelihood ratio test

- ▶ Because the large p-value indicates that we would not reject `fm2` in favor of `fm1`, we prefer the more parsimonious `fm2`.
- ▶ This conclusion is consistent with the AIC (Akaike's Information Criterion) and the BIC (Bayesian Information Criterion) values for which "smaller is better".
- ▶ We can also use a Bayesian approach, where we regard the parameters as themselves being random variables, is assessing the values of such parameters. A currently popular Bayesian method is to use sequential sampling from the conditional distribution of subsets of the parameters, given the data and the values of the other parameters. The general technique is called *Markov chain Monte Carlo* sampling.
- ▶ The `lme4` package has a function called `mcmcsamp` to evaluate such samples from a fitted model. At present, however, there seem to be a few "infelicities", as Bill Venables calls them, in this function.

## Comparing the models

- ▶ Model `fm1` contains model `fm2` in the sense that if the parameter values for model `fm1` were constrained so as to force the correlation, and hence the covariance, to be zero, and the model were re-fit, we would get model `fm2`.
- ▶ The value 0, to which the correlation is constrained, is not on the boundary of the allowable parameter values.
- ▶ In these circumstances a likelihood ratio test and a reference distribution of a  $\chi^2$  on 1 degree of freedom is suitable.

```
> anova(fm2, fm1)
```

```
Data: sleepstudy
Models:
fm2: Reaction ~ Days + (1 | Subject) + (0 + Days | Subject)
fm1: Reaction ~ Days + (Days | Subject)
      Df      AIC      BIC logLik  Chisq Chi Df Pr(>Chisq)
fm2  5 1762.05 1778.01 -876.02
fm1  6 1763.99 1783.14 -875.99 0.0609      1      0.805
```

## Likelihood ratio tests on variance components

- ▶ As for the case of a covariance, we can fit the model with and without the variance component and compare the fit quality.
- ▶ As mentioned previously, the likelihood ratio is a reasonable test statistic for the comparison but the "asymptotic" reference distribution of a  $\chi^2$  does not apply because the parameter value being tested is on the boundary.
- ▶ The p-value computed using the  $\chi^2$  reference distribution should be conservative (i.e. greater than the p-value that would be obtained through simulation).

```
> fm3 <- lmer(Reaction ~ Days + (1 | Subject), sleepstudy)
> anova(fm3, fm2)
```

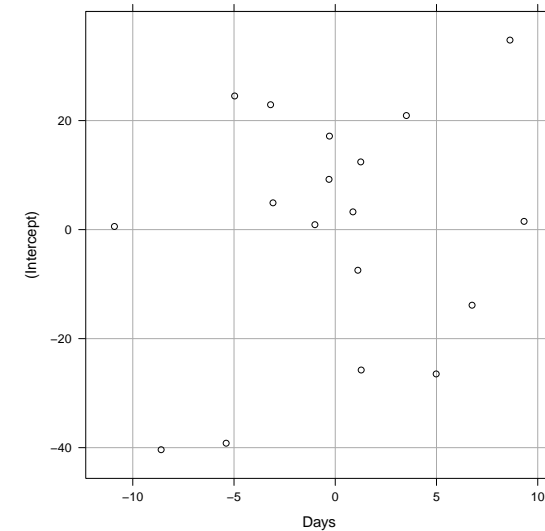
```
Data: sleepstudy
Models:
fm3: Reaction ~ Days + (1 | Subject)
fm2: Reaction ~ Days + (1 | Subject) + (0 + Days | Subject)
      Df      AIC      BIC logLik  Chisq Chi Df Pr(>Chisq)
fm3  4 1802.10 1814.87 -897.05
fm2  5 1762.05 1778.01 -876.02 42.053      1 8.885e-11
```

## Conditional means of the random effects

```
> (rr2 <- ranef(fm2))
```

```
$Subject
  (Intercept)    Days
308  1.5138201    9.3241219
309 -40.3749106   -8.5997562
310 -39.1816682   -5.3881596
330  24.5182906   -4.9689806
331  22.9140345   -3.1941494
332   9.2219311   -0.3085136
333  17.1560765   -0.2872253
334  -7.4515945    1.1160651
335   0.5774092  -10.9067061
337  34.7689483    8.6282046
349 -25.7541540    1.2807723
350 -13.8642119    6.7568576
351   4.9156063   -3.0753411
352  20.9294539    3.5124498
369   3.2587508    0.8731102
370 -26.4752098    4.9841221
371   0.9055257   -1.0053610
372  12.4219020    1.2584893
```

