The problems seen in the field of biology, where observations are taken from clusters of repeated measurements taken from related units. So Linear Mixed Effects Model (LMM) which is a model with both fixed and random effects and used when the independence assumption of OLS regression is violated from the dependence of observations taken from repeated measurements or within clusters.

Case study:

Here Reaction variable: repeated measure for each subject and taken 10 times. This is what makes the data longitudinal, because each subject was measured multiple times. In this case, the Reaction variable is measuring the reaction time of the subject on a particular day.

Belenky et al. [2003] we consider the group of 18 subjects who were restricted to 3 hours of sleep/night for first 10 days of the trial.

2 covariates are Days (number of days of sleep deprivation) and Subject (**which day** each measurement happened)

(Fixed effect predictor variable) Days: ranges (0 to 9).

(Predictor variable & random effect) Subject : were numbered like 308, 309 310, 330, 331, 332, 333, 334, 335, 337, 349, 350, 351, 352, 369, 370, 371, 372.

Goal: accurately model the effect of sleep deprivation on the number of days a subject goes without sleep.

Problem: Multiple observations are taken from each subject which violates the assumption of independence**.**

Model and Methods

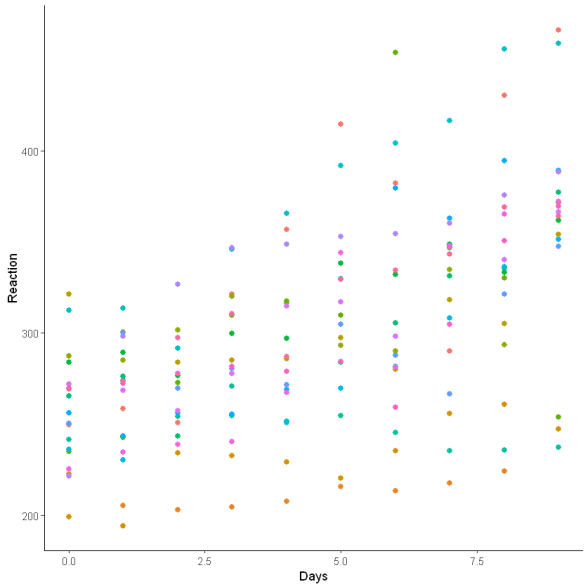
Intra-Class Correlation analysis

1 Plotting partial pooling in mixed-effects models
2. Random intercept, fixed predictor in individual level
3. Random intercept, random slope
4. Random intercept, individual and group level predictor
5. Random intercept, cross-level interaction
6. Model With Correlated and uncorrelated Random Effects

```
#null-model

model_null <- lmer(Reaction ~ 1 + (1 | Subject), data=sleepDat)

# Intraclass Correlation Coefficient < 0.5 hence, poor reliability
 Adjusted ICC: 0.395
 Conditional ICC: 0.395
```

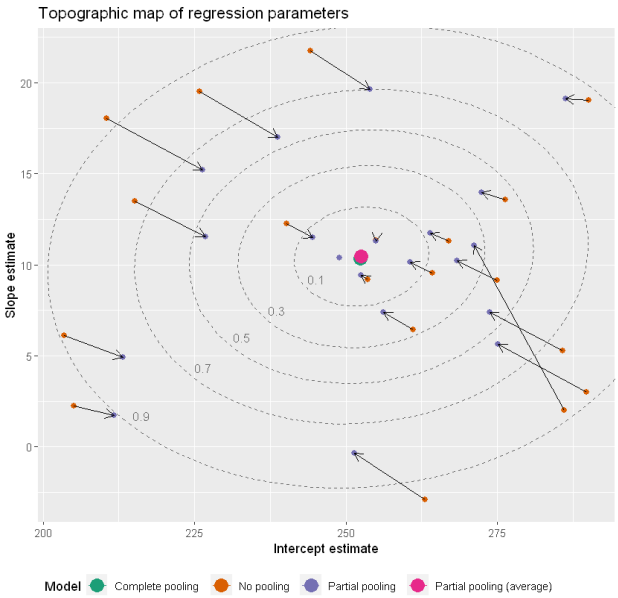


Reaction vs. Days by Subject

Each dots' colour = subject the observation belongs to trend example- orange are near the 200 reaction time range and the subject corresponding to purple dots are all near 350 reaction time.

