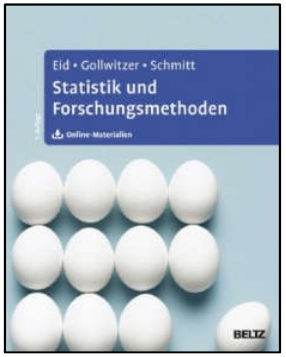
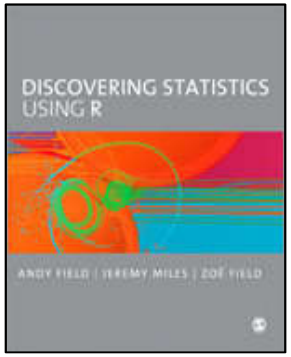
Literatur

Einführende Literatur

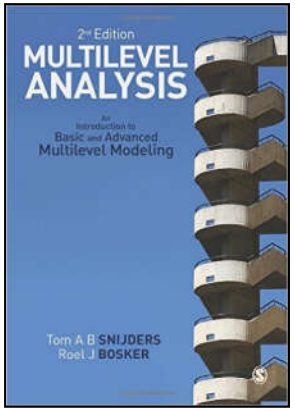
- Eid, M., Gollwitzer, M., & Schmitt, M. (2017). *Statistik und Forschungsmethoden* (5. Aufl.). Basel: Beltz. [Kap. 20]
- Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R.* Thousand Oaks: Sage. [Kap. 19]

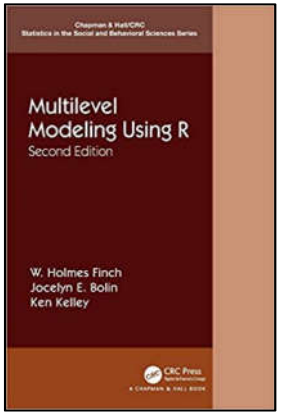




Vertiefende Literatur

- Finch, W. H., Bolin, J. E., & Kelley, K. (2019). *Multilevel Modeling Using R* (2. Aufl.). CRC Press.
- Snijders, T. A., & Bosker, R. J. (2012). *Multilevel Analysis* (2. Aufl.). Los Angeles: Sage.





Ausblick

1. **Hierarchische Datenstrukturen**
2. Modell der hierarchischen Regression
3. Forschungsfragen
4. Inferenzstatistische Tests
5. Anwendungsbeispiel

Anhänge

- Beispielanalysen in R
- R-Funktionen

Struktur hierarchischer Daten

- **Studiendesign** in psychologischen Erhebungen

  Es ist nicht praktikabel alle Personen zu testen (Vollerhebung).

  Es reicht eine Auswahl an Personen (Stichprobenziehung*)*

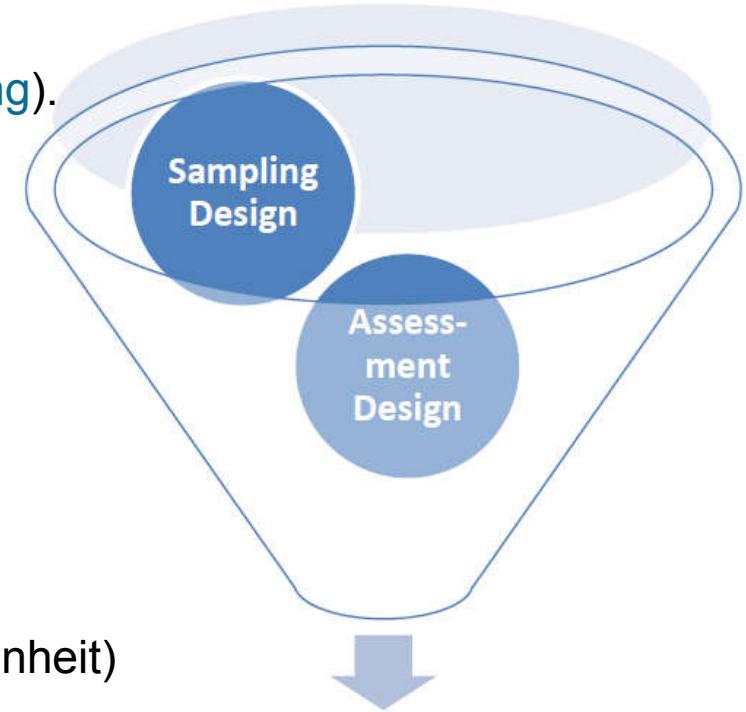
- Einfache Zufallsstichproben sind oft nicht zu realisieren (z.B. Zufallsauswahl von Schülern in Deutschland), da …

  … zu teuer, organisatorisch unpraktisch, schwierig individuelle Informationen von Person (= Testleistung) mit Kontextinformation (z.B. über Schule) zu verknüpfen usw.

- Mehrstufige Zufallsstichproben (zweistufige Auswahl):

  1. Stufe: Zufallsauswahl von Schulen (= erste Stichprobeneinheit)

  2. Stufe: Zufällige Auswahl Klassen innerhalb von Schulen (= zweite Stichprobeneinheit)



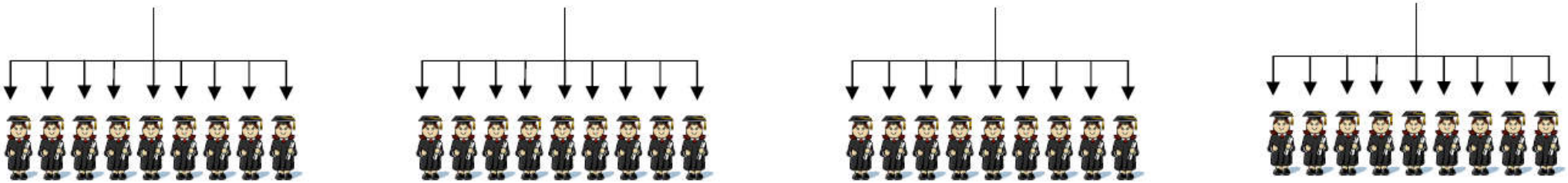
Bei mehrstufigen Zufallsstichproben sind Beobachtungen innerhalb eines Clusters / Gruppe / Klasse nicht mehr unabhängig von einander, sondern werden durch den **gemeinsamen Kontext** beeinflusst.