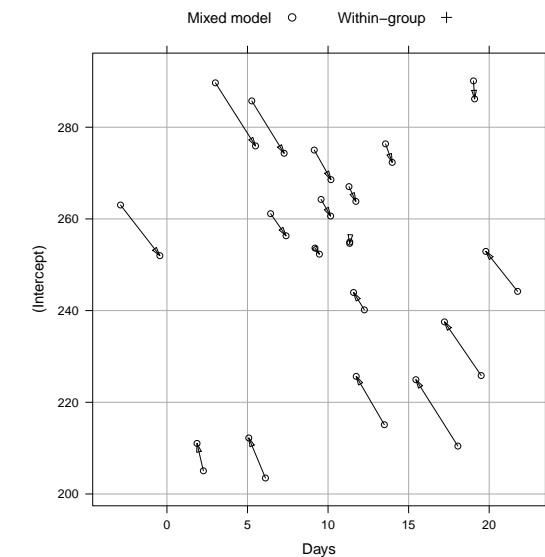
## Scatterplot of the conditional means



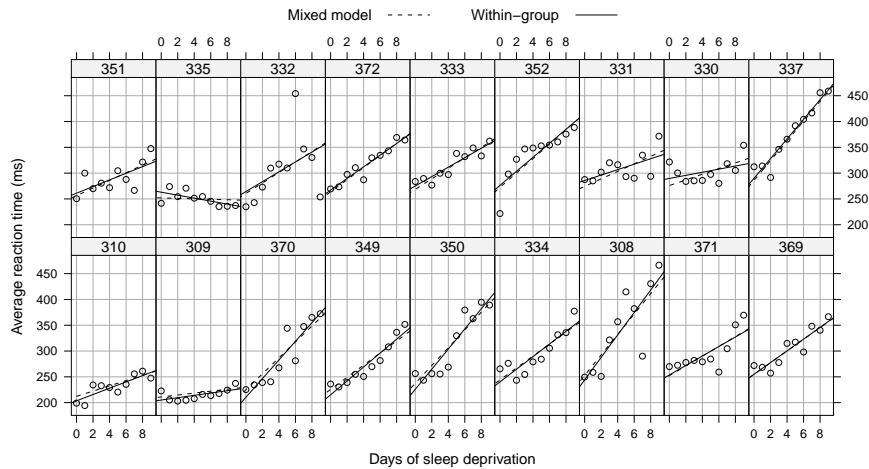
## Comparing within-subject coefficients

- ▶ For this model we can combine the conditional means of the random effects and the estimates of the fixed effects to get conditional means of the within-subject coefficients.
- ▶ These conditional means will be “shrunk” towards the fixed-effects estimates relative to the estimated coefficients from each subject’s data. John Tukey called this “borrowing strength” between subjects.
- ▶ Plotting the shrinkage of the within-subject coefficients shows that some of the coefficients are considerably shrunk toward the fixed-effects estimates.
- ▶ However, comparing the within-group and mixed model fitted lines shows that large changes in coefficients occur in the noisy data. Precisely estimated within-group coefficients are not changed substantially.

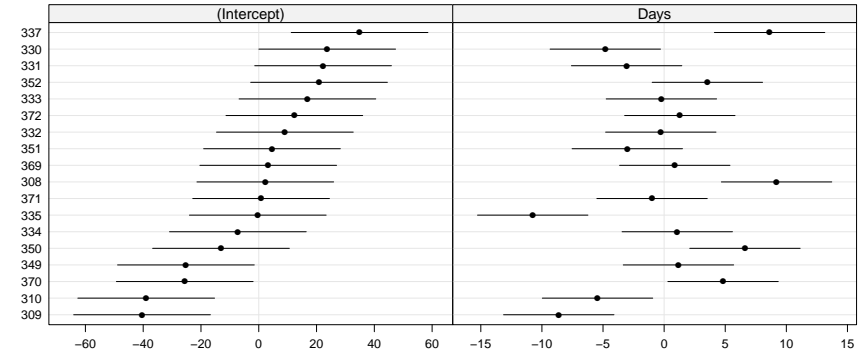
## Estimated within-group coefficients and BLUPs



## Observed and fitted



## Plot of prediction intervals for the random effects



Each set of prediction intervals have constant width because of the balance in the experiment.

## Conclusions from the example

- ▶ Carefully plotting the data is enormously helpful in formulating the model.
- ▶ It is relatively easy to fit and evaluate models to data like these, from a balanced designed experiment.
- ▶ We consider two models with random effects for the slope and the intercept of the response w.r.t. time by subject. The models differ in whether the (marginal) correlation of the vector of random effects per subject is allowed to be nonzero.
- ▶ The “estimates” (actually, the conditional means) of the random effects can be considered as penalized estimates of these parameters in that they are shrunk towards the origin.
- ▶ Most of the prediction intervals for the random effects overlap zero.