To avoid this variation issue we can to fit a regression model for individual subject but it might create Type I error.

For modelling the differences in each subject, I fit a model which assigns a random intercept to each subject.



Partial pooling model for visualization:

The model needs to estimate variability around fixed effects so we can pool information from all the lines all together to improvise the estimates of individual lines. As we have18 trend lines for the participants so we can make guess about the trend lines for new participants as well.

lmer(Reaction ~ 1 + Days + (1 + Days | Subject)

Here we can see that the estimates are pulling towards the eclipse centre which is created by `**ellipse**()` [it t takes a covariance matrix, a centre value, and quantile/confidence level and returns the points from an oval around the centre at the given confidence level] As there is shrinkage near 250 - 275

1. **Random intercept, fixed predictor in individual level** i.e. The random effect will be subject to account for the variation between different subjects.

sleepModel2<-lmer(Reaction ~ Days + (1|Subject), sleepDat)

$Y = x_{ij}^T\beta + \gamma_i + \epsilon_{ij}$

Y_{ij} = response of the j^{th} member cluster i, i = 1,.., m, j = 1, ..., n_i

m = number of clusters (Subjects = 18)

n_i = size of cluster i (n = 10)

x_{ij}^T = covariate vector of j^{th} member of cluster i for fixed effects

β = fixed effect paramter γ = random effect parameter

$\gamma_i \sim N(0, D)$

$\epsilon \sim N(0, \Sigma_i)$

D = Covariance matrix of Random Effects γ_i

Σ_i = covariance matrix of error vector ϵ_i in cluster i

```
Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.2257 -0.5529  0.0109  0.5188  4.2506

Random effects:
 Groups   Name        Variance Std.Dev.
 Subject  (Intercept) 1378.2   37.12
 Residual              960.5   30.99
Number of obs: 180, groups:  Subject, 18

Fixed effects:
            Estimate Std. Error t value
(Intercept) 251.4051     9.7467   25.79
Days         10.4673     0.8042   13.02

Correlation of Fixed Effects:
     (Intr)
Days -0.371
```

```
$Subject
    (Intercept)     Days
308    292.1888 10.46729
309    173.5556 10.46729
310    188.2965 10.46729
330    255.8115 10.46729
331    261.6213 10.46729
332    259.6263 10.46729
333    267.9056 10.46729
334    248.4081 10.46729
335    206.1230 10.46729
337    323.5878 10.46729
349    230.2089 10.46729
350    265.5165 10.46729
351    243.5429 10.46729
352    287.7835 10.46729
369    258.4415 10.46729
370    245.0424 10.46729
371    248.1108 10.46729
372    269.5209 10.46729
```

Name	Estimateβ	Std. Error
(Intercept)	251.40	9.74
Days	10.46	0.80

Table 2. Random Intercept Model Summary of Fixed Effects

This table shows the coefficients and standard error for the random intercept model and that as days increases so does reaction time

Random Effects Summary from Random Intercept Model

Groups	Name	Variance	SD
Subject	(Intercept)	1378.2	37.12
Residual		960.5	30.99

From the table, the variance of the random effect accounts i.e. Subject accounts for $\frac{1378.2}{960.5+1378.2} = 59\%$ of the total variance explained by the fixed effect, Days and here residual variance is the variance which is not explained by the model, ϵ. The differences between subject account for 59% of the variance.