Beispiel: Schüler in Schulklassen

Stichprobendesign:
2. Schulklasse
1. Schüler

Ebene 2: Schulklasse

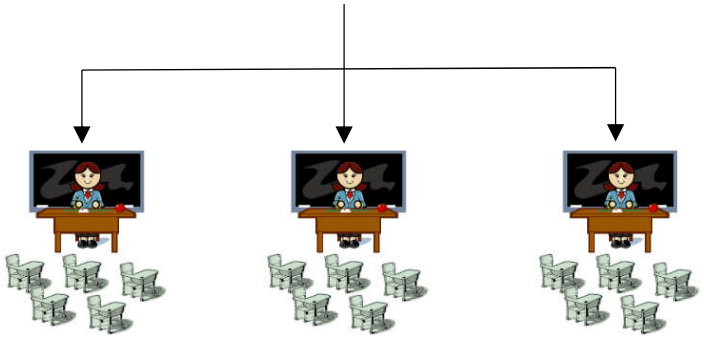
Ebene 1: Schüler



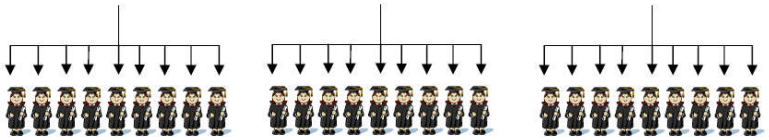
Beispiel: Schüler in Klassen in Schulen

Ebene 3: Schule

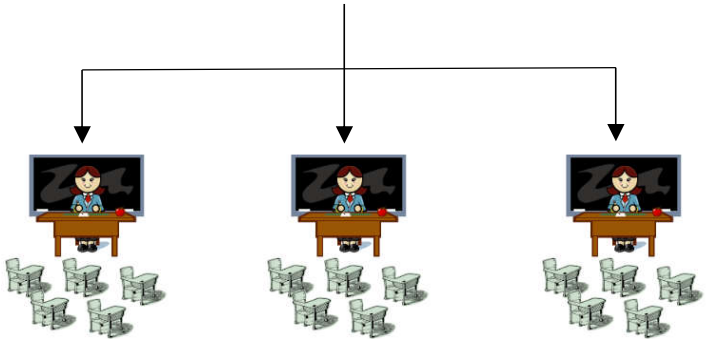
Ebene 2: Schulklasse

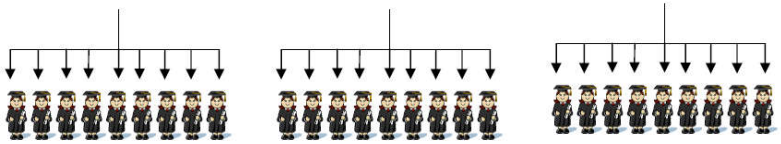


Ebene 1: Schüler



Stichprobendesign:
3. Schule
2. Schulklasse
1. Schüler



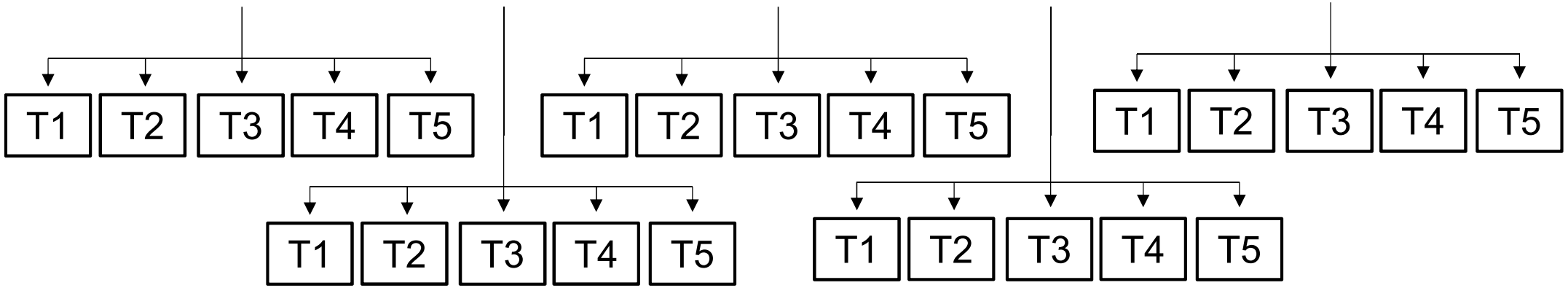


Beispiel: Messzeitpunkte in Personen

Stichprobendesign:
2. Person
1. Messzeitpunkt

Ebene 2: Personen

Ebene 1: Messzeitpunkt



Hierarchische Strukturen

- **Hierarchische Strukturen** sind weit verbreitet in unserem Alltag und daher auch in der psychologischen Forschung:

  Viele gesellschaftliche Phänomene können als ein System mit unterschiedlichen Ebenen beschrieben werden.

  Diese Ebenen (z.B. Schüler – Schulklasse - Schule) weisen eine hierarchische Ordnung auf (d.h. man kann zwischen Makro- und Mikroebenen unterscheiden).

  In diesem System sind die Untersuchungsobjekte / -einheiten auf diesen Ebenen angeordnet.

Man sagt …

Objekte / Einheiten einer niedrigeren Ebene (z.B. Personen) sind innerhalb von Objekten / Einheiten einer höheren Ebene (z.B. Schulklassen) **verschachtelt** (engl.: *nested*).

oder:

Man sagt …

Objekte / Einheiten einer niedrigeren Ebene sind Elemente von Objekten / Einheiten einer höheren Ebene (z.B. Schüler in Schulklassen).

Variablen in hierarchischen Strukturen

- Die **abhängige Variable** (Kriterium) wird auf der niedrigsten Ebene (Querschnitt: meist Personen, Längsschnitt: meist Messzeitpunkte) erfasst (z.B. Schulleistung, subjektives Wohlbefinden, Arbeitszufriedenheit)
- Die **unabhängigen Variablen** (Prädiktor) können sowohl auf der niedrigsten Ebene wie auch höheren Ebenen (Querschnitt: z.B. Schulklasse, Firma, Krankenhaus; Längsschnitt: meist Personen) gemessen werden, z.B.:

	Querschnitt	Längsschnitt
Ebene 1:	Leistungsmotivation auf Personenebene	Leistungsmotivation auf Zeitebene
Ebene 2:	Lehrmethode auf Klassenebene	Geschlecht auf Personenebene

- Die Forschungsfragen beziehen sich auf Unterschiede und Ähnlichkeiten von Objekten der niedrigsten Ebene (z.B. Schüler) und Objekten der höheren Ebene (z.B. Schulklassen).
- Problem: Beobachtungen auf der niedrigsten Eben werden durch den **gemeinsamen Kontext** (z.B. Klassenmerkmale) beeinflusst.

Beispiele hierarchischer Daten

1. Auswahl von Schulen zur Untersuchung neuer Lehrmethoden: alle Schüler innerhalb einer Schule werden derselben Treatment-Bedingung angehören

Schüler in Klassen

2. Arbeitszufriedenheit von Mitarbeitern verschiedener Unternehmensabteilungen

Arbeitnehmer in Arbeitsbereichen

3. Qualität von Therapiesitzungen bewertet durch Patienten/Patientinnen verschiedener Therapeuten (einer Therapieeinrichtung)

Individuelle Sitzungen in Patienten in Therapeuten

Folgen hierarchischer Datenstrukturen

- Hierarchische Datenstrukturen führen dazu, dass die **Varianz der abhängigen Variable**

  durch Unterschiede der Schüler (Ebene 1) und

  durch Unterschiede der Schulklassen (Ebene 2) beeinflusst werden.

- Diese Abhängigkeit zwischen einzelnen Untersuchungsobjekten / -einheiten muss bei der statistischen Auswertung berücksichtigt werden:

Hierarchische lineare Modelle (= Mehrebenanalysen)

- Das **Ignorieren** von Mehrebenenstrukturen führt zum Risiko falscher Schlüsse bei …

  der Interpretation von Zusammenhangs- und Beeinflussungsstrukturen sowie

  der inferenzstatistischen Absicherung von Regressionsgewichten.

