Mixed Effects Models in R

Semih Can Aktepe

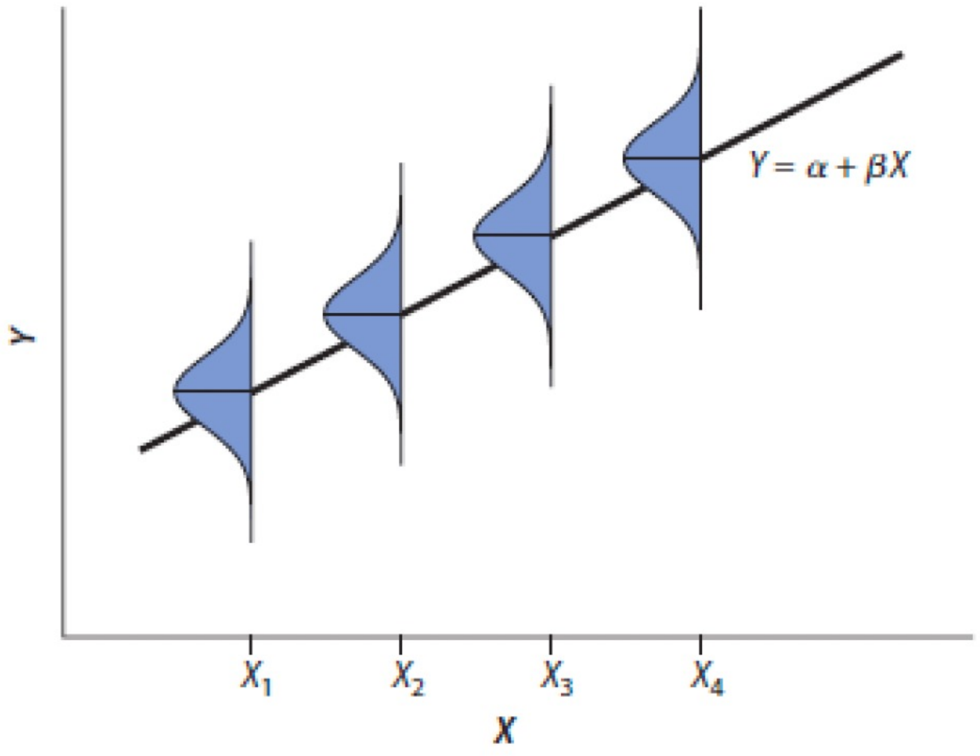
16.07.2024 Giessen

Outline

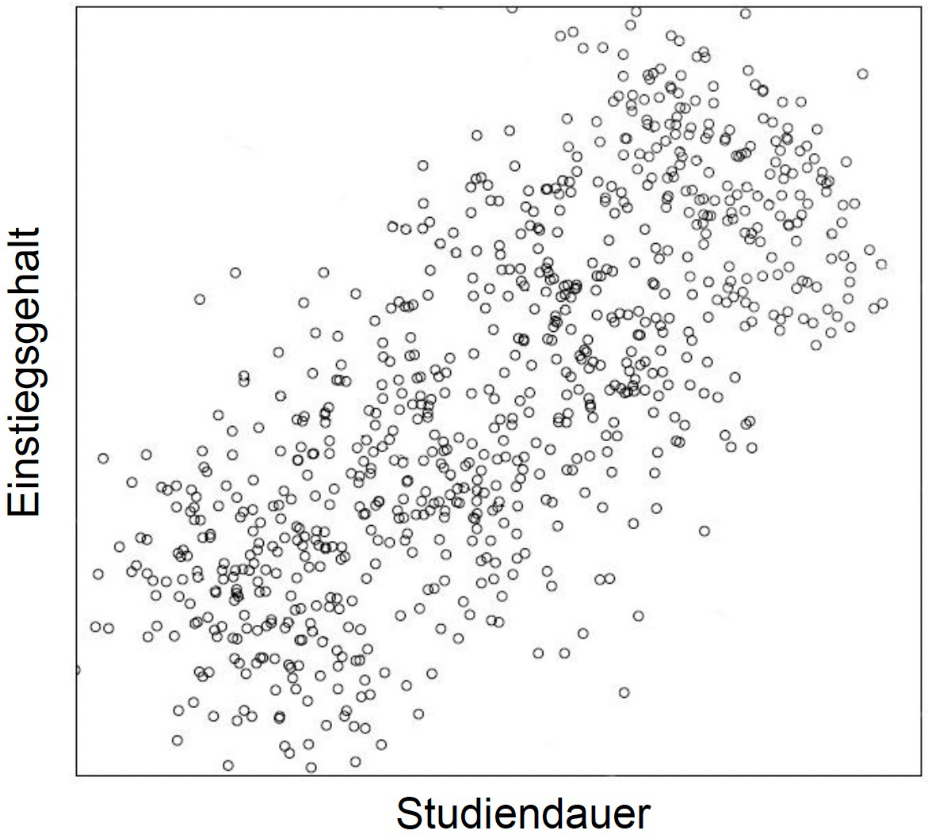
- Hierarchical Data
- Mixed Models
  - Basic Idea
  - Model Types
  - Intraclass Correlation
- Model Comparison
- Mixed Models in Action
- Practice (Analyzing Reaction Times)

Linear Regression

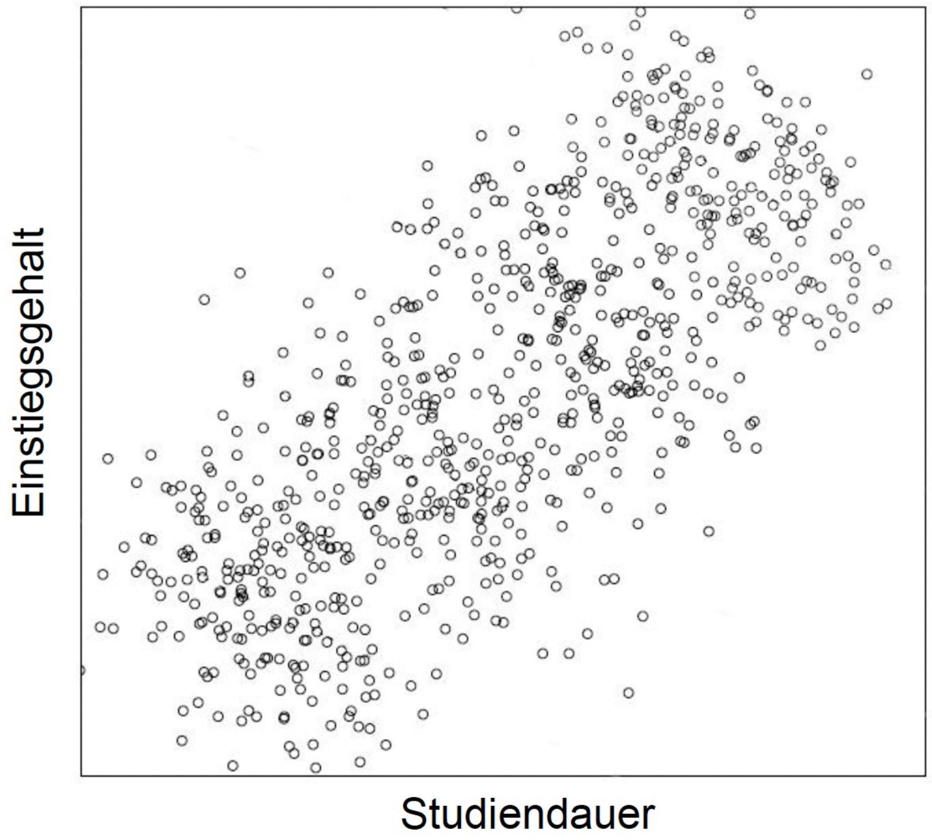
- Central Assumptions
  - Residuals are independent and identically distributed (iid).
  - Residuals do NOT depend on the random variable (e.g., participants, items).
- Violation of Assumtions
  - Observations are often dependent.
  - Through clusters of multi-level structures
  - Through repeated measures within people

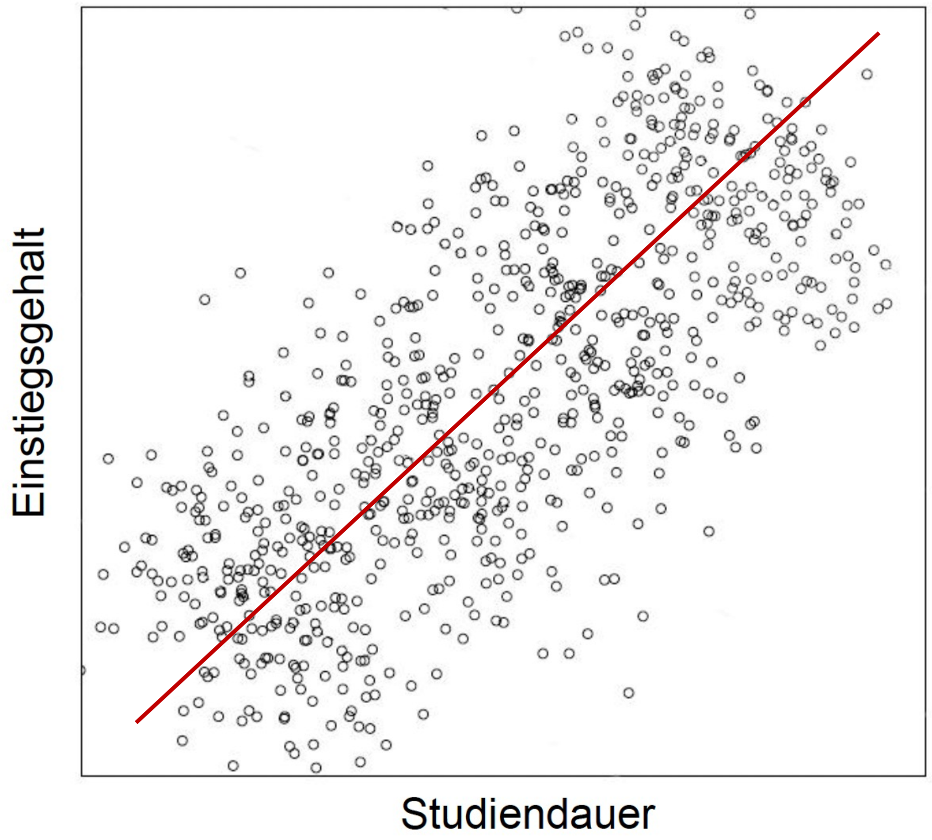


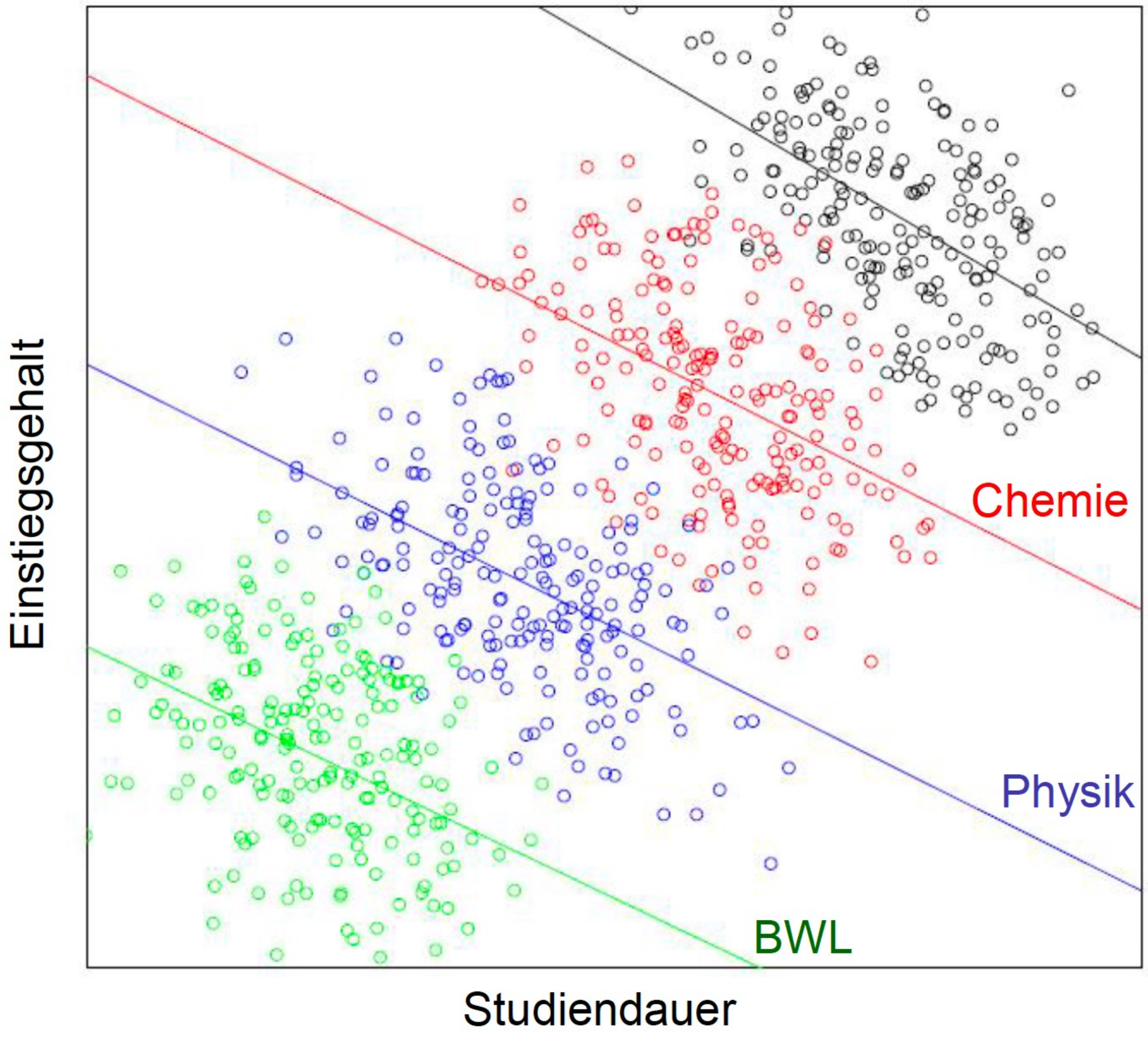
Example



Example


Einstiegsgehalt
Studiendauer


Einstiegsgehalt
Studiendauer

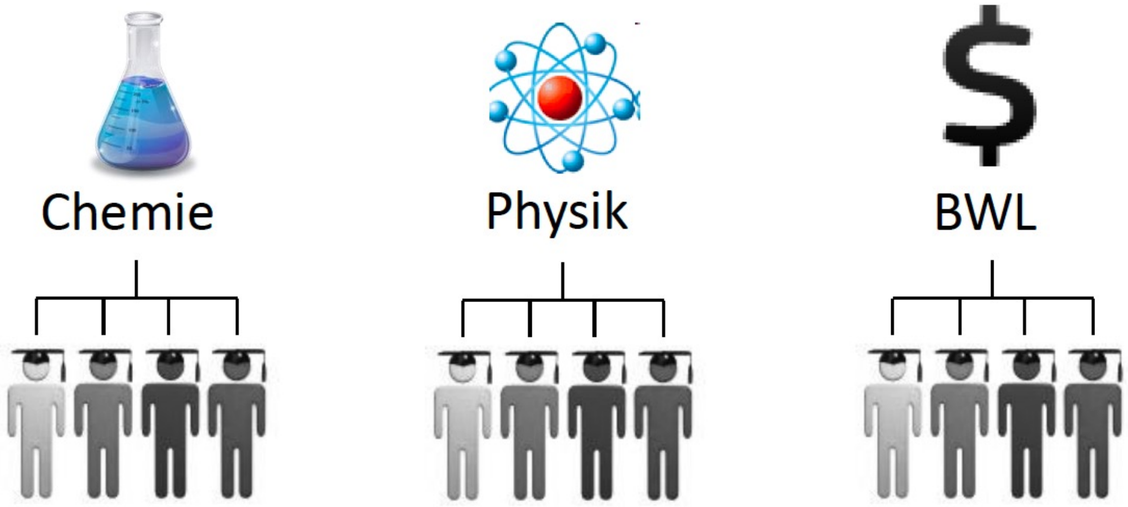

Chemie
Physik
BWL
Einstiegsgehalt
Studiendauer

Outline

- Hierarchical Data
- Mixed Models
  - Basic Idea
  - Model Types
  - Intraclass Correlation
- Model Comparison
- Mixed Models in Action
- Practice (Analyzing Reaction Times)

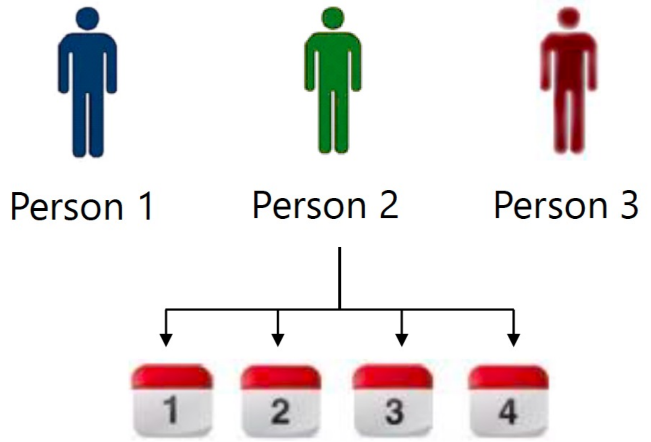
Hierarchical Data

- Data Structure
  - Several hierarchical organized levels
  - Clearly defined units within each level
  - Each unit is assigned to a unit of the next higher level (nested, clustered)



Consequences of Hierarchical Data

- More observations within one unit
  - Often not independent
  - Measurements within one unit are correlated.
  - Dependent observations provide less information than the independent ones.
  - Observations within a unit are influenced from each other.
    - Patients share disease history.
    - Sentences share syntactic structure.
    - Students share classroom context.



Consequences of Hierarchical Data

- Observations within a unit are more similar to each other than observation between the units.
- Violation of independence (data is not iid.)
- Ignoring such dependencies results in
  - Drastically underestimated standard errors
  - Misleading estimates and correlations

Data Structure

- Required data format is long format.
  - One observation per row
  - Units such as people, items, clinics are indexed.

ID	Messzeit-punkt	Kriterium
1	0	14
1	1	11
1	2	16
1	3	16
2	0	13
2	1	14
2	2	11
2	3	7
3	0	20
3	1	15
3	2	17
3	3	14
4	0	12
4	1	10
4	2	9
4	3	8
…	…	…

Fixed-Effects Analyses

- Small number of clusters (e.g., experimental groups, item types, etc.)
  - Cluster variable as categorical predictor
  - Assumption: Separate regression per cluster, separate error variance
  - Deal with potential dependencies between people, items etc.
  - Control differences in intercept and slope

Problem

- What if we have 10 clinics and 50 patients? 500 indicator variables?
- We are not interested in differences in patients.
- But if there is a difference, how can we introduce that to the model?

- Solution: Mixed (Multi-level, Hierarchical) Models

Outline

- Hierarchical Data
- Mixed Models
    - Basic Idea
    - Model Types
    - Intraclass Correlation
- Model Comparison
- Mixed Models in Action
- Practice (Analyzing Reaction Times)

Basic Idea

Level 1:
(für jede Person *i* in Gruppe *j*)

$$Y_{i,j} = \beta_{0,j} + \beta_{1,j} \cdot X_{i,j} + e_{i,j}$$

Level 2:
(für jede Gruppe *j*)

$$\beta_{0,j} = \gamma_{0(\beta_0)} + u_{j(\beta_0)}$$
$$\beta_{1,j} = \gamma_{0(\beta_1)} + u_{j(\beta_1)}$$

Basic Idea

$$
\begin{aligned}
Y_{i,j} &= \boxed{\beta_{0,j}} + \boxed{\beta_{1,j}} \cdot X_{i,j} + e_{i,j} \\
&= \left(\gamma_{0(\beta_0)} + u_{j(\beta_0)}\right) + \left(\gamma_{0(\beta_1)} + u_{j(\beta_1)}\right) \cdot X_{i,j} + e_{i,j} \\
&= \boxed{\gamma_{0(\beta_0)}} + \boxed{\gamma_{0(\beta_1)}} \cdot X_{i,j} + \boxed{u_{j(\beta_1)}} \cdot X_{i,j} + \boxed{u_{j(\beta_0)}} + e_{i,j}
\end{aligned}
$$

Mittlerer Achsenabschnitt

Mittlere Steigung

Abweichung der Steigung für Gruppe *j* von der mittleren Steigung: Level-2 Residuum

Abweichung im Achsenabschnitt für Gruppe *j* vom mittleren Achsenabschnitt: Level-2 Residuum

Level-1 Residuum: Abweichung zwischen beobachteten & erwarteten Wert für Beobachtung *i* in Gruppe *j*

Basic Idea

$$Y_{i,j} = \underbrace{\gamma_{0(\beta_0)} + \gamma_{0(\beta_1)} \cdot X_{i,j}}_{\text{Gruppen\underline{un}spezifische Effekte}} + \underbrace{u_{j(\beta_1)} \cdot X_{i,j} + u_{j(\beta_0)}}_{\text{Gruppenspezifische Effekte}} + e_{i,j}$$

(fester Teil, „fixed effects“)

(zufälliger Teil, „random effects“)

Fixed-Effects vs. Random-Effects

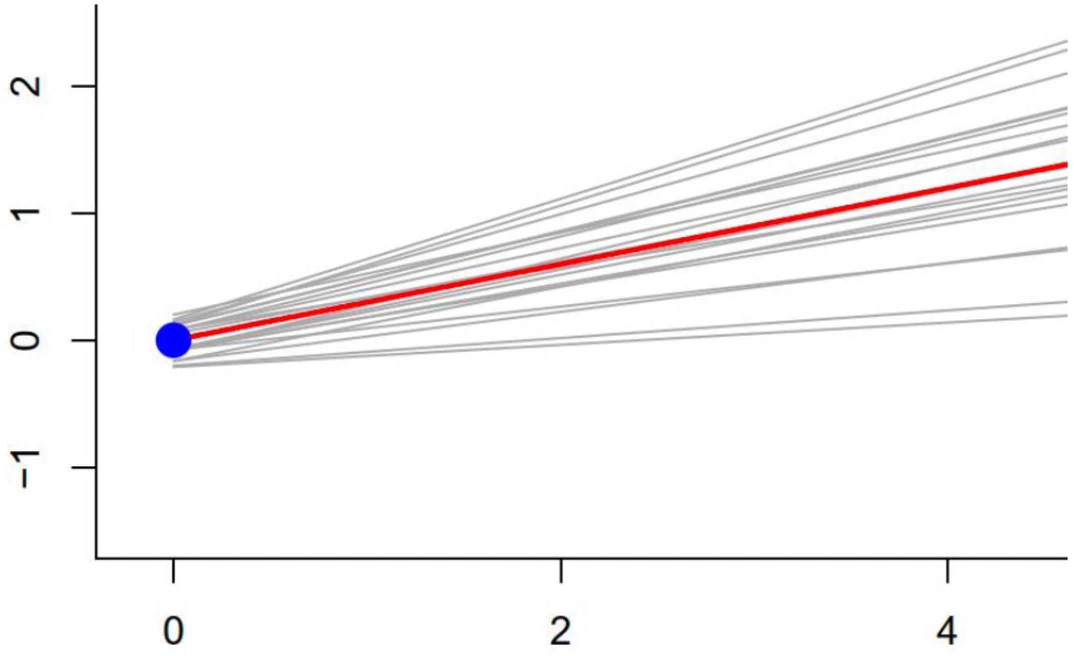
- There are different types of dicrete factors.
- How exacly should the factors be modeled?
  - Fixed-effects: group differences, experimental manipulations etc.
  - Random-effects: group relations, e.g., individuals, items,
- Fixed-effects are as in the linear regression.
  - Y = Intercept + Slope * X + Error
- Dummy/Effect/Contrast-coding for discrete factors.
  - Mean differences between the groups are estimated as fixed effects.

Outline

- Hierarchical Data
- Mixed Models
  - Basic Idea
  - Model Types
  - Intraclass Correlation
- Model Comparison
- Mixed Models in Action
- Practice (Analyzing Reaction Times)

Model Types

- Random intercept, random slope model
- Both intercept and slopes vary by random-effect factors.



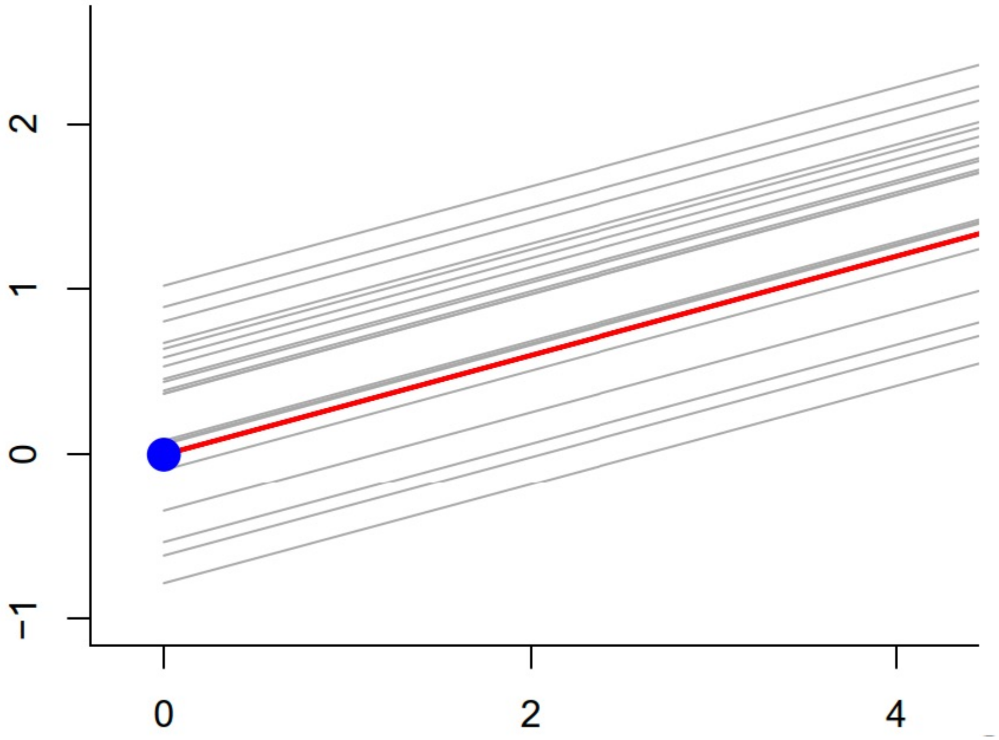
Level 1: $$Y_{i,j} = \beta_{0,j} + \beta_{1,j} \cdot X_{i,j} + e_{i,j}$$

Level 2: $$\beta_{0,j} = \gamma_{0(\beta_0)} + u_{j(\beta_0)}$$

$$\beta_{1,j} = \gamma_{0(\beta_1)} + u_{j(\beta_1)}$$

Model Types

- Random intercept, fixed slope
- Intercept varies by random-effect factors, but slope is the same for each factor.



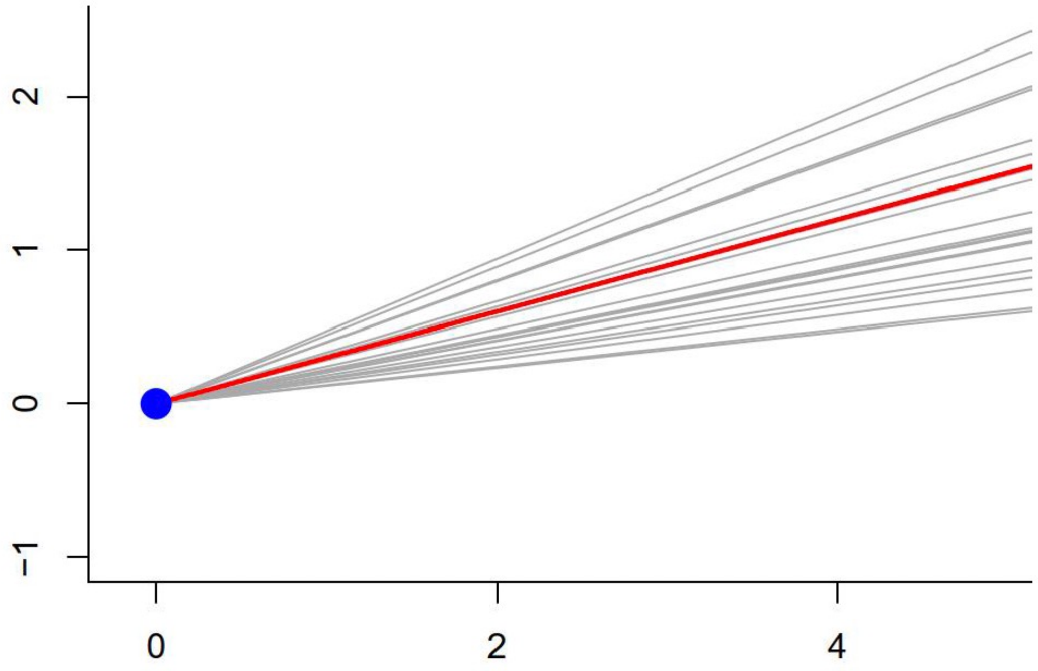
Level 1: $$Y_{i,j} = \beta_{0,j} + \beta_{1,j} \cdot X_{i,j} + e_{i,j}$$

Level 2: $$\beta_{0,j} = \gamma_{0(\beta_0)} + u_{j(\beta_0)}$$

$$\beta_{1,j} = \gamma_{0(\beta_1)}$$

Model Types

- Fixed intercept, random slope (make no sense!)
- Intercept is fixed for random-effect factor, but slope varies for each.



Level 1: $$Y_{i,j} = \beta_{0,j} + \beta_{1,j} \cdot X_{i,j} + e_{i,j}$$

Level 2: $$\beta_{0,j} = \gamma_{0(\beta_0)}$$

$$\beta_{1,j} = \gamma_{0(\beta_1)} + u_{j(\beta_1)}$$

Model Types

- Mixing fixed and random components

Level 1: $$Y_{i,j} = \beta_{0,j} + \beta_{1,j} \cdot X_{1,i,j} + \beta_{2,j} \cdot X_{2,i,j} + e_{i,j}$$

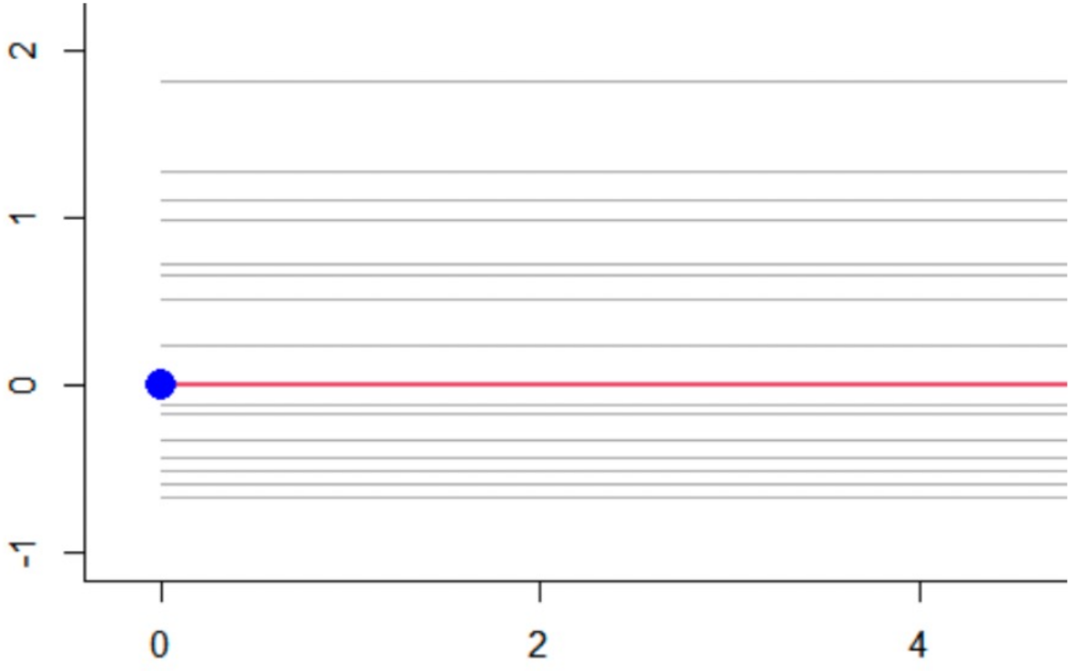
Level 2: $$\beta_{0,j} = \gamma_{0(\beta_0)} + u_{j(\beta_0)}$$

$$\beta_{1,j} = \gamma_{0(\beta_1)} + u_{j(\beta_1)}$$

$$\beta_{2,j} = \gamma_{0(\beta_2)}$$

Model Types

- Intercept only (Null-model)
- No predictor!
- Used to compare if a model with a predictor is any better.



Level 1: $Y_{i,j} = \beta_{0,j} + e_{i,j}$

Level 2: $\beta_{0,j} = \gamma_{0(\beta_0)} + u_{j(\beta_0)}$

Einsetzen: $Y_{i,j} = \gamma_{0(\beta_0)} + u_{j(\beta_0)} + e_{i,j}$

Which model to build?

- Substential question is which fixed- and random-effects will be included in the model?
- Fixed-effects are easy!
- There are two approaches for random-effects:
  - From from basic to complex (Baguley, 2012)
  - From complex to basic (Barr et al., 2013)
    - Problem: maximal model does not always converge and not easily interpretable (Bates et al, 2018).

Outline

- Hierarchical Data
- Mixed Models
  - Basic Idea
  - Model Types
  - Intraclass Correlation
- Model Comparison
- Mixed Models in Action
- Practice (Analyzing Reaction Times)

Intraclass Correlation (ICC)

- What proportion of the variance is due to group differences?
- How similar are the units within the same group?
- Expected correlation between two observation from the same unit (e.g., two measurements from one participant)

$$\mathrm{ICC} = \rho = \frac{\sigma^2_{u(\beta_0)}}{\sigma^2_{u(\beta_0)} + \sigma^2_{\epsilon}}$$

Intraclass Correlation (ICC)

- If ICC = 0 then no group differences
- If ICC > 0 then evidence for systematic group differences (dependency between the observations)

- Rule of Thumb: ICC > 0.05 requires Mixed Model!

Shrinkage

- https://m-clark.github.io/posts/2019-05-14-shrinkage-in-mixed-models/

Outline

- Hierarchical Data
- Mixed Models
    - Basic Idea
    - Model Types
    - Intraclass Correlation
- Model Comparison
- Mixed Models in Action
- Practice (Analyzing Reaction Times)

Model Comparison

- Model estimation provides Deviance = -2 log(Likelihood)
- Decide whether one model is more likely than the other.

$$\Delta_{\text{Devianz}} = 2 \cdot [\log(\text{Likelihood}_{M_1}) - \log(\text{Likelihood}_{M_0})]$$

- Evaluation of the model fit relative to the number of parameters k.

Akaike Information Criterion: $\text{AIC} = -2 \cdot \log(\text{Likelihood}) + 2k$

Bayesian Information Criterion: $\text{BIC} = -2 \cdot \log(\text{Likelihood}) + \log(N)k$

Sample Size

- The more observation, the better the model fit.
- Sufficiently large sample necessary for stable parameter estimates (especially of variances)

- Random-effects: 30-50 observation pro factor, preferably 100. (e.g., 40 measurements per participant)
- Fixed-effects: approximately 20 units per factor group (e.g., 20 participant per each level of experimental manipulation)

Outline

- Hierarchical Data
- Mixed Models
  - Basic Idea
  - Model Types
  - Intraclass Correlation
- Model Comparison
- Mixed Models in Action
- Practice (Analyzing Reaction Times)

Power Calculation

- Smallest Effect Size Of Interest (SESOI)
- Simulation-based power calculation