Mixed Effects Models

Phillip M. Alday

05. August 2014

Today's data: Sleep Study

- `Reaction`: RT in ms

Today's data: Sleep Study

- `Reaction`: RT in ms
- `Days`: day 0 is normal sleep baseline (interval, i.e. Numeric)

Today's data: Sleep Study

- `Reaction`: RT in ms
- `Days`: day 0 is normal sleep baseline (interval, i.e. Numeric)
- `Subject`: numbered (categorical, non ordinal, i.e. Factor)

In R

```
> library(lme4)

## Loading required package: Matrix
## Loading required package: Rcpp

> data(sleepstudy)
> library(lattice)
> str(sleepstudy)

## 'data.frame':    180 obs. of  3 variables:
##  $ Reaction: num  250 259 251 321 357 ...
##  $ Days    : num  0 1 2 3 4 5 6 7 8 9 ...
##  $ Subject : Factor w/ 18 levels "308","309","310",..: 1
```

A quick warning

- timeo danaos et dona ferentes!

A quick warning

- timeo danaos et dona ferentes!
- Relax, it'll be okay.

Back to basics

Statistical Primitives

Basic methods all possible (if tedious) to calculate by hand:

- linear regression (OLS)

Statistical Primitives

Basic methods all possible (if tedious) to calculate by hand:

- linear regression (OLS)
- t-test

Statistical Primitives

Basic methods all possible (if tedious) to calculate by hand:

- linear regression (OLS)
- t-test
- F-test and AN(C)OVA

Statistical Primitives

Basic methods all possible (if tedious) to calculate by hand:

- linear regression (OLS)
- t-test
- F-test and AN(C)OVA
- χ^2-test, including LR-variant if you have a log-table

Statistical Primitives

Basic methods all possible (if tedious) to calculate by hand:

- linear regression (OLS)
- t-test
- F-test and AN(C)OVA
- χ^2-test, including LR-variant if you have a log-table
- (correlations)

General linear model

- just about everything in statistics based on the general linear model

General linear model

- just about everything in statistics based on the general linear model
- χ^2, *t*-test and its extension factorial ANOVA included

General linear model

- just about everything in statistics based on the general linear model
- χ^2, *t*-test and its extension factorial ANOVA included
- *t*-test between groups thus fully equivalent to coefficient tests in linear regression

General linear model

- just about everything in statistics based on the general linear model
- χ^2, t-test and its extension factorial ANOVA included
- t-test between groups thus fully equivalent to coefficient tests in linear regression
- ANOVA (F-test) thus fully equivalant to test F-test for overall model fit in linear gression

Repeated Measures

- independence assumption violated in repeated measures

Repeated Measures

- independence assumption violated in repeated measures
- repeated measures ANOVA based on a convenient trick for the special case involving categorial predictors

Repeated Measures

- independence assumption violated in repeated measures
- repeated measures ANOVA based on a convenient trick for the special case involving categorial predictors
- non trivial to extend this trick to ordinal or continuous predictors

Repeated Measures

- independence assumption violated in repeated measures
- repeated measures ANOVA based on a convenient trick for the special case involving categorial predictors
- non trivial to extend this trick to ordinal or continuous predictors
- r.m. ANOVA also restricted to modelling repetition in one dimension, while collapsing in all other dimensions

Repeated Measures

- independence assumption violated in repeated measures
- repeated measures ANOVA based on a convenient trick for the special case involving categorial predictors
- non trivial to extend this trick to ordinal or continuous predictors
- r.m. ANOVA also restricted to modelling repetition in one dimension, while collapsing in all other dimensions
  - traditionally separate F_1 (ANOVA by subjects) and F_2 (ANOVA by items) to model crossed random effects

Repeated Measures

- independence assumption violated in repeated measures
- repeated measures ANOVA based on a convenient trick for the special case involving categorial predictors
- non trivial to extend this trick to ordinal or continuous predictors
- r.m. ANOVA also restricted to modelling repetition in one dimension, while collapsing in all other dimensions
  - traditionally separate F_1 (ANOVA by subjects) and F_2 (ANOVA by items) to model crossed random effects
  - Clark (1973): combine these two tests into a single measure

Repeated Measures

- independence assumption violated in repeated measures
- repeated measures ANOVA based on a convenient trick for the special case involving categorial predictors
- non trivial to extend this trick to ordinal or continuous predictors
- r.m. ANOVA also restricted to modelling repetition in one dimension, while collapsing in all other dimensions
  - traditionally separate F_1 (ANOVA by subjects) and F_2 (ANOVA by items) to model crossed random effects
  - Clark (1973): combine these two tests into a single measure
- ANOVA sensitive to unbalanced designs and empty cells

Impasse

- all the usual stuff can be expressed as a variant of (generalized) linear regression

Impasse

- all the usual stuff can be expressed as a variant of (generalized) linear regression
- except repeated measures ANOVA

Impasse

- all the usual stuff can be expressed as a variant of (generalized) linear regression
- except repeated measures ANOVA
- so we have a choice

Impasse

- all the usual stuff can be expressed as a variant of (generalized) linear regression
- except repeated measures ANOVA
- so we have a choice
    - a detailed, full model with lots of subjects and items

Impasse

- all the usual stuff can be expressed as a variant of (generalized) linear regression
- except repeated measures ANOVA
- so we have a choice
  - a detailed, full model with lots of subjects and items
  - or

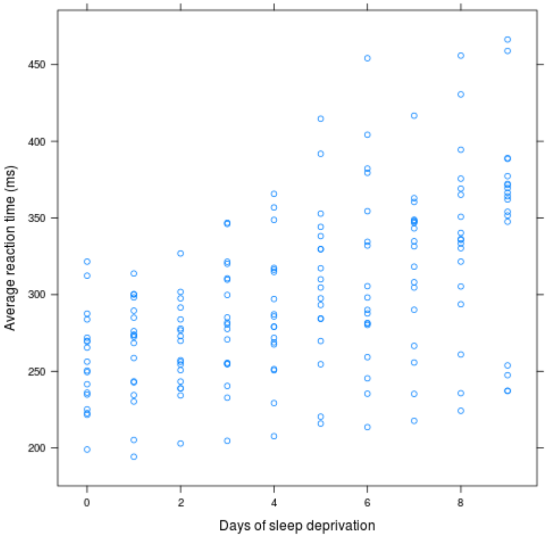
Impasse

- all the usual stuff can be expressed as a variant of (generalized) linear regression
- except repeated measures ANOVA
- so we have a choice
  - a detailed, full model with lots of subjects and items
  - or
  - ramming everything into a factorial model

What happens if we use linear regression on repeated measures data?

Linear Regression

```
> # simple scatter plot
> sleep.xy <- xyplot(Reaction ~ Days,data=sleepstudy,
+                    xlab = "Days of sleep deprivation",
+                    ylab = "Average reaction time (ms)")
```


Average reaction time (ms)
450
400
350
300
250
200
0
2
4
6
8
Days of sleep deprivation

Make a linear model

- basic line, no error term: $y = mx + b$

Make a linear model

- basic line, no error term: $y = mx + b$
- dep $=$ slope*indep $+$ baseline.offset

Make a linear model

- basic line, no error term: $y = mx + b$
- dep = slope*indep + baseline.offset
- outcome = (model) + error

Fit a line

- Fit a line to observed data with magic and matrices:

Fit a line

- Fit a line to observed data with magic and matrices:
  - $Y = \beta_1 X + \beta_0 + \epsilon$

Fit a line

- Fit a line to observed data with magic and matrices:
  - $Y = \beta_1 X + \beta_0 + \epsilon$
  - $Y = \beta_2 X + \beta_1 X + \beta_0 + \epsilon$

Fit a line

- Fit a line to observed data with magic and matrices:
  - $Y = \beta_1 X + \beta_0 + \epsilon$
  - $Y = \beta_2 X + \beta_1 X + \beta_0 + \epsilon$
  - $Y = \beta_3 X + \beta_2 X + \beta_1 X + \beta_0 + \epsilon$

Fit a line

- Fit a line to observed data with magic and matrices:
  - $Y = \beta_1 X + \beta_0 + \epsilon$
  - $Y = \beta_2 X + \beta_1 X + \beta_0 + \epsilon$
  - $Y = \beta_3 X + \beta_2 X + \beta_1 X + \beta_0 + \epsilon$
  - $\ldots$

Fit a line

- Fit a line to observed data with magic and matrices:
  - $Y = \beta_1 X + \beta_0 + \epsilon$
  - $Y = \beta_2 X + \beta_1 X + \beta_0 + \epsilon$
  - $Y = \beta_3 X + \beta_2 X + \beta_1 X + \beta_0 + \epsilon$
  - ...
- R has this built in:

```
> sleep.lm <- lm(Reaction~Days,data=sleepstudy)
```

Fit a line

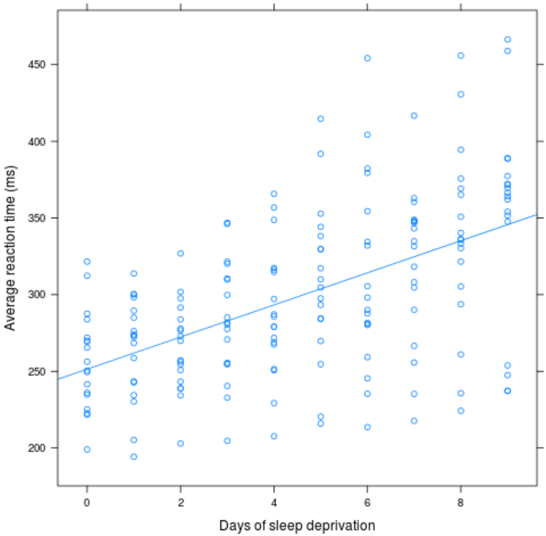
- Fit a line to observed data with magic and matrices:
  - $Y = \beta_1 X + \beta_0 + \epsilon$
  - $Y = \beta_2 X + \beta_1 X + \beta_0 + \epsilon$
  - $Y = \beta_3 X + \beta_2 X + \beta_1 X + \beta_0 + \epsilon$
  - ...
- R has this built in:

```
> sleep.lm <- lm(Reaction~Days,data=sleepstudy)
```

- additional predictors with + (no interaction) or * (interaction)

Add a regression line with lattice graphics

```
> # p for points, r for regression
> sleep.xy <- update(sleep.xy,type=c("p","r"))
```



Average reaction time (ms)
200
250
300
350
400
450
0
2
4
6
8
Days of sleep deprivation

Model summary

```
> summary(sleep.lm)

##
## Call:
## lm(formula = Reaction ~ Days, data = sleepstudy)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -110.85  -27.48    1.55   26.14  139.95
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   251.41       6.61   38.03  < 2e-16 ***
## Days           10.47       1.24    8.45  9.9e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '
##
## Residual standard error: 47.7 on 178 degrees of freedom
## Multiple R-squared:  0.286,  Adjusted R-squared:  0.282
## F-statistic: 71.5 on 1 and 178 DF,  p-value: 9.89e-15
```

Not a great fit!

Sidebar: ANOVA

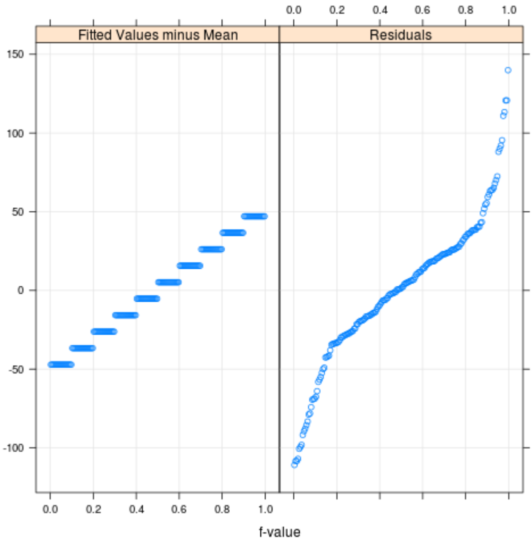
```
> anova(sleep.lm)

## Analysis of Variance Table
##
## Response: Reaction
##            Df Sum Sq Mean Sq F value  Pr(>F)
## Days        1 162703  162703    71.5 9.9e-15 ***
## Residuals 178 405252    2277
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '
```

But still not a great fit!

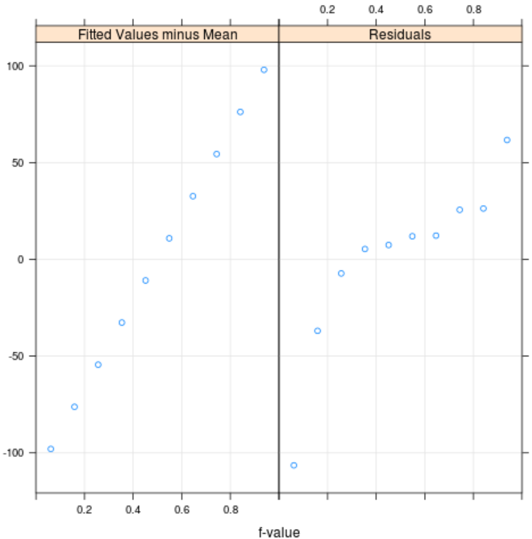
Residuals for all data

```
> rfs(sleep.lm)
```



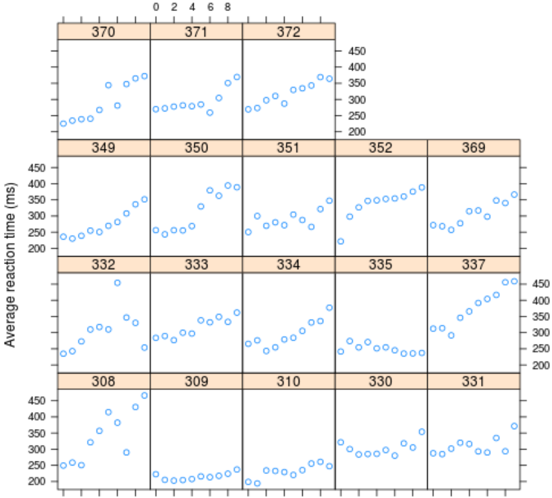
Residuals for a single subject

```
> sleep.lm.vp1 <- lm(Reaction ~ Days,
+                     data=sleepstudy[sleepstudy$Subject=="308",])
> rfs(sleep.lm.vp1)
```



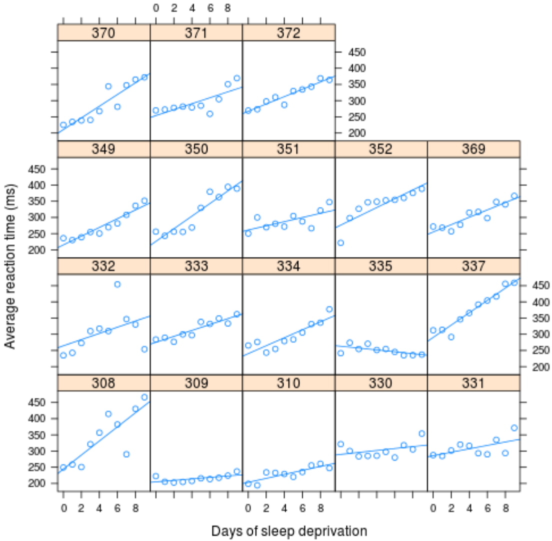
Models for single subjects

```
> sleep.xy.bysubj <- xyplot(Reaction ~ Days|Subject,
+                          data=sleepstudy,
+                          xlab = "Days of sleep deprivation",
+                          ylab = "Average reaction time (ms)")
> sleep.xy.bysubj
```



With regression lines

```
> sleep.xy.bysubj <- update(sleep.xy.bysubj,type=c("p","r"))
> sleep.xy.bysubj
```



What do repeated measures actually do to the data?

Variance and Repeated Measures

- inter- and intra- variance

Variance and Repeated Measures

- inter- and intra- variance
- random jitter from our choice of sample population

Variance and Repeated Measures

- inter- and intra- variance
- random jitter from our choice of sample population
- each subject fulfills a certain "condition", but random error pro instance of the condition

Variance and Repeated Measures

- inter- and intra- variance
- random jitter from our choice of sample population
- each subject fulfills a certain "condition", but random error pro instance of the condition
- similar idea for item analysis in linguistic designs

Fixed vs Random Effects

Subjects as fixed effect?

- only when we want to make intrasample predictions

Fixed vs Random Effects

Subjects as fixed effect?

- only when we want to make intrasample predictions
- i.e. sample==population

Fixed vs Random Effects

Subjects as fixed effect?

- only when we want to make intrasample predictions
- i.e. sample==population
- fixed means known variance / manipulation

Fixed vs Random Effects

Subjects as fixed effect?

- only when we want to make intrasample predictions
- i.e. sample==population
- fixed means known variance / manipulation
- fixed-effects: directed, preferably "exhaustive" manipulation

Fixed vs Random Effects

Subjects as random effects?

- "random" means unknown variance

Fixed vs Random Effects

Subjects as random effects?

- “random” means unknown variance
- error term is a random effect

Fixed vs Random Effects

Subjects as random effects?

- “random” means unknown variance
- error term is a random effect
- correction for the error resulting from our particular choice of sample

Fixed vs Random Effects

Subjects as random effects?

- "random" means unknown variance
- error term is a random effect
- correction for the error resulting from our particular choice of sample
- correction per grouping for slope and intercept possible

Fixed vs Random Effects

Subjects as random effects?

- "random" means unknown variance
- error term is a random effect
- correction for the error resulting from our particular choice of sample
- correction per grouping for slope and intercept possible
- error term per grouping!

Mixed Effects Models

- "Mixed" because both fixed random effects are used

```
> sleep.lmer <- lmer(Reaction ~ Days + (1|Subject),
+                    data=sleepstudy)
```

```
>     ?formula
```

Mixed Effects Models

- "Mixed" because both fixed random effects are used
- Same basic formula syntax `dep ~ indep | group`

```
> sleep.lmer <- lmer(Reaction ~ Days + (1|Subject),
+                    data=sleepstudy)
```

```
>     ?formula
```

Mixed Effects Models

- "Mixed" because both fixed random effects are used
- Same basic formula syntax `dep ~ indep | group`
- additional `(indep|group)` terms for random effects

```
> sleep.lmer <- lmer(Reaction ~ Days + (1|Subject),
+                    data=sleepstudy)
```

```
>     ?formula
```

Mixed Effects Models

- "Mixed" because both fixed random effects are used
- Same basic formula syntax `dep ~ indep | group`
- additional `(indep|group)` terms for random effects

```
> sleep.lmer <- lmer(Reaction ~ Days + (1|Subject),
+                           data=sleepstudy)
```

- More info here, here, and here

```
>     ?formula
```

Model Summary

```
>    summary(sleep.lmer)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Reaction ~ Days + (1 | Subject)
##    Data: sleepstudy
##
## REML criterion at convergence: 1786
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.226 -0.553  0.011  0.519  4.251
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 1378     37.1
##  Residual              960     31.0
## Number of obs: 180, groups:  Subject, 18
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  251.405      9.747    25.8
## Days          10.467      0.804    13.0
##
## Correlation of Fixed Effects:
##      (Intr)
## Days -0.371
```

Model Summary I

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Reaction ~ Days + (1 | Subject)
##    Data: sleepstudy
##
## REML criterion at convergence: 1786
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
## -3.226 -0.553  0.011  0.519  4.251
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 1378     37.1
##  Residual              960     31.0
## Number of obs: 180, groups:  Subject, 18
##
```

Model Summary II

```
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  251.405      9.747    25.8
## Days          10.467      0.804    13.0
##
## Correlation of Fixed Effects:
##      (Intr)
## Days -0.371
```

Fixed effect structure

- Package ez

Fixed effect structure

- Package `ez`
  - `ezMixed()` as a convenience for exploring fixed effects

Fixed effect structure

- Package `ez`
  - `ezMixed()` as a convenience for exploring fixed effects
  - `ezPredict()` useful for plotting regression lines

Fixed effect structure

- Package `ez`
  - `ezMixed()` as a convenience for exploring fixed effects
  - `ezPredict()` useful for plotting regression lines
- Package `effects`

Fixed effect structure

- Package `ez`
  - `ezMixed()` as a convenience for exploring fixed effects
  - `ezPredict()` useful for plotting regression lines
- Package `effects`
- Package `lmerTest`

Fixed effect structure

- Package `ez`
  - `ezMixed()` as a convenience for exploring fixed effects
  - `ezPredict()` useful for plotting regression lines
- Package `effects`
- Package `lmerTest`
- Package `languageR`

Fixed effect structure

- Package `ez`
  - `ezMixed()` as a convenience for exploring fixed effects
  - `ezPredict()` useful for plotting regression lines
- Package `effects`
- Package `lmerTest`
- Package `languageR`
- Package `LMERConvenienceFunctions`

Fixed effect structure

- Package `ez`
  - `ezMixed()` as a convenience for exploring fixed effects
  - `ezPredict()` useful for plotting regression lines
- Package `effects`
- Package `lmerTest`
- Package `languageR`
- Package `LMERConvenienceFunctions`
- Package `lmtest`

Models

```
> sleep.lmer <- update(sleep.lmer,REML=FALSE)
> null <- update(sleep.lmer, . ~ (1|Subject))
```

Likelihood-ratio test via model comparison

```
> # can only be used for nested models!
> anova(null,sleep.lmer)
```

```
## Data: sleepstudy
## Models:
## null: Reaction ~ (1 | Subject)
## sleep.lmer: Reaction ~ Days + (1 | Subject)
##            Df  AIC  BIC logLik deviance Chisq Chi Df Pr(>Chisq)
## null        3 1917 1926   -955     1911
## sleep.lmer  4 1802 1815   -897     1794   116      1     <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Random effect structure

- combine by-subject and by-item analyses in one step

Random effect structure

- combine by-subject and by-item analyses in one step
- cf. Clark (1973)

Random effect structure

- Early idea: build up from minimal structure until improvements don't bring you anything on ANOVA (R. Baayen, Davidson, and Bates 2008)

Random effect structure

- Early idea: build up from minimal structure until improvements don't bring you anything on ANOVA (R. Baayen, Davidson, and Bates 2008)
- New idea: Use the most complicated random effects structure possible (Barr et al. 2013)

Random effect structure

Possible random effect structures for ONE fixed factor:

1. Intercepts only by random factor:
   `(1 | random.factor)`

Random effect structure

Possible random effect structures for ONE fixed factor:

1. Intercepts only by random factor:
   `(1 | random.factor)`
2. Slopes only by random factor:
   `(0 + fixed.factor | random.factor)`

Random effect structure

Possible random effect structures for ONE fixed factor:

1. Intercepts only by random factor:
   `(1 | random.factor)`
2. Slopes only by random factor:
   `(0 + fixed.factor | random.factor)`
3. Intercepts and slopes by random factor:
   `(1 + fixed.factor | random.factor)`

Random effect structure

Possible random effect structures for ONE fixed factor:

1. Intercepts only by random factor:
   `(1 | random.factor)`
2. Slopes only by random factor:
   `(0 + fixed.factor | random.factor)`
3. Intercepts and slopes by random factor:
   `(1 + fixed.factor | random.factor)`
4. Intercept and slope, separately, by random factor:
   `(1 | random.factor) + (0 + fixed.factor | random.factor)`

Models

```
> sleep.lmer.slopes <- update(sleep.lmer,
+                               . ~ Days + (1+Days|Subject))
> sleep.lmer.slopes.int <- update(sleep.lmer,
+                                   . ~ Days + (0+Days|Subject))
```

Comparing Models

```
>     # can only be used for nested models!
>     anova(sleep.lmer,sleep.lmer.slopes, sleep.lmer.slopes.int)
```

```
## Data: sleepstudy
## Models:
## sleep.lmer: Reaction ~ Days + (1 | Subject)
## sleep.lmer.slopes.int: Reaction ~ Days + (0 + Days | Subject)
## sleep.lmer.slopes: Reaction ~ Days + (1 + Days | Subject)
##                       Df  AIC  BIC logLik deviance Chisq Chi Df Pr(>Chisq)
## sleep.lmer             4 1802 1815   -897     1794
## sleep.lmer.slopes.int  4 1782 1795   -887     1774  20.0      0     < 2e-16
## sleep.lmer.slopes      6 1764 1783   -876     1752  22.1      2     1.6e-05
##
## sleep.lmer
## sleep.lmer.slopes.int ***
## sleep.lmer.slopes     ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Judging Fit

- `anova()` function for `lmer()` provided for convenience and parallel to `lm()`

Judging Fit

- `anova()` function for `lmer()` provided for convenience and parallel to `lm()`
- χ^2 comparisons valid ONLY for nested models

Judging Fit

- `anova()` function for `lmer()` provided for convenience and parallel to `lm()`
- χ^2 comparisons valid ONLY for nested models
- use AIC or BIC otherwise

Judging Fit

- `anova()` function for `lmer()` provided for convenience and parallel to `lm()`
- χ^2 comparisons valid ONLY for nested models
- use AIC or BIC otherwise
  - no absolute good or bad

Judging Fit

- `anova()` function for `lmer()` provided for convenience and parallel to `lm()`
- χ^2 comparisons valid ONLY for nested models
- use AIC or BIC otherwise
  - no absolute good or bad
  - "smaller is better"

Judging Fit

- `anova()` function for `lmer()` provided for convenience and parallel to `lm()`
- χ^2 comparisons valid ONLY for nested models
- use AIC or BIC otherwise
  - no absolute good or bad
  - "smaller is better"
  - hard to determine what a significant difference is

Judging Fit

- `anova()` function for `lmer()` provided for convenience and parallel to `lm()`
- χ^2 comparisons valid ONLY for nested models
- use AIC or BIC otherwise
  - no absolute good or bad
  - "smaller is better"
  - hard to determine what a significant difference is
  - tips on AIC

Judging Fit

- `anova()` function for `lmer()` provided for convenience and parallel to `lm()`
- χ^2 comparisons valid ONLY for nested models
- use AIC or BIC otherwise
    - no absolute good or bad
    - "smaller is better"
    - hard to determine what a significant difference is
    - tips on AIC
- Use `REML=FALSE` when comparing models!

Judging Fit

- `anova()` function for `lmer()` provided for convenience and parallel to `lm()`
- χ^2 comparisons valid ONLY for nested models
- use AIC or BIC otherwise
  - no absolute good or bad
  - "smaller is better"
  - hard to determine what a significant difference is
  - tips on AIC
- Use `REML=FALSE` when comparing models!
- More advanced techniques for testing in package `pbkrtest`

relationship to AN(C)OVA

Relationship to ANOVA

- `ezANOVA()` depends on `aov()` which depends on `lm()`

Relationship to ANOVA

- `ezANOVA()` depends on `aov()` which depends on `lm()`
- `anova()` can be used to compare existing `lm()`s

Relationship to ANOVA

- `ezANOVA()` depends on `aov()` which depends on `lm()`
- `anova()` can be used to compare existing `lm()`s
- linear models compared with F and t tests

Relationship to ANOVA

- `ezANOVA()` depends on `aov()` which depends on `lm()`
- `anova()` can be used to compare existing `lm()`s
- linear models compared with F and t tests
- no continuous predictors with ANOVA

Relationship to ANOVA

- `ezANOVA()` depends on `aov()` which depends on `lm()`
- `anova()` can be used to compare existing `lm()`s
- linear models compared with F and t tests
- no continuous predictors with ANOVA
- ANOVA works on per-subject item averages and examines variance over subjects for each condition

Relationship to ANOVA

- `ezANOVA()` depends on `aov()` which depends on `lm()`
- `anova()` can be used to compare existing `lm()`s
- linear models compared with F and t tests
- no continuous predictors with ANOVA
- ANOVA works on per-subject item averages and examines variance over subjects for each condition
- MEMs work at an individual trial level and can accomodate empty cells and unbalanced designs!

Relationship to ANOVA

- `ezANOVA()` depends on `aov()` which depends on `lm()`
- `anova()` can be used to compare existing `lm()`s
- linear models compared with F and t tests
- no continuous predictors with ANOVA
- ANOVA works on per-subject item averages and examines variance over subjects for each condition
- MEMs work at an individual trial level and can accomodate empty cells and unbalanced designs!
- `anova()` can be used on individual `lm()`s and `lmer()`s to produce more traditional ANOVA-style output

ANOVA

```
anova(sleep.lmer.slopes)
```

```
## Analysis of Variance Table
##      Df Sum Sq Mean Sq F value
## Days  1  31798   31798    48.5
```

ANOVA

```
library(car)
Anova(sleep.lmer.slopes)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: Reaction
##      Chisq Df Pr(>Chisq)
## Days  48.5  1    3.2e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '
```

Wait, where are the p-values?

lmer, p-values and all that

- degrees of freedom not as trivial as you were led to believe in your basic stats courses

lmer, p-values and all that

- degrees of freedom not as trivial as you were led to believe in your basic stats courses
- not at all clear what a good way is to calculate this in the general case for mixed effects models

lmer, p-values and all that

- degrees of freedom not as trivial as you were led to believe in your basic stats courses
- not at all clear what a good way is to calculate this in the general case for mixed effects models
- t-distribution approximates normal (z) distribution for sufficiently high degrees of freedom

lmer, p-values and all that

- degrees of freedom not as trivial as you were led to believe in your basic stats courses
- not at all clear what a good way is to calculate this in the general case for mixed effects models
- t-distribution approximates normal (z) distribution for sufficiently high degrees of freedom
- so treat t-values as z values, which are significant at $\alpha = 0.05$ when $|z| > 2$ (cf. R. Baayen, Davidson, and Bates 2008)

lmer, p-values and all that

- degrees of freedom not as trivial as you were led to believe in your basic stats courses
- not at all clear what a good way is to calculate this in the general case for mixed effects models
- t-distribution approximates normal (z) distribution for sufficiently high degrees of freedom
- so treat t-values as z values, which are significant at $\alpha = 0.05$ when $|z| > 2$ (cf. R. Baayen, Davidson, and Bates 2008)
- forget p-values and traditional notions of significance behind!

Don't get us started

- focus on estimation, not significance

Don't get us started

- focus on estimation, not significance
- the purely fixed-effects model was able to demonstrate significance of the main effect

Don't get us started

- focus on estimation, not significance
- the purely fixed-effects model was able to demonstrate significance of the main effect
- but provided a poor overall description (fit) of the data – lousy estimate

Don't get us started

- focus on estimation, not significance
- the purely fixed-effects model was able to demonstrate significance of the main effect
- but provided a poor overall description (fit) of the data – lousy estimate
- this is the foundation of the new statistics (cf. Cummings 2014)

Don't get us started

- focus on estimation, not significance
- the purely fixed-effects model was able to demonstrate significance of the main effect
- but provided a poor overall description (fit) of the data – lousy estimate
- this is the foundation of the new statistics (cf. Cummings 2014)
- after all, "the goal is precision"

When things go wrong

Possible warning messages

- Convergence warnings: you don't have enough data for the proposed model structure

Possible warning messages

- Convergence warnings: you don't have enough data for the proposed model structure
- Singular: perfect multicollinearity (at least one variable is linear combination of the others)

Possible warning messages

- Convergence warnings: you don't have enough data for the proposed model structure
- Singular: perfect multicollinearity (at least one variable is linear combination of the others)
- Not positive definite: matrix not greater than "zero"; too much correlation / collinearity, not enough data

Generalized linear mixed models (GLMMs)

Extensions of linear models to non-linear data

- traditional linear models can be extended to model other types of data such as binary (e.g. yes/no responses)

Extensions of linear models to non-linear data

- traditional linear models can be extended to model other types of data such as binary (e.g. yes/no responses)
- basically works by strapping a transformation (link function) onto the front and back ends – R does this for you!

Extensions of linear models to non-linear data

- traditional linear models can be extended to model other types of data such as binary (e.g. yes/no responses)
- basically works by strapping a transformation (link function) onto the front and back ends – R does this for you!
- fixed effects: `glm()`

Extensions of linear models to non-linear data

- traditional linear models can be extended to model other types of data such as binary (e.g. yes/no responses)
- basically works by strapping a transformation (link function) onto the front and back ends – R does this for you!
- fixed effects: `glm()`
- mixed effects: `glmer()`

Family types

- binomial: (aka logistic regression) `binary ~ continuous`

Family types

- binomial: (aka logistic regression) `binary ~ continuous`
- Gaussian: (normal linear regression) `continuous ~ continuous` and `continuous ~ categegorial`

Family types

- binomial: (aka logistic regression) `binary ~ continuous`
- Gaussian: (normal linear regression) `continuous ~ continuous` and `continuous ~ categegorial`
- Gamma: `continuous ~ exp(continuous)` (exponential response)

Family types

- binomial: (aka logistic regression) `binary ~ continuous`
- Gaussian: (normal linear regression) `continuous ~ continuous` and `continuous ~ categegorial`
- Gamma: `continuous ~ exp(continuous)` (exponential response)
- Poisson: `count ~ continuous`

Family types

- binomial: (aka logistic regression) `binary ~ continuous`
- Gaussian: (normal linear regression) `continuous ~ continuous` and `continuous ~ categegorial`
- Gamma: `continuous ~ exp(continuous)` (exponential response)
- Poisson: `count ~ continuous`
- (inverse.gaussian, quasi, quasibinomial, quasipoisson)

Binomial models

- casuality of grouping

Binomial models

- casuality of grouping
  - traditional t-test vs. detection prediction

Binomial models

- casuality of grouping
  - traditional t-test vs. detection prediction
  - turn the traditional models on their head

Binomial models

- casuality of grouping
  - traditional t-test vs. detection prediction
  - turn the traditional models on their head
- existence / evidence for *a priori* categories

Binomial models

- casuality of grouping
  - traditional t-test vs. detection prediction
  - turn the traditional models on their head
- existence / evidence for *a priori* categories
  - connecting theory and empiry

Binomial models

- casuality of grouping
  - traditional t-test vs. detection prediction
  - turn the traditional models on their head
- existence / evidence for *a priori* categories
  - connecting theory and empiry
  - difficult vs non difficult violations

Binomial models

- casuality of grouping
  - traditional t-test vs. detection prediction
  - turn the traditional models on their head
- existence / evidence for *a priori* categories
  - connecting theory and empiry
  - difficult vs non difficult violations
- `behavioral ~ eeg`

Binomial models

- casuality of grouping
  - traditional t-test vs. detection prediction
  - turn the traditional models on their head
- existence / evidence for *a priori* categories
  - connecting theory and empiry
  - difficult vs non difficult violations
- `behavioral ~ eeg`
  - performance (cf. Vanrullen 2011)

Binomial models

- casuality of grouping
  - traditional t-test vs. detection prediction
  - turn the traditional models on their head
- existence / evidence for *a priori* categories
  - connecting theory and empiry
  - difficult vs non difficult violations
- `behavioral ~ eeg`
  - performance (cf. Vanrullen 2011)
  - anomaly detection

Even more advanced extensions

- General additive models extend linear models to arbitrary smooth functions

Even more advanced extensions

- General additive models extend linear models to arbitrary smooth functions
- Variants also available for mixed effects: generalized additive mixed models, implemented in R with the `gamm4` package

That's it, but I've added a bunch of further reading after this slide. . .

(More) References I

- FAQ from the mailing list (lots to absorb at first, but a good place to keep going back to)

(More) References II

- Generalized Linear Mixed Models (in English)

- (Kliegl et al. 2010)
- (Roehm, Sorace, and Bornkessel-Schlesewsky 2012)

More Academic Bibliography I

Baayen, R.H., D.J. Davidson, and D.M. Bates. 2008. "Mixed-Effects Modeling with Crossed Random Effects for Subjects and Items." *Journal of Memory and Language* 59 (390–412). doi:http://dx.doi.org/10.1016%2Fj.jml.2007.12.005.

Barr, Dale J, Roger Levy, Christoph Scheepers, and Harry J Tily. 2013. "Random Effects Structure for Confirmatory Hypothesis Testing: Keep It Maximal." *Journal of Memory and Language* 68: 255–78. doi:http://dx.doi.org/10.1016%2Fj.jml.2012.11.001.

Clark, Herbert H. 1973. "The Language-as-Fixed-Effect Fallacy: A Critique of Language Statistics in Psychological Research." *Journal of Verbal Learning and Verbal Behavior* 12: 335–59. doi:http://dx.doi.org/10.1016/S0022-5371(73)80014-3.

Cummings, Geoff. 2014. "The New Statistics: Why and How." *Psychological Science* 20 (10): 1–23. doi:10.1177/0956797613504966.

More Academic Bibliography II

Kliegl, Reinhold, Ping Wei, Michael Dambacher, Ming Yan, and Xiaolin Zhou. 2010. "Experimental Effects and Individual Differences in Linear Mixed Models: Estimating the Relationship Between Spatial, Object, and Attraction Effects in Visual Attention." *Front Psychol* 1: 238. doi:10.3389/fpsyg.2010.00238.

Roehm, Dietmar, Antonella Sorace, and Ina Bornkessel-Schlesewsky. 2012. "Processing Flexible Form-to-Meaning Mappings: Evidence for Enriched Composition as Opposed to Indeterminacy." *Language and Cognitive Processes* iFirst: 1–31. doi:10.1080/01690965.2012.712143. `http://www.tandfonline.com/doi/abs/10.1080/01690965.2012.712143`.

Vanrullen, Rufin. 2011. "Four Common Conceptual Fallacies in Mapping the Time Course of Recognition." *Frontiers in Psychology* 2 (00365). doi:10.3389/fpsyg.2011.00365. `http://www.frontiersin.org/Journal/Abstract.aspx?s=851&name=perception_science&ART_DOI=10.3389/fpsyg.2011.00365`.

More Academic Bibliography III

Wieling, Martijn, John Nerbonne, and R Harald Baayen. 2011. "Quantitative Social Dialectology: explaining Linguistic Variation Geographically and Socially." *PLoS One* 6 (9): e23613. doi:10.1371/journal.pone.0023613.