From the Coefficients for the random intercept model its evident that each subject has a different intercept but the same slope to help account for differences between subjects

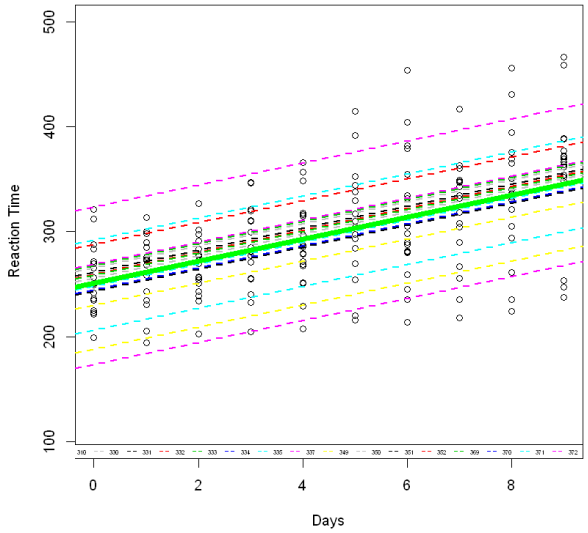


Figure. Random Intercept Linear Mixed Effects Model of Reaction Time vs. Days by Subject The thick Green line indicates the fixed effects model

The other dashed lines correspond to each individual subject

Here all subject share the same slope but a different intercept.

2. Random intercept, random slope (Unconditional growth model):

From our previous model, random effect looks like (Days|Subject) as we want to consider days effect on subjects **sleepModel3<-lmer(Reaction ~ Days + (Days|Subject), data= sleepDat)**

```
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 Subject  (Intercept) 612.10   24.741
          Days         35.07    5.922   0.07
 Residual             654.94   25.592
Number of obs: 180, groups:  Subject, 18

Fixed effects:
            Estimate Std. Error t value
(Intercept)  251.405      6.825  36.838
Days          10.467      1.546   6.771
```

The estimates of the standard deviations of the random effects for the intercept and the slope are 24.74 ms and 5.92 ms/day. The fixed-effects coefficients, β, are 251.4 ms and 10.46 ms/day for the intercept and slope.

Similar to the Random Intercept Model but with another term:

u_{ij}^{T} = covariate vector of the j^{th} member of

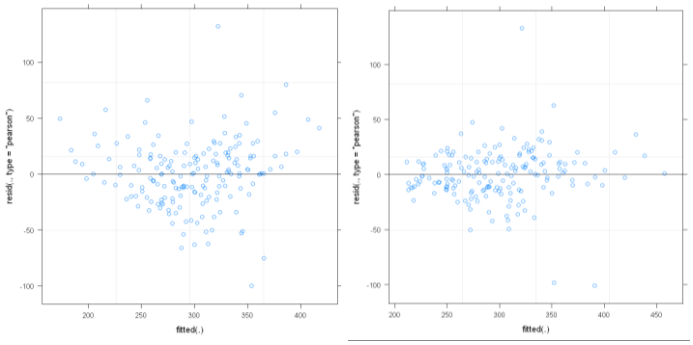
cluster i for random effects

This table shows that $\frac{612.10+35.07}{612.10+35.07+654.94} = 49.6\%$ of the variance is remaining after using for fixed effects discussed by subject.

Name	Estimateβ	Std. Error
(Intercept)	251.40	6.8
Days	10.46	1.54

Random Effects Summary from Random Intercept & Slope Model			
Groups	Name	Variance	SD
Subject	(Intercept)	612.10	24.74
	Days	35.07	5.92
Residual		654.94	25.592

Homogeneity of variance test: from both figures they do not look to have any visible pattern and spread points and so the homogeneity of variance assumpti on for both models looks to be satisfied.



Likelihood Ratio Test for Significance of Fixed Effects in Model		
Model	Test Statistic	P-Value
Random Intercept	116.46	3.764461e-27
Random Slope	23.53	1.22564e-06

p-value 1.22564e-06, indicates that including Days as a fixed effect is significant to the model.

Restricted Likelihood Ratio Test (Random Effects)		
Model	RLRT	P-Value
Random Intercept	107.2	<2.2e-16
Random Slope	42.7	<2.2e-16

The p-value in both model gives evidence in favor of the alternative hypothesis that the inclusion of the random intercept and random slope differs from zero

sleepDat$mygrp <- sample(1:5, size = 180, replace = TRUE)

#random intercept, individual and group level predictor

lmer(Reaction ~ Days + mygrp + (1 + Days | Subject), data=sleepDat)

```
Random effects:
 Groups   Name        Std.Dev. Corr
 Subject  (Intercept) 25.235
          Days         5.885   0.05
 Residual             25.595
Number of obs: 180, groups:  Subject, 18
Fixed Effects:
(Intercept)         Days        mygrp
     247.96        10.41         1.20
```

#random intercept, cross-level interaction

lmer(Reaction ~ Days * mygrp + (1 + Days | Subject), data=sleepDat)

```
Random effects:
 Groups   Name        Std.Dev. Corr
 Subject  (Intercept) 25.191
          Days         5.876   0.02
 Residual             25.408
Number of obs: 180, groups:  Subject, 18
Fixed Effects:
(Intercept)         Days        mygrp   Days:mygrp
   260.0935       7.4870      -2.8386       0.9423
```

Adding random groups did not make any noticeable change in the model, even in cross level interaction standard deviation seems similar to our **Random Intercept & Slope Model**

`CorrelatedRE: Reaction ~ 1 + Days + (1 + Days | Subject)`

the within-subject correlation between the random intercepts and the random slopes is low, as 0.06. Hence there is no evidence that the initial reaction time has systematic impact on the pace of increasing reaction time with sleep restriction

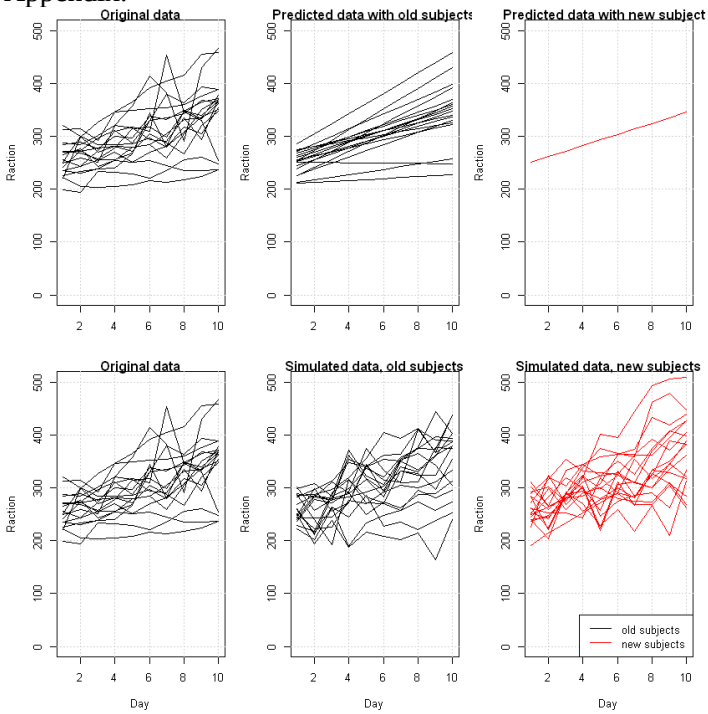
`UncorrelatedRE: Reaction ~ 1 + Days + (1 | Subject) + (0 + Days | Subject)`

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
UncorrelatedRE	5	1762.003	1777.968	-876.0016	1752.003	NA	NA	NA
CorrelatedRE	6	1763.939	1783.097	-875.9697	1751.939	0.06391066	1	0.8004185

Discussion

The Random Intercept model shows that each subject has a different starting reaction time but all subjects reaction time increases as the number of days of sleep deprivation increases whereas Random Slope model shows that each subject has a different starting reaction time but all but 1 subject has an increase in reaction time with an increase in the number of days with sleep deprivation. Each subjects reaction time increases or decreases is different between subjects. The $\chi 2$ statistic is 0.06 which is small and p-value of 0.80 indicating that the extra parameter in model CorrelatedRE compared to UncorrelatedRE does not have any significantly better fit. So its preferable to use the reduced UncorrelatedRF model. From these 2 models we can say that a subject's initial reaction time and when their reaction time is affected by sleep deprivation doesn't have strong relationship.

Appendix:



Days per subject in original, predicted, and simulated data of 18 new subjects.