Verzerrte Interpretationen: Disaggregation

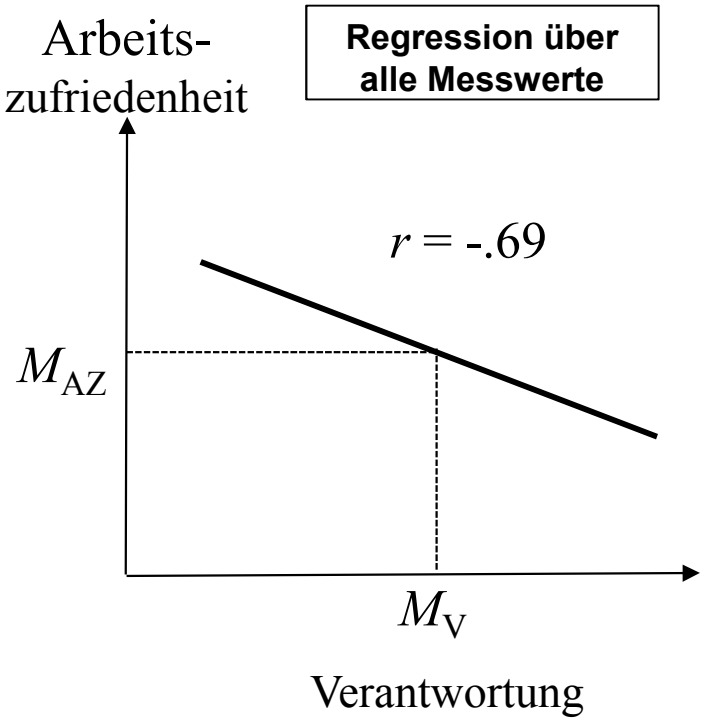
- Die Zusammenhänge zwischen den Variablen werden ausschließlich auf der Personenebene (Ebene 1) getestet ohne die unterschiedlichen Gruppen (Ebene 2) zu berücksichtigen (z.B. mit multipler OLS Regression):
- Es wird angenommen, dass alle Gruppen homogen sind und daher die Effekte auf der niedrigeren Ebene nicht beeinflussen.
- **Nachteil**:

  Mögliche Unterschiede zwischen Gruppen werden nicht berücksichtigt

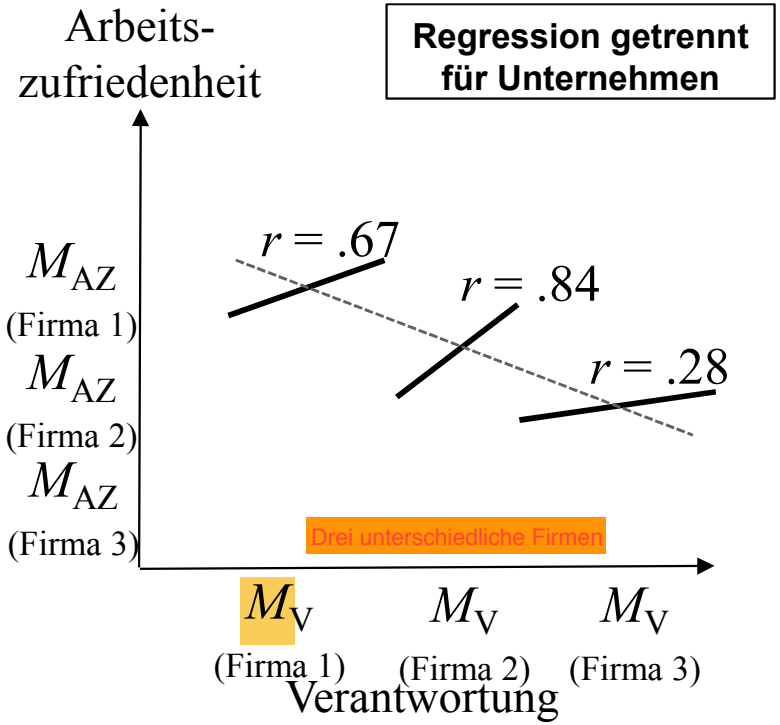
  Die Regressionskoeffizienten stellen eine Mischung aus Effekten innerhalb der Gruppen (z.B. Schüler) und zwischen den Gruppen (z.B. Schulklassen) dar und können zu fehlerhaften Interpretationen führen.

Verzerrte Interpretationen: Disaggregation

Beispiel: Wie ist der Zusammenhang zwischen dem Ausmaß an Verantwortung und der Arbeitszufriedenheit der Mitarbeiter in drei Unternehmen?



hier wäre die Conclusion, gebt den Leuten nicht zuviel Verantwortung



Hier dann plötzlich positiver ZH zwischen Verantwortung und Arbeitszufriedenheit

- Der Zusammenhang zwischen zwei Variablen ändert sich, wenn man für eine Drittvariable kontrolliert (**Simpson-Paradox**).

Offensichtlich unterscheiden sich die Firmen hier sehr stark in ihrer durchschnittlichen Arbeitszufiedenheit

- Mögliche Drittvariable auf Firmenebne: wirtschaftlicher Erfolg

Verzerrte Interpretationen: Aggregation

- Verzerrte Schlüsse können auch erfolgen, wenn die Daten auf der höheren Ebene zusammengefasst werden; d.h. es wird der Mittelwert der abhängigen Variable innerhalb jeder Gruppe gebildet.
- **Nachteile**:

  Die Varianz innerhalb der Gruppen wird ignoriert und es wird angenommen, dass die Einheiten auf Ebene 1 (z.B. Mitarbeiter) homogen sind.

  Unterschiedliche Stichprobengrößen der einzelnen Gruppen werden ignoriert, weshalb Variable unterschiedliche genau erfasst werden (größere Gruppen haben i.R. auch reliablere Variablen als kleinere Gruppen).

Wichtig

Ein **ökologischer Fehlschluss** (»ecological fallacy«) liegt vor, wenn eine Effekt, der auf Gruppenebene (Ebene 2) gefunden wurde, so interpretiert wird, als würde er auf Personenebene (Ebene 1) zutreffen.

Beispiel: Auf der Ebene der Firmen gibt es einen negativen Zusammenhang zwischen Verantwortung und Arbeitszufriedenheit, während er auf Mitarbeiterebne positiv ist.

Verzerrte Inferenzstatistik

- Eine Grundannahme der multiplen Regression ist Unkorreliertheit der Residuen.
- Eine **Verletzung dieser Annahme** führt zu einer …

  Unterschätzung der Standardfehler,

  erhöhten Wahrscheinlichkeit signifikanter Ergebnisse, da die Standardfehler zu klein sind, und

  $$t = \frac{b_j}{\sigma_{b_j}} \qquad \text{mit} \quad df = n - m - 1$$

  Zunahmen der Fehler 1. Art (eigentlich richtige Nullhypothesen werden verworfen).

Beispiel: Mitarbeiter der gleichen Firma sind sich wahrscheinlich nicht nur hinsichtlich ihrer Verantwortung ähnlicher, sondern auch auf anderen nicht beobachteten Variablen (z.B. Arbeitsklima), welche die Arbeitszufriedenheit beeinflussen. Ändert sich nun das Arbeitsklima in einer Firma, wirkt sich das nicht nur auf einen sondern alle Mitarbeiter aus.

Je größer die Unterschiede auf dem gemessenen Merkmal **zwischen** den Gruppen auf Ebene 2 (z.B. Firmen, Schulen) im Vergleich zu den Unterschieden **innerhalb** dieser Gruppen (z.B. Mitarbeiter in Firmen, Schüler in Schulen) sind, desto stärker ist die Unabhängigkeitsannahme verletzt.

Verzerrte Inferenzstatistik

Setzt die Varianz auf höheren Ebene ins Verhältnis zur Gesamtvarianz

Intra-Klassen-Korrelation (intra-class correlation, ICC):

$$ICC = \rho = \frac{Var_{\text{Ebene 2}}}{Var_{\text{Gesamt}}} = \frac{Var_{\text{Ebene 2}}}{Var_{\text{Ebene 1}} + Var_{\text{Ebene 2}}}$$

Werte zwischen 0 und 1. 0= es gibt keine Abhängigkeiten

- ρ (rho) gibt den Anteil der Varianz zwischen den Gruppen an der Gesamtvarianz an.
- Alternative Interpretation: ρ ist die Korrelation der abhängigen Variablen zweier Individuen, die zufällig aus derselben Gruppe gezogen wurden.
- Nimmt Werte zwischen 0 und 1 an.
- Ist ρ gering (z.B. < .05), dann liegt nur wenig Ebene 2-Varianz vor und hierarchische Strukturen verzerren Inferenzstatistiken nur unwesentlich.
- Ist ρ gleich 1, dann liegt nur Ebene 2-Varianz vor und Personen innerhalb der Gruppen unterscheiden sich nicht (Aggregation ist zulässig).

Hierarchische lineare Modelle

Ziel: Untersuchung von **individuellen Effekten** (Ebene 1) und **Kontexteffekten** (Ebene 2) sowie deren Interaktion.

Hierarchische lineare Modelle =

Mehrebenenanalyse =

Random effects Modelle =

Mixed effects Modelle

Ausblick

1. Hierarchische Datenstrukturen
2. **Modell der hierarchischen Regression**
3. Forschungsfragen
4. Inferenzstatistische Tests
5. Anwendungsbeispiel

Anhänge

- Beispielanalysen in R
- R-Funktionen

Bivariate lineare Regression

- Zur Wiederholung das Modell der einfachen linearen Regression zur Vorhersage <u>einer</u> abhängigen Variable durch <u>eine</u> unabhängige Variable:

$$y_i = \beta_0 + \beta_1 \quad x_i + \varepsilon_i$$

y_i = abhängige Variable für Person i

x_i = unabhängige Variable für Person i

β_0 = Intercept des Modells

β_1 = Steigungsparameter (Slope)

ε_i = Residuum

- Es wird angenommen, dass für alle Personen der gleiche Intercept und Slope zutreffend ist.
- Wenn jedoch Personen in verschiedenen Kontexten (z.B. Schulklassen, Unternehmen) untersucht werden, ist es wahrscheinlich, dass unterschiedliche Intercepts und Slopes für jede Gruppe gelten.
- Nur wenn es keinen Effekte der Gruppe (Ebene 2) aufgrund der hierarchischen Datenstruktur gibt, dann sind alle Intercepts und Slopes für alle Gruppen gleich.

Hierarchische bivariate lineare Regression

Bivariate Regression:

$$y_i = \beta_0 + \beta_1 \quad x_i + \varepsilon_i$$

y_i = abhängige Variable für Person i

x_i = unabhängige Variable für Person i

β_0 = Intercept des Modells

β_1 = Steigungsparameter (Slope)

ε_i = Residuum für Person i

Hierarchische Regression:

$$y_{ij} = \beta_{0j} + \beta_{1j} \quad x_{ij} + \varepsilon_{ij}$$

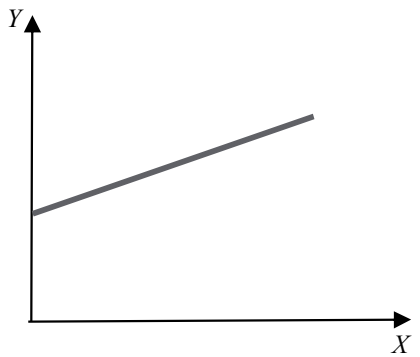
y_i = abhängige Variable für Person i in Gruppe j

x_i = unabhängige Variable für Person i in Gruppe j

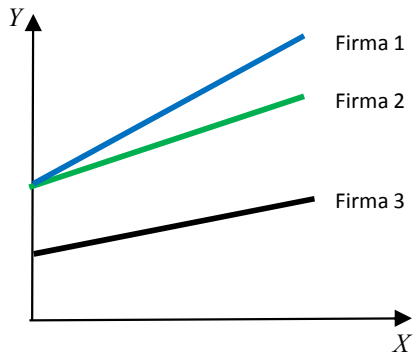
β_{0j} = Intercept des Modells in Gruppe j

β_{1j} = Steigungsparameter (Slope) in Gruppe j

ε_{ij} = Residuum für Person i in Gruppe j



Es wird ein Intercept und Slope für alle Gruppen angenommen!



Es werden gruppenspezifische Intercepts und Slopes angenommen!

Random Intercept

- Indem der Intercept über die Gruppen variieren kann, entsteht ein **random Intercept**.

Hierarchische Regression:

$$y_{ij} = \beta_{0j} + \beta_{1j} \quad x_{ij} + \varepsilon_{ij}$$

Random Intercept:

$$\beta_{0j} = \boxed{\gamma_{00}} + \boxed{\upsilon_{0j}}$$

fixed effect

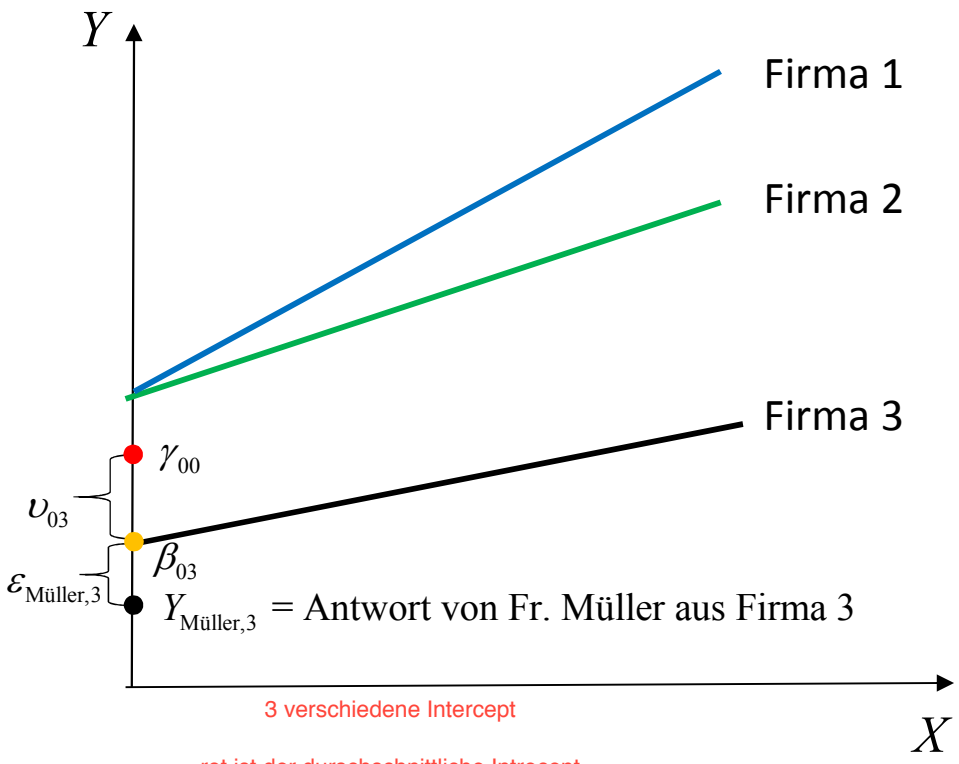
random effect

zufälliger Effekt weil er von Firma zu Firma anders sein kann

β_{0j} = Intercept des Modells in Gruppe j

γ_{00} = Gesamtintercept (mittlerer Intercept über alle Gruppe)

υ_{0j} = Abweichung des Intercepts in Gruppe j vom Gesamtintercept



3 verschiedene Intercept

rot ist der durschschnittliche Intrecept

Gelb die Abweichung vom Durchschnitt

Annahme: Die Intercept unterscheiden sich zufällig von den Gruppen

Random Slope

- Indem der Slope über die Gruppen variieren kann, entsteht ein **random Slope**.

Hierarchische Regression:

$$y_{ij} = \beta_{0j} + \beta_{1j} \quad x_{ij} + \varepsilon_{ij}$$

Random Slope:

$$\beta_{1j} = \boxed{\gamma_{10}} + \boxed{\upsilon_{1j}}$$

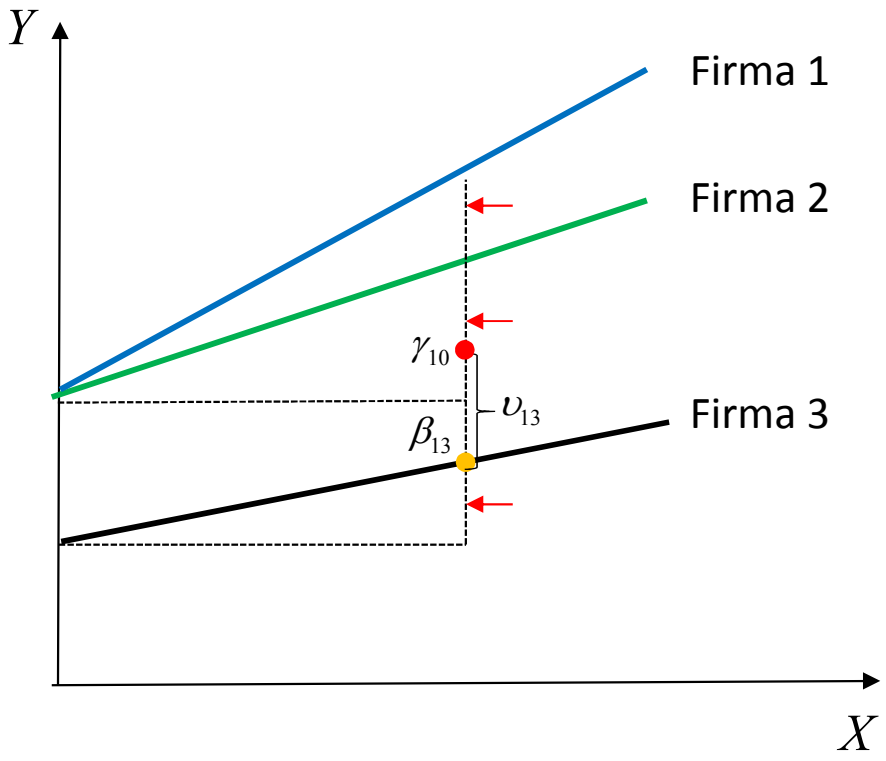
fixed effect

random effect

β_{1j} = Slope des Modells in Gruppe j

γ_{10} = Gesamtslope (mittlerer Slope über alle Gruppe)

υ_{1j} = Abweichung des Slopes in Gruppe j vom Gesamtslope



Hierarchisches Modell

- Die Kombination der Regressionsgleichungen auf Ebene 1 und 2 führt zum hierarchischen Regressionsmodell:

Ebene 1:

$$y_{ij} = \beta_{0j} + \beta_{1j} \quad x_{ij} + \varepsilon_{ij}$$

Ebene 2:

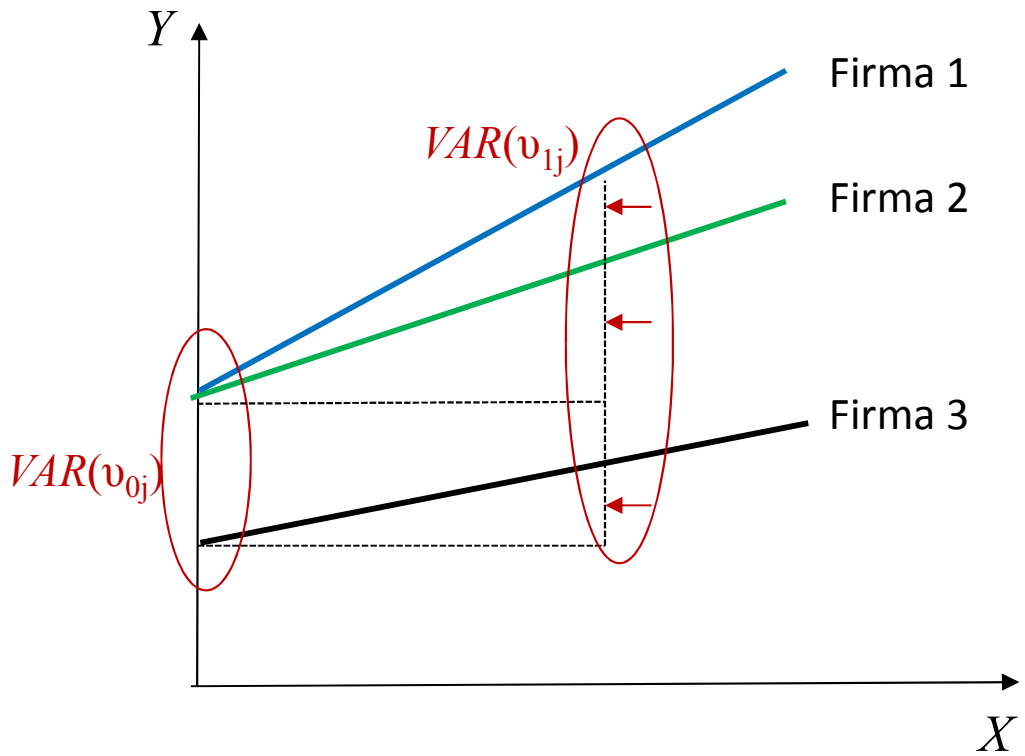
$$\beta_{0j} = \gamma_{00} + \upsilon_{0j} \quad \text{(random intercept)}$$

$$\beta_{1j} = \gamma_{10} + \upsilon_{1j} \quad \text{(random slope)}$$

Ebene 1 + 2:

$$y_{ij} = (\gamma_{00} + \boxed{\upsilon_{0j}}) + (\gamma_{10} + \boxed{\upsilon_{1j}}) \quad x_{ij} + \varepsilon_{ij}$$

Vergleichbare Interpretation wie bei der einfachen bivariaten Regression



Zusätzliche Fehlerterme für unterschiedliche Gruppen

Ausblick

1. Hierarchische Datenstrukturen
2. Modell der hierarchischen Regression
3. **Forschungsfragen**
4. Inferenzstatistische Tests
5. Anwendungsbeispiel

Anhänge

- Beispielanalysen in R
- R-Funktionen

1. Varianz innerhalb und zwischen Kontexten

- **Fragestellung**:

  Können individuelle Unterschiede sowohl durch Unterschiede zwischen Personen und auch Unterschiede zwischen Kontexten, in denen sich die Personen befinden, erklärt werden?

Zur Erinnerung:

Varianzaufteilung:

Varianz *innerhalb* + Varianz *zwischen* Gruppen

$$ICC = \rho = \frac{Var_{\text{Ebene 2}}}{Var_{\text{Ebene 1}} + Var_{\text{Ebene 2}}} = \frac{Var(\upsilon_0)}{Var(\varepsilon) + Var(\upsilon_0)}$$

Intercept-Only Modell

Ebene 1:

$$y_{ij} = \beta_{0j} + \varepsilon_{ij}$$

Ebene 2:

$$\beta_{0j} = \gamma_{00} + \upsilon_{0j}$$ (random intercept)

Ebene 1 + 2:

$$y_{ij} = \gamma_{00} + \upsilon_{0j} + \varepsilon_{ij}$$

Drückt die Varianz der zufälligen Ebenen. Varianz des Zufallseffekts

Beispiel: Wieviel Prozent der Varianz in der Arbeitszufriedenheit kann durch Unterschiede der Mitarbeiter erklärt werden und wieviel Prozent sind durch unterschiedliche Firmenkontexte erklärbar?

2. Effekte auf der Personenebene

- **Fragestellung**:

  Welche Personenmerkmale (Ebene 1) hängen signifikant mit der abhängigen Variable zusammen?

  Wieviel Varianz erklären Personenvariablen in der abhängigen Variable?

Beispiel: Wieviel Prozent der Varianz in der Arbeitszufriedenheit kann durch die wahrgenommene Verantwortung der Mitarbeiter erklärt werden?

Random Intercept Modell

Ebene 1:

$$y_{ij} = \beta_{0j} + \beta_{1j} \quad x_{ij} + \varepsilon_{ij}$$

Ebene 2:

$$\beta_{0j} = \gamma_{00} + \upsilon_{0j}$$ (random intercept)

$$\beta_{1j} = \gamma_{10}$$ (fixed slope)

Ebene 1 + 2:

$$y_{ij} = \gamma_{00} + \upsilon_{0j} + \gamma_{10} \quad x_{ij} + \varepsilon_{ij}$$

3. Effekte auf der Gruppenebene

- **Kompositionseffekte**:

  Welche Merkmale der Personenstichprobe (-komposition) hängen mit der abhängigen Variable zusammen?

  Es gibt einen äquivalenten Effekt auf der Personenebene.

- **Kontexteffekte**:

  Welche (strukturellen) Merkmale der Situation / des Kontexts hängen mit der abhängigen Variable zusammen?

  Es gibt keinen äquivalenten Effekt auf der Personenebene

Beispiel: Erklärt die durchschnittliche Verantwortung in einer Firma (Kompositionseffekt) oder der wirtschaftliche Erfolg der Firma (Kontexteffekt) die Arbeitszufriedenheit der Mitarbeiter über die individuelle Verantwortung hinaus?

Random Intercept Modell

Ebene 1:

$$y_{ij} = \beta_{0j} + \beta_{1j} \quad x_{ij} + \varepsilon_{ij}$$

Ebene 2:

Prädiktor w auf Gruppenebene

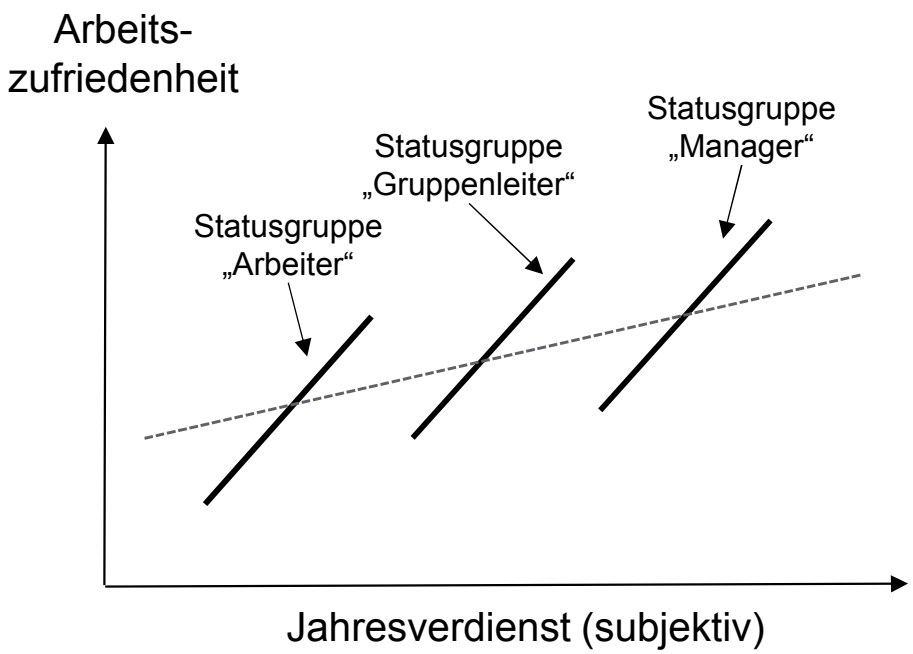
$$\beta_{0j} = \gamma_{00} + \gamma_{01} \quad w_{1j} + \upsilon_{0j} \quad \text{(random intercept)}$$

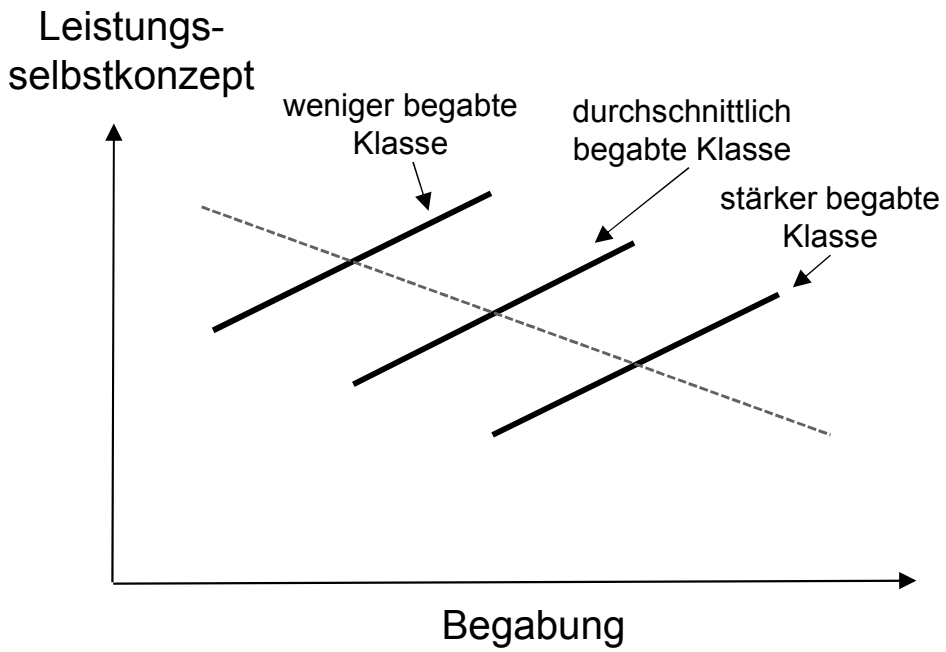
$$\beta_{1j} = \gamma_{10} \quad \text{(fixed slope)}$$

Ebene 1 + 2:

$$y_{ij} = \gamma_{00} + \gamma_{01} \quad w_{1j} + \upsilon_{0j} + \gamma_{10} \quad x_{ij} + \varepsilon_{ij}$$

Beispiele für Kompositionseffekte





Klassen die im Durschnitt besser sind, haben ein schlechteres Selbstkonzept. Sind alle gut, fühle ich mich schlechter als in einem Umfeld mit durchschnittlichen Leistungen

- Der Innerhalb-Gruppen-Effekt (Jahresverdienst innerhalb einer Statusgruppe) auf die Arbeitszufriedenheit ist stärker als der zwischen-Gruppen-Effekt (durchschnittlicher Jahresverdienst einer Statusgruppe).

- Der Innerhalb-Klassen-Effekt (Begabung innerhalb einer Schulklasse) auf das Leistungsselbstkonzept ist positiv; der Zwischen-Klassen-Effekt (durchschnittliche Begabung einer Klasse) auf das Leistungsselbstkonzept ist hingegen negativ („Big-Fish-Little-Pond-Effekt“, Marsh et al., 2001).

4. Unterschiedliche Personeneffekte

- **Fragestellung**:

  Hängt der Zusammenhang zwischen der abhängigen Variable und Personenmerkmalen auf Ebene 1 von einem Prädiktor auf Gruppenebne ab?

> *Beispiel*: Ist der Zusammenhang zwischen Arbeitszufriedenheit und der wahrgenommenen Verantwortung der Mitarbeiter in allen Firmen gleich?

- Die Intercepts und Slopes auf Ebene 2 sind nicht unabhängig von einander sondern können kovariieren, d.h. $Cov(\upsilon_0, \upsilon_1) \neq 0$.
- Die Varianzen und Kovarianzen der Ebene 2 werden daher meist in Form einer Matrix dargestellt:

$$\Sigma_\upsilon = \begin{bmatrix} \sigma^2_{\upsilon_0} & \sigma^2_{\upsilon_0,\upsilon_1} \\ \sigma^2_{\upsilon_1,\upsilon_0} & \sigma^2_{\upsilon_1} \end{bmatrix}$$

$COV(\upsilon_{0j}, \upsilon_{1j}) = COV(\upsilon_{1j}, \upsilon_{0j})$

$VAR(\upsilon_{0j})$ = Varianz des random Intercept

$VAR(\upsilon_{1j})$ = Varianz des random Slope

Random Coefficient Modell

Ebene 1:

Regressionsgewicht

$$y_{ij} = \beta_{0j} + \beta_{1j} \; x_{ij} + \varepsilon_{ij}$$

Ebene 2:

$$\beta_{0j} = \gamma_{00} + \upsilon_{0j} \quad \text{(random intercept)}$$

$$\beta_{1j} = \gamma_{10} + \upsilon_{1j} \quad \text{(random slope)}$$

Ebene 1 + 2:

Durschn. Regressionsgewicht über alle Gruppen hinweg

$$y_{ij} = \gamma_{00} + \upsilon_{0j} + (\gamma_{10} + \upsilon_{1j}) \; x_{ij} + \varepsilon_{ij}$$

Durchschnittliche Effekt über alle Gruppen hinweg

Zufallsvarianz

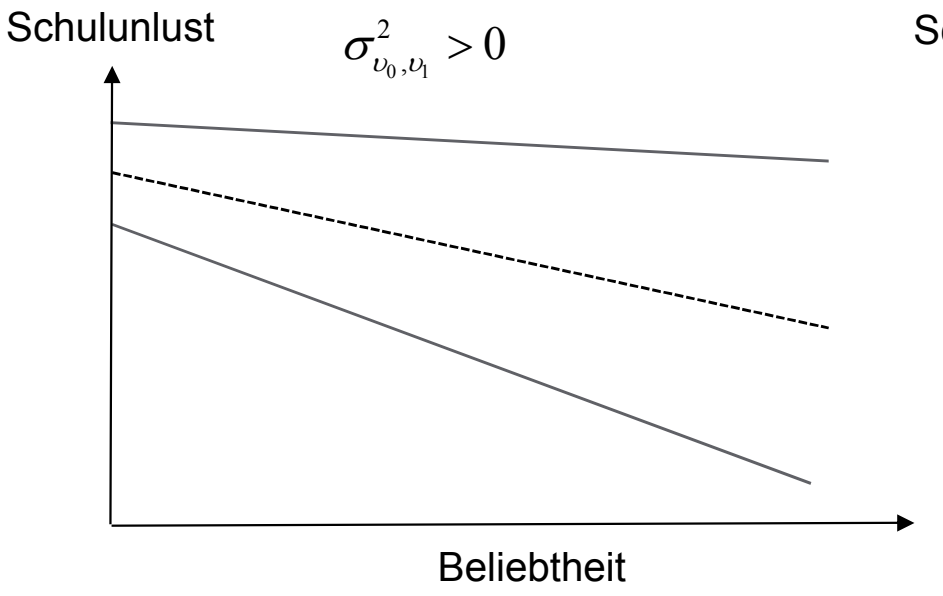
Zufallsvarianz

Fehlervarianz

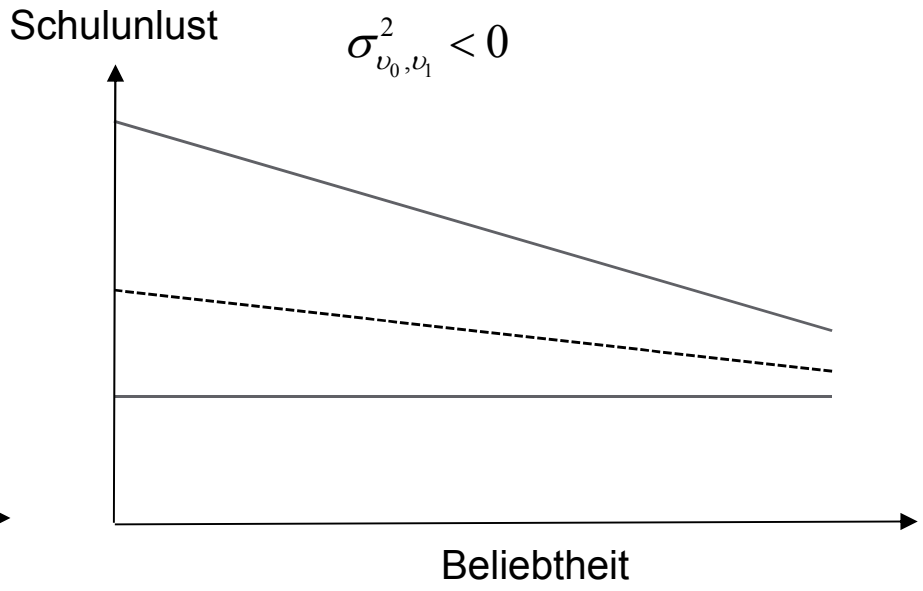
4. Unterschiedliche Personeneffekte

- Beispiele für unterschiedliche Kovarianzen zwischen den zufälligen Effekten:

Positiver Zusammenhang



Negativer Zusammenhang



5. Interaktionseffekte

- **Fragestellung**:

  Variiert der Zusammenhang zwischen der abhängigen Variable und Personenmerkmalen auf Ebene 1 zwischen verschiedenen Kontexten?

Beispiel: Wird der Zusammenhang zwischen Arbeitszufriedenheit und der wahrgenommenen Verantwortung der Mitarbeiter durch den wirtschaftlichen Erfolg der Firma beeinflusst?

Cross-level Interaktion

Ebene 1:

$$y_{ij} = \beta_{0j} + \beta_{1j} \quad x_{ij} + \varepsilon_{ij}$$

Ebene 2:

Prädiktor w auf Gruppenebene

$$\beta_{0j} = \gamma_{00} + \upsilon_{0j} \qquad \text{(random intercept)}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} \quad w_{1j} + \upsilon_{1j} \qquad \text{(random slope)}$$

Ebene 1 + 2:

$$y_{ij} = \gamma_{00} + \upsilon_{0j} + (\gamma_{10} + \gamma_{11} \quad w_{1j} + \upsilon_{1j}) \quad x_{ij} + \varepsilon_{ij}$$

Unterschiedliche Modelle

Analyseziel	Modell
Random Intercept Modelle	
1. Varianzaufteilung	Intercept-Only Modell
2. Effekte auf der Personenebene	Random Intercept mit Prädiktoren auf Ebene 1
3. Effekte auf der Gruppenebene	Random Intercept mit Prädiktoren auf Ebene 2
Random Coefficient Modelle	
4. Unterschiedliche Effekte zwischen Gruppen	Random Intercepts und Slopes
5. Interaktion zwischen Gruppen- und Personeneffekten	Random Intercepts und Slopes mit Prädiktoren auf Ebene 1 und 2

Ausblick

1. Hierarchische Datenstrukturen
2. Modell der hierarchischen Regression
3. Forschungsfragen
4. **Inferenzstatistische Tests**
5. Anwendungsbeispiel

Anhänge

- Beispielanalysen in R
- R-Funktionen

Annahmen

- Für inferenzstatistischen Tests müssen in der Population vergleichbare Annahmen / **Voraussetzungen** wie für die einfache lineare Regression gelten:

1. Die abhängige Variable ist kontinuierlich (metrisch)
2. Alle Variablen müssen eine Varianz größer als 0 aufweisen (keine Konstanten).
3. Für alle Personen gilt die gleiche lineare Regressionsgleichung (Linearität).
4. Die Kriteriumswerte (bzw. Residuen) sind statistisch unabhängig voneinander.

   $Cov\left(\varepsilon_i, \varepsilon_j\right) = 0$ für alle Personen i und j
5. Für jede Kombination von Prädiktorwerten weisen die Residuen …

   - den Erwartungswert 0 auf,

   $\mathrm{E}\left(\varepsilon \mid x_1, \ldots x_m\right) = 0$ für alle m Prädiktoren

   - die gleiche Varianz auf (Varianzhomogenität = Homoskedastizität),

   $Var\left(\varepsilon \mid x_1, \ldots x_m\right) = \sigma_\varepsilon^2 > 0$ für alle m Prädiktoren
6. Für jede Kombination von Prädiktorwerten weisen die Residuen …

   - eine Normalverteilung auf.

   $(\varepsilon \mid \mathbf{x}) \sim \mathrm{N}\left(0, \sigma_\varepsilon^2\right)$

$\sim \mathrm{N}(m, v)$ meint „ist normalverteilt mit Mittelwert m und Varianz v“

Annahmen

- Zusätzlich formuliert die hierarchische Regression auch Annahmen / **Voraussetzungen** über die **zufälligen Effekte**:

7. Die zufälligen Effekte weisen für jede Kombination von Prädiktorwerten …

   - einen Erwartungswert von 0 auf.

   $\mathrm{E}(\upsilon \mid x_1, \ldots x_m) = 0$ für alle m Prädiktoren

   - eine multivariate Normalverteilung auf.

   $(\upsilon \mid \mathbf{x}) \sim \mathrm{N}(0, \Sigma_\upsilon)$

   mit

$$\Sigma_\upsilon = \begin{bmatrix} \sigma^2_{\upsilon_0} & \sigma^2_{\upsilon_0,\upsilon_1} \\ \sigma^2_{\upsilon_1,\upsilon_0} & \sigma^2_{\upsilon_1} \end{bmatrix}$$

$VAR(\upsilon_{0j})$ = Varianz des random Intercept

$COV(\upsilon_{0j}, \upsilon_{1j}) = COV(\upsilon_{1j}, \upsilon_{0j})$

$VAR(\upsilon_{1j})$ = Varianz des random Slope

Da die zufälligen Effekte der Ebene 2 korreliert sein können, werden die Varianzen (Diagonale) und Kovarianzen (nicht-Diagonale) der zufälligen Effekte häufig in Form von Matrizen dargestellt.

8. Die Residuen auf Ebene 1 und 2 sind statistisch unabhängig voneinander.

   $Cov(\varepsilon, \upsilon) = 0$

Schätzverfahren

- Das bei der nicht-hierarchischen, linearen OLS-Regression eingesetzte Schätzverfahren anhand des Least-Squares Kriterium ist zur Schätzung von zufälligen Varianzkomponenten nicht geeignet.
- Die rechnerische Bestimmung der Modellparameter (z.B. Regressionsgewichte, Varianzen) erfolgt in der Regel anhand von Maximum-Likelihood (ML) Schätzverfahren.

  Das Spezifische der ML-Verfahren besteht darin, dass die Parameter des Modells so geschätzt werden, dass die Wahrscheinlichkeit (die Likelihood L) dafür, dass die gegebenen Daten im Sinne des Modells zu erklären sind, maximal wird.

  Die Likelihoodfunktion bezeichnet die Wahrscheinlichkeit aller Daten (beobachtete Varianzen, Kovarianzen und Mittelwerte aller Variable) in Abhängigkeit der Modellparameter.

  Das ML-Verfahren ist ein iteratives Verfahren, d.h. die besten Schätzungen werden schrittweise ermittelt, indem verschiedene Werte für die Modellparameter angenommen werden und jeweils die Likelihoodfunktion bestimmt werden. Ändert sich die Likelihood bei zwei aufeinanderfolgenden Iterationen nur noch unwesentlich, werden die aktuellen Werte als beste Schätzungen für die Modellparameter verwendet.

Inferenzstatistik für Modellparameter

- ML-Verfahren schätzen mit den fixen Effekten γ gleichzeitig auch deren Standardfehler SE_γ.
- Die Prüfgröße $z_\gamma = \gamma / SE_\gamma$ ist approximativ standardnormalverteilt.
- z_γ kann zur Prüfung der Nullhypothese, dass $\gamma = 0$ ist, herangezogen werden.
- z_γ setzt normalverteilte Residuen voraus. Simulationsstudien zeigen:

  Bei einer Verletzung der Normalverteilