Linear Mixed-Effects Models

(aka Statistics III)

Bernd Figner
b.figner@psych.ru.nl

Week 2: February 10, 2014

1

Today: Theory

- **Some admin stuff: Take-home exam date**
- **Brief recap from last week**
- **Theoretical background**
  - Some formulas and procedures
  - Intercepts and slopes
  - Fixed and random effects
- **Hands-on: My first mixed model**
- **Homework/lab session**

2

Take-Home Exam Date

In-class exam date: April 7

Take-home

- 2 weeks to finish the take-home exam
- If you want deadline 1 week before in-class...

→Hand out of take-home instruction: **March 17**

→Deadline to hand it in: **March 31**

3

Today: Theory

- Some admin stuff: Take-home exam date
- **Brief recap from last week**
- Theoretical background
  - Some formulas and procedures
  - Intercepts and slopes
  - Fixed and random effects
- Hands-on: My first mixed model
- Homework/lab session

4

Linear Mixed-Effects Models

- Many names (multilevel, fixed-and-random effects, ...)
- Main idea: like linear regression, but for grouped data
  - **Repeated-measures data**
  - Longitudinal data
  - "Nested" aka "hierarchical" data
- Many advantages, compared to traditional approaches (aggregating, rmANOVA), including:
- Data analyzed at multiple levels (trial, participant, ...)
  - No loss of information
  - More of the involved processes can be modeled

5

Advantages cont.

- Can better handle
  - unbalanced designs
  - heteroscedasticity
  - missing data (up to a point)
- Continuous and categorical predictors
- More weight given to more “trustworthy” data, e.g., participants with
  - more data
  - more reliable data
- Often better power (compared to rmANOVA etc)

6

But: Increased Complexity

- No formula-based solution
- Instead: Numeric estimating procedure that iterates until convergence
  - Takes time; sometimes no convergence
  - Math more complicated than ANOVA, lin reg, ...
- Relatively new; still under development
- Less agreement, more opinionated
- Different “traditions” or “schools of thought”
  → e.g., mixed-models versus multilevel models

7

Today: Theory

- Some admin stuff: Take-home exam date
- Brief recap from last week
- **Theoretical background**
  - Some formulas and procedures
  - Intercepts and slopes
  - Fixed and random effects
- Hands-on: My first mixed model
- Homework/lab session

8

Repeated Measures

- Each participant contributes more than 1 data point to the dependent variable (DV)
  - from 2 to hundreds or thousands
  - e.g., tasks like Stroop, Go/No-Go, AAT, ...
- Observations within a participant resemble each other → not independent
- „Grouping factor:“ participant
- Traditional approaches to deal with such data
  - repeated-measures ANOVA
  - averaging (creating „scores“)

9

Back to the Stroop Example

Multi-Trial Response Time Tasks

Stroop: Name the color (don‘t read the word)

- 24 congruent trials: **blue**, **red**, **yellow**, **green**
- 24 incongruent: **blue**, **red**, **yellow**, **green**

10

Data Structure

- **1920 data points**
  - 40 participants (20 f, 20 m)
  - 48 trials (24 congruent, 24 incongruent)
- **"Level 1"** (smallest data unit): **trial-level**
  - RT (= DV)
  - Type of trial: congruent or incongruent
  - Accuracy: did participant respond correctly or not
  - Trial order
  - ...
- **"Level 2:" participant-level**
  - Gender
  - Individual differences: IQ, Age, ...
  - Experimental condition
  - ...

11

Hypotheses

- **H1: Congruency**
  - Longer response times in incongruent than congruent trials
- **H2: Gender differences**
  - Main effect: Females are faster than males
  - Interaction: Congruency effect smaller in females than males

12

Two-step approach?

(Step 1) Separate linear regression per participant

- **DV:** 48 RTs
- **IVs**
  - Congruency (congruent, incongruent)
  - Trial number, ...

→ **For each participant, we get several coefficients**

(Step 2) Do tests on all participants' coefficients

- Congruency coefficient different from 0?
- Difference in intercept between females and males?
- Gender difference in Congruency coefficient?

→ **Good idea, but worthy of improvement**

13

To keep things simple for now

- **"Level 1"** (smallest data unit): **trial-level**
  - RT (= DV)
  - Type of trial: congruent or incongruent
  - Accuracy: did participant respond correctly or not
  - Trial order
  - ...
- **"Level 2:" participant-level**
  - Gender: as main effect and interaction with Congruency
  - Individual differences: IQ, Age, ...
  - Experimental condition
  - ...

14

Today: Theory

- Some admin stuff: Take-home exam date
- Brief recap from last week
- **Theoretical background**
  - Some formulas and procedures
  - Intercepts and slopes
  - Fixed and random effects
- Hands-on: My first mixed model
- Homework/lab session

15

Formulas

- Different types of notation in the literature
- Ultimately the same; highlight different aspects
- We'll encounter different ones in the course
- This one quite intuitive (→ *multi*-level)

16

Estimated Parameters

Let i = trial number, j = subject number
Congruency: 0=congruent; 1=incongruent; female=0, male=1

Equation 1: $RT_{ij} = B_{0j} + B_{1j} * \text{Congruency} + R_{ij}$

- B_{0j} = mean RT for subject j for congruent trials
- B_{1j} = difference in RT between incongruent and congruent trials for subject j
- R_{ij} = unexplained variance in RT across all trials (after accounting for mean RT across participants and effect of trial congruence)

- Equation 2: $B_{0j} = G_{00} + G_{01} * \text{Gender} + U_{0j}$
- Equation 3: $B_{1j} = G_{10} + G_{11} * \text{Gender} + U_{1j}$

17

- Equation 1: $RT_{ij} = B_{0j} + B_{1j} * \text{Congruency} + R_{ij}$

Equation 2: $B_{0j} = G_{00} + G_{01} * \text{Gender} + U_{0j}$

- G_{00} = mean RT for female participants
  - participants weighted in inverse proportion to their error variance
- G_{01} = overall size of the gender difference in RT
  - participants weighted in inverse proportion to their error variance
- U_{0j} = unexplained variance in mean RT
  - after accounting for participants' gender

- Equation 3: $B_{1j} = G_{10} + G_{11} * \text{Gender} + U_{1j}$

18

- Equation 1: $RT_{ij} = B_{0j} + B_{1j} * \text{Congruency} + R_{ij}$
- Equation 2: $B_{0j} = G_{00} + G_{01} * \text{Gender} + U_{0j}$

Equation 3: $B_{1j} = G_{10} + G_{11} * \text{Gender} + U_{1j}$

- G_{10} = effect of incongruent trials on RT
  - participants weighted in inverse proportion to their error variance
- G_{11} = gender difference in congruence-RT relationship
  - participants weighted in inverse proportion to their error variance
- U_{1j} = unexplained variance in congruence-RT relationship
  - after accounting for participants' gender

19

Mixed model to run in R

$RT_{ij} = G_{00} + G_{01}*\text{Gender} + G_{10}*\text{Congruency} +$
$G_{11}*\text{Gender}*\text{Congruency} + R_{ij} + U_{0j} + U_{1j}*\text{Congruency}$

Note: Congruency slope varying across participants

R, pretty please do the following...

... find parameter values that are most likely to lead to the observed (actual) data: → *Maximum Likelihood*

Put simply: Find best fit between model and data

20

Iterative Procedure: Iteration 1

- For each of 40 participants, run separate Level-1 regressions
  - For each of 40 participants, obtain a value of B_{0j} and a value of B_{1j}
    → save these values into Level-2 data file
  - For *each trial of each participant*, compute:
    $R_{ij} = RT_{ij} – B_{0j} – B_{1j} * \text{Congruency}$
  - For each of 40 participants, compute $V_j = var(R_{ij})$

21

Iteration 1 cont.

- With Level-2 data file, run two Level-2 regressions using B_{0j}, B_{1j}, and Gender
  - Obtain *one value* for each of the following:
    - G_{00} [average female RT]
    - G_{01} [gender difference coefficient]
    - G_{10} [In/Congruency coefficient]
    - G_{11} [Gender x In/Congruency coefficient]

22

Iteration 1 cont.

- For each of 40 participants, compute
  - $U_{0j} = B_{0j} - G_{00} - G_{01}$ * Gender [unexpl. var. mean RT, after gender]
  - $U_{1j} = B_{1j} - G_{10} - G_{11}$ * Gender [unexpl. var. congr/RT, after gender]
- Define
  - $TAU_{00} = var(U_{0j})$
  - $TAU_{11} = var(U_{1j})$
  - $TAU_{01} = covar(U_{0j}, U_{1j})$

23

Iteration 1 cont.

- For each of 40 participants, compute:
  - $L_{00j} = TAU_{00} / \{TAU_{00} + (V_j / \sqrt{n_j})\}$
  - same for L_{11j} and L_{01j}
- L_j ~ reliability of participant j
  - Note: for a particular participant
    - L_j ~ 0 if *large* variance in participant's trials
    - L_j ~ 1 if *no* variance in participant's trials
  - Participants with more trials → larger L_j's

24

Iteration 2

For each of 40 participants, compute NEW Betas:

- $B0j_{new} = L_j{*}B0j_{old} + (1 - L_j){*}(G00_{old} + G01_{old}{*}Gender)$
- $B1j_{new} = L_j{*}B1j_{old} + (1 - L_j){*}(G10_{old} + G11_{old}{*}Gender)$
- Compute $Rij_{new} = RT_{ij} - B0j_{new} - B1j_{new}{*}$ Congruency
- Compute $Vj_{new} = var(Rij_{new})$

25

Iteration 2 cont.

Run Level-2 regression with $B0j_{new}$, $B1j_{new}$

- Obtain: $G00_{new}$, $G01_{new}$, $G10_{new}$, $G11_{new}$

Compute

- $U0j_{new} = B0j_{new} - G00_{new} - G01_{new}{*}Gender$
- $TAU00_{new} = var(U0j_{new})$
- $L00j_{new} = TAU00_{new} / \{TAU00_{new} + (Vj_{new}/\sqrt{nj})\}$
- etc

26

And so on...

- Repeat above procedure...
- ...until two successive iterations...
- ...cause *very small* change in the likelihood function
- likelihood: measure how well the model fits the data

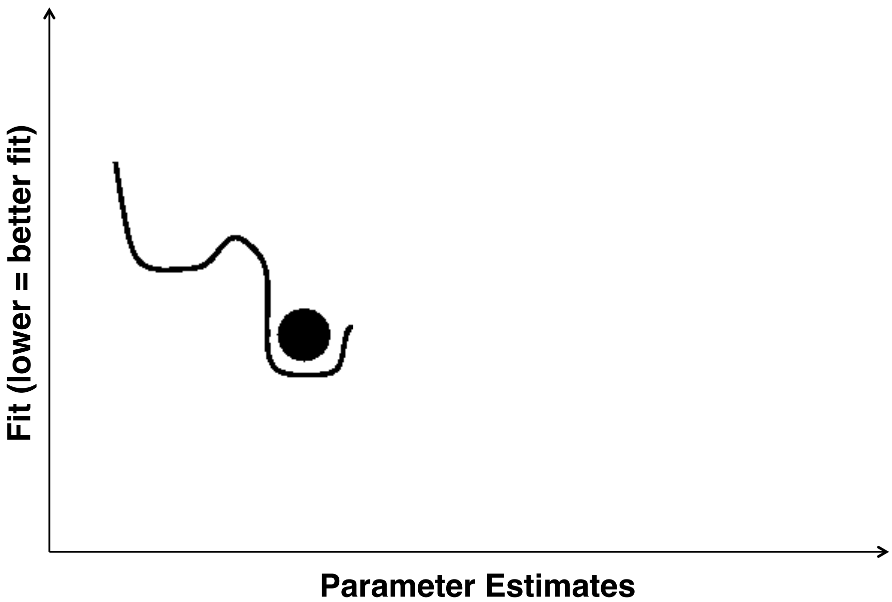
Put simply

→ Repeat until the model doesn't improve further

→ or until the maximum number of iterations is reached → Non-convergence warning!

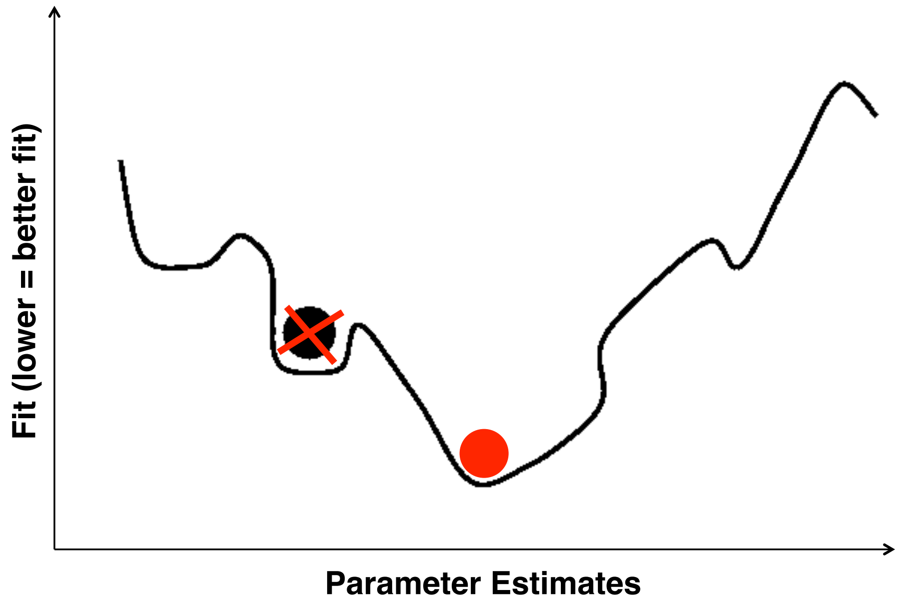
27

The Big Picture



28

The REALLY Big Picture



29

Optimization Procedure

- Complicated → science on its own...
  - ML: Maximum Likelihood
  - REML: Restricted Maximum Likelihood
- SAS, SPSS procedures: not public, but proprietary
- lme4: different "optimizers" available
  - bobyqa; Nelder-Mead; ...
  - default: bobyqa in most recent lme4 version
- Different optimizers: sometimes slightly different results
- Model non-convergence → try different optimizer

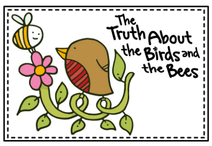
30

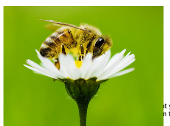
Today: Theory

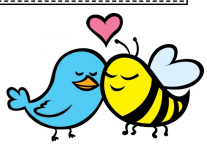
- Some admin stuff: Take-home exam date
- Brief recap from last week
- **Theoretical background**
  - Some formulas and procedures
  - Intercepts and slopes
  - Fixed and random effects
- Hands-on: My first mixed model
- Homework/lab session

31

Ok, it‘s time for that talk...







32

Fixed and Random Effects

33

Fixed and Random Effects

- Mixed Models: like linear regression, but for grouped (i.e., non-independent) data
- How do MMs handle dependence in the data?
- **They contain 2 different types of "effects"**
  - Fixed effects
  - Random effects

BTW: Some say we should call them *parameters*, not effects

34

Fixed Effects

→ What you know from typical linear regressions

- **(Fixed) Intercept**
  - Overall mean when all predictors are 0
- **(Fixed) Slope(s)**
  - Each slope describes relationship between 1 predictor and DV

35

- In **mixed-models**, we have those fixed effects, too
- But in addition, we have **random effects**
  - **Random intercept(s)**
    - Allow us to account for the non-independence in the data
    - E.g.: that each participant contributes several data points
  - **Random slope(s)**
    - Allow us to model participant ("group") differences in the DV-predictor relationships
    - E.g.: participants differ in magnitude of Stroop effect
  - And some more cool stuff (covariances between random effects) → for later

36

So far so good

- Now we know what they are good for in MMs
- But what are they? What's the difference between fixed and random effects? When is an effect fixed or random?

This is where things get a bit more complicated...

Unlike Fixed Effects, Random Effects are

- Not a single value (like fixed coefficient or intercept), but...
- A **Distribution of values**, described by mean (=0) and variance

37

Fixed and Random Effects

Lots of different definitions and conceptualizations

- The mathematical definition is not controversial
- The "meaning/interpretation" is controversial
- What do random versus fixed effects mean, what are their implications, when should which be used, … differs across authors

Good news

It doesn’t matter that much for us, because the more technical and pragmatic aspects are much more straightforward (and relevant, if I may add…)

38

A "Classical" Definition

Fixed effects

- The levels of the variable represent exhaustively all the values that exist in the **population** (i.e., not just in the current sample)
- Typical example: **Gender**
  → There are only males and females in the population (well…)
- Often also: Levels of a variable that are **directly manipulated** and **repeatable**
  → E.g., **Experiments**
  - The experimenter generates specific experimental conditions
  - Drug vs. placebo; congruent vs. incongruent; low versus high loss amount; ... (within-subject or between-subject)

39

Random effects

- The levels occurring in the **current sample** are only a subset of the levels occurring in the population (example: participants)
- I.e., the current levels in the sample are a (more or less) random sample of what exists in the population
- If we want to **generalize** from the sample to the population, we need to treat such variables as **random** effects
- If we treat them as fixed effects, we **must not** make generalizations about the population

→ **This is uncool and not what we want, if we want to test the significance of an effect!**

→ **Thus: Significance tests typically for fixed effects**

40

Different Conceptualizations...

Gelman & Hill (2007; p. 245)

- List 5 different definitions found in the literature (that contradict each other at least to some extent)
- Different authors give different (and sometimes unhelpful) advice
- G&H avoid the terms and instead talk about "modeled" and "unmodeled" coefficients (not sure this helps so much…)

41

More pragmatic (Gelman, 2005; same as in R/lme4)

Fixed Effects

Coefficients and intercepts that are **constant** across all "groups"

Random Effects

Coefficients and intercepts that **vary** across "groups"

→ **"Groups?!?"**

- Participants in the case of repeated measures
- Classes in the case of students nested in classes
- ...

42

Rules of Thumb

Participants

- Typically (more or less) random sample of students/ patients/...

→Virtually always to be modeled as random effects

→ Random **intercept**

Stimuli

- **Sometimes carefully created and varied**
- Risky choice task: Number of loss cards, gain and loss amount
- Fixed effects! (fixed slopes)

→ But: participants can differ in these effects, thus:

→ **fixed + random slopes**

- **Fixed part:** what is common across all participants ("average" slope)
- **Random part:** how participants differ from this average slope (variation around the average effect)

Stimuli (continued)

- **Sometimes they are a random sample**
  - **AAT with happy/angry faces** (chosen from a data base)
    - Depicted face ("identity") not of interest, only whether it is happy or angry

    → Happy/angry expression → fixed effect
    → "Identity" of face → random effect
  - **Stroop:** some colors chosen from all possible colors
  - **Lexicographic tasks:** words and non-words; ...

→ Just like participants represent a random sample, such stimuli represent also a random sample

→ **Random intercepts** for (a) participants and (b) stimuli

→ **"Crossed"** or **"orthogonal"** random effects

45

All confused?

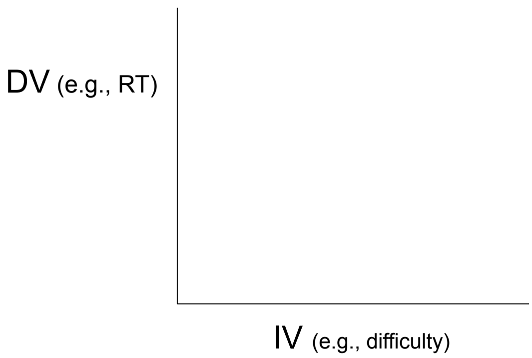
46

Let's try some graphs

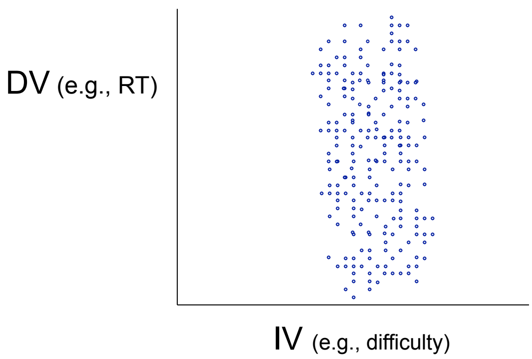
A simple but common case: Random intercepts for participants

(example: a repeated-measures task)

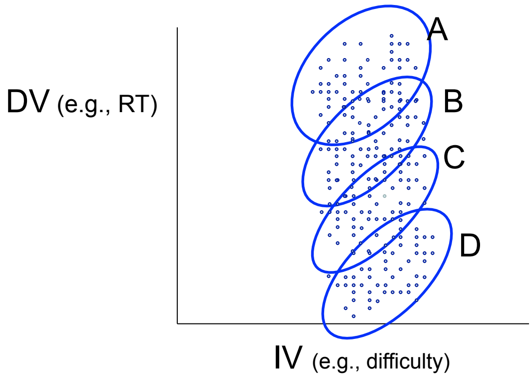
A simple data set: 4 participants
A case for **random intercepts**…



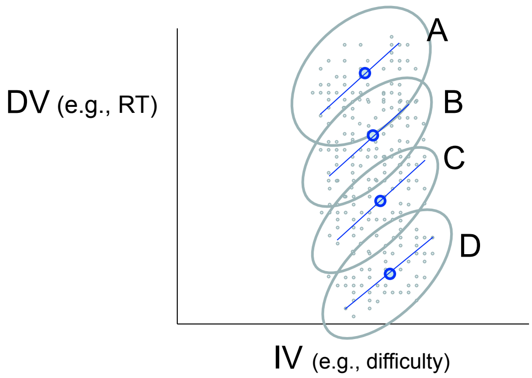
A simple data set: 4 participants
A case for **random intercepts**…



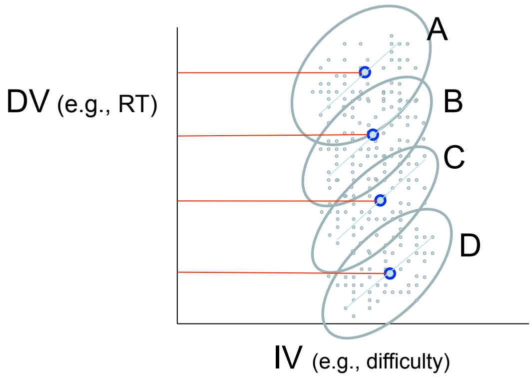
A simple data set: 4 participants
A case for **random intercepts**…



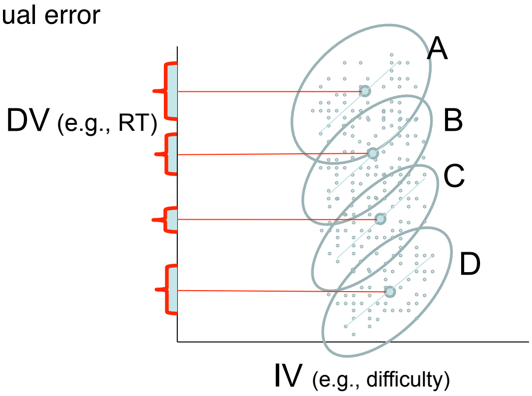
A simple data set: 4 participants
A case for **random intercepts**…



A simple data set: 4 participants
A case for **random intercepts**…

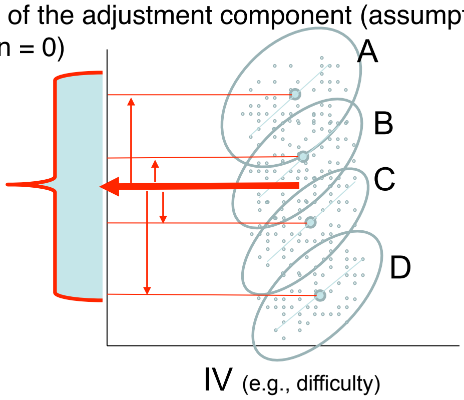


- Much variance due to each participant's average slowness
- Once we account for that, all show similar IV-DV relationship (slope)
- **The model's job:** Where does what variance come from?
  - "grouping factors" (here: participants)
  - IVs
  - residual error



- **Intercept: 2 components (fixed and random)**
  - Overall average intercept (like a grand mean) → **fixed**
  - Participant-specific "adjustments" to this overall intercept → **random**

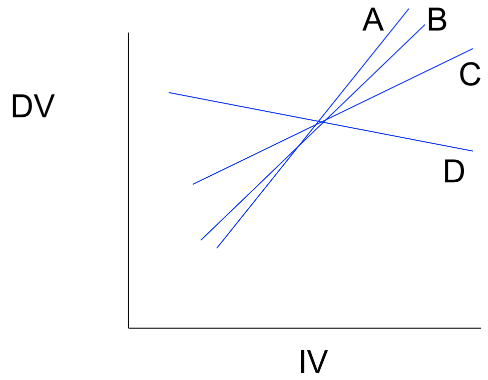
→ **2 parameters are estimated**
  - Value of fixed intercept
  - **Variance** of the adjustment component (assumption: normal distribution with mean = 0)



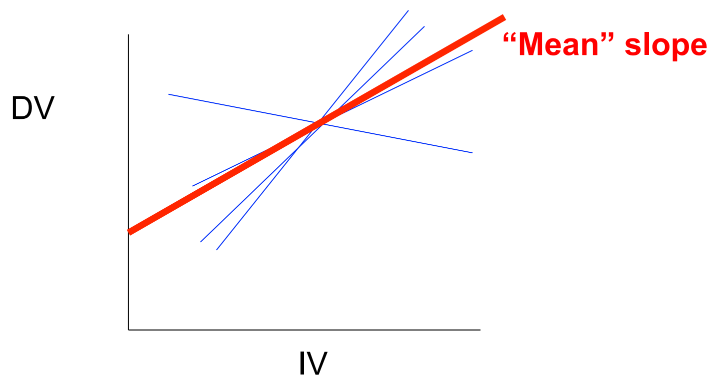
The same idea for random SLOPES

56

Random slopes

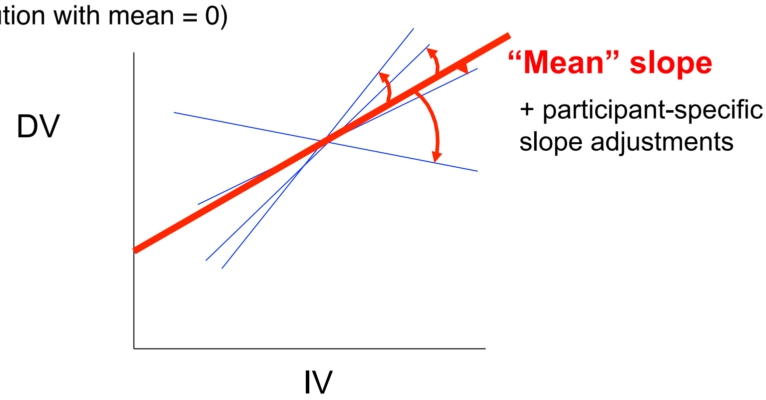


Random slopes

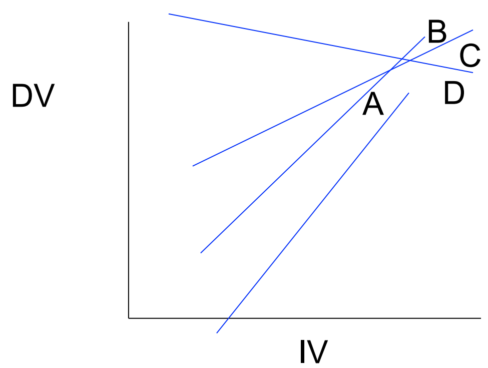


- **Slope: 2 components (fixed and random)**
  - Overall average slope → **fixed**
  - Participant-specific "adjustments" → **random**

→ **2 parameters are estimated**
  - Value of fixed slope
  - **Variance** of the adjustment component (assumption: normal distribution with mean = 0)



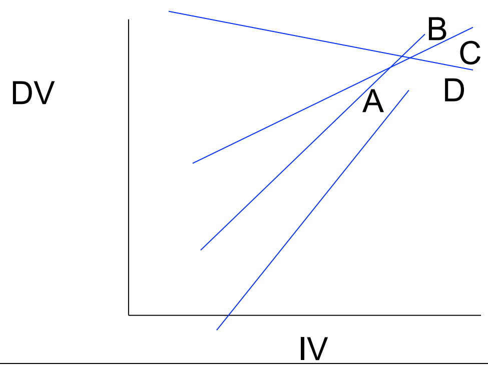
And typically intercepts **and** slopes can be random…



Estimated parameters

- **Intercept:** 2 parameters
  → fixed part + variance of adjustments
- **Slope:** 2 parameters
  → fixed part + variance of adjustments
- Often 1 more: **Covariance** between random Intercept and Slope Adjustments

→ E.g., participants with greater intercept have less steep slope



Real Data

62

Sleep Study

```
library(lme4)
?sleepstudy
```

Sleep Deprivation Study (*N* = 18)

- The average reaction time per day for subjects in a sleep deprivation study.
- On day 0 the subjects had their normal amount of sleep.
- Starting that night they were restricted to 3 hours of sleep per night; this was done for 9 consecutive nights.
- The observations represent the average reaction time on a series of tests given each day to each subject.

63

Let's have a look!

```
sleepstudy[1:20,]
```

```
   Reaction Days Subject
1  249.5600    0     308
2  258.7047    1     308
3  250.8006    2     308
4  321.4398    3     308
5  356.8519    4     308
6  414.6901    5     308
7  382.2038    6     308
8  290.1486    7     308
9  430.5853    8     308
10 466.3535    9     308
11 222.7339    0     309
12 205.2658    1     309
13 202.9778    2     309
14 204.7070    3     309
15 207.7161    4     309
16 215.9618    5     309
17 213.6303    6     309
18 217.7272    7     309
19 224.2957    8     309
20 237.3142    9     309
```

64

What's the data structure?

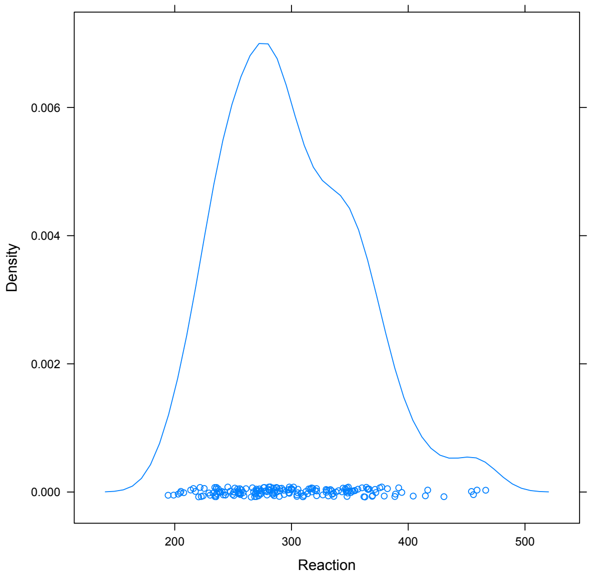
```
with(sleepstudy, table(Subject, Days))

       Days
Subject 0 1 2 3 4 5 6 7 8 9
    308 1 1 1 1 1 1 1 1 1 1
    309 1 1 1 1 1 1 1 1 1 1
    310 1 1 1 1 1 1 1 1 1 1
    330 1 1 1 1 1 1 1 1 1 1
    331 1 1 1 1 1 1 1 1 1 1
    332 1 1 1 1 1 1 1 1 1 1
    333 1 1 1 1 1 1 1 1 1 1
    334 1 1 1 1 1 1 1 1 1 1
    335 1 1 1 1 1 1 1 1 1 1
    337 1 1 1 1 1 1 1 1 1 1
    349 1 1 1 1 1 1 1 1 1 1
    350 1 1 1 1 1 1 1 1 1 1
    351 1 1 1 1 1 1 1 1 1 1
    352 1 1 1 1 1 1 1 1 1 1
    369 1 1 1 1 1 1 1 1 1 1
    370 1 1 1 1 1 1 1 1 1 1
    371 1 1 1 1 1 1 1 1 1 1
    372 1 1 1 1 1 1 1 1 1 1
```

65

Distribution of the DV

```
with(sleepstudy, densityplot(Reaction))
```



66

Hypothesis

- After more nights of sleep deprivation, participants have longer RTs

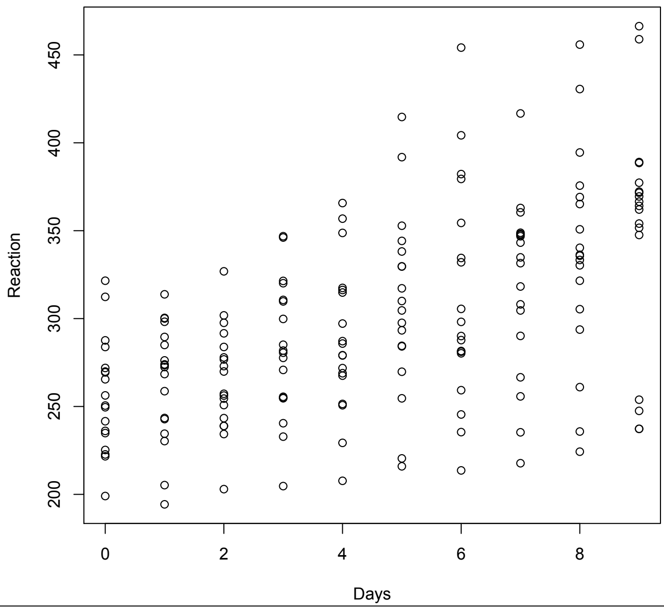
Our analysis?

- Like linear regression: `lm(Reaction~Days, data=sleepstudy)`
- but accounting for:
  - Repeated-measures (non-independence)
  - Participants might differ in
    - "Baseline Speed" (→ random intercept)
    - Sleep deprivation ("Days") effect (→ random slope)

67

RT as a function of "Days?"

- Hmm, scatterplot with Days on x and RT on y axis?



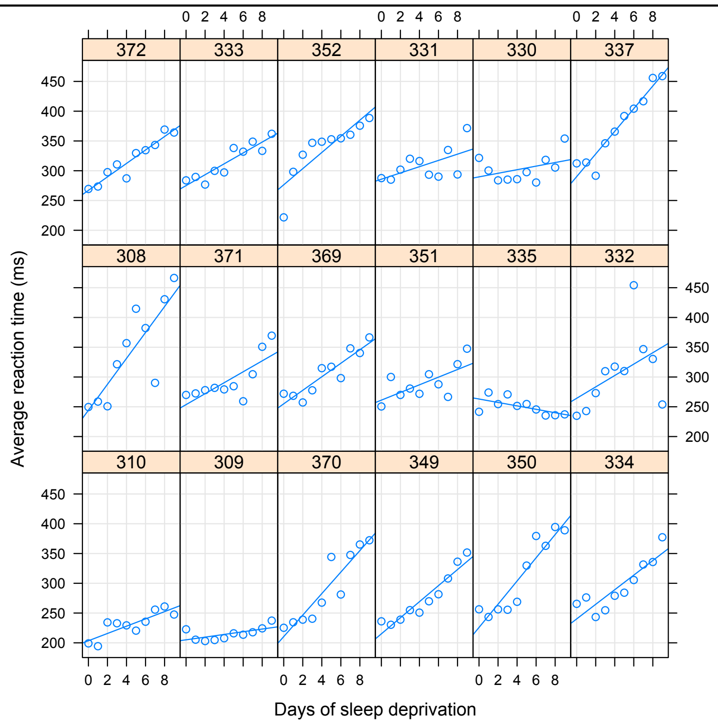
68

Differences in RT and Days Effect?

- Plot RT ~ Days relationship for each participant
- Pretty: use ggplot2!
- Lazy: use code from `?sleepstudy`

```
xyplot(Reaction ~ Days | Subject, sleepstudy,
   type = c("g","p","r"),
   index = function(x,y) coef(lm(y ~ x))[1],
   xlab = "Days of sleep deprivation",
   ylab = "Average reaction time (ms)",
   aspect = "xy")
```

69



70

Evidence for individual differences in

- RT on day 0 (→ random intercepts)
- Effect of sleep deprivation on RT (→ random slopes)

BUT: This is a repeated-measures data set, thus

- **Always** include random intercept per participant
  → To account for non-independence!
- **Always** include random slopes for within-subject effects
  → Barr, Levy, Scheepers, & Tily (2013)
  – Not doing so can lead to inflated Type I errors!

71

Today's Homework/Lab Session

(A) Instructions from IT how to run R in basement

Also on: BlackBoard → Course Documents

(B) Reading

- Work through this tutorial:
  http://www.bodo-winter.net/tutorial/bw_LME_tutorial.pdf
- Baayen et al. (2008) (thoroughly!)
  http://www.sfs.uni-tuebingen.de/~hbaayen/publications/baayenDavidsonBates.pdf

(C) Sign up here (don't post homework-related questions there!)

- https://stat.ethz.ch/mailman/listinfo/r-sig-mixed-models
- Read through all emails from now on (don't worry if you don't understand everything)

→ **Deadline** for finishing homework: February 17, 2014

72

Hand in a text (paper or pdf) that goes something like this:

I, [first name, last name]—also known as S... [student number]—solemnly declare that I have finished my homework for week 2 of the mixed-models class. In particular, I:

- tested out the lab computers, working through the IT instructions we received and sent Bernd (b.figner@psych.ru.nl) an email if it worked and/or sent an email to Bernd and Wolter Jansen (w.jansen@ru.nl) if it did not work, including the error message(s) I received.
- read and worked through this tutorial: http://www.bodo-winter.net/tutorial/bw_LME_tutorial.pdf
- thoroughly read the Baayen et al. (2008) paper, ignoring the `pvals.fnc()` stuff as this is not recommended anymore
- signed up for the R-sig-mixed-models list and read all emails

Signed, [signature; place; date]

73

Instructions for lab computers

1. Login with your S number and youw password.
2. Open R3.02 from the Program Files or from the path "c:/windows/program files/R/R-3.0.2/bin/x64/Rgui.exe".
3. Try to install one of the packages needed for the course.
4. If you get an error message like Access Denied, do the following:
5. Open My Computer
6. Open your homedirectory (usually drive Z: or H:)
7. Look for the folder Documents.
8. Check wether you have documents in that folder, if so, backup them to a higher level.
9. After that, erase Documents.
10. Open R3.02 from the Program Files or from the path "c:/windows/program files/R/R-3.0.2/bin/x64/Rgui.exe".
11. Try to install one of the packages needed for the course.
12. If there is still an error:
13. Open My Computer
14. Select your homedirectory with one click of your mouse.
15. Use the right button from your mouse.
16. Select disconnect.
17. Open R3.02 from the Program Files or from the path "c:/windows/program files/R/R-3.0.2/bin/x64/Rgui.exe".
18. Try to install one of the packages needed for the course.

74

That's it for today's lecture!
Questions or comments?

À bientôt au kelder!
-1.55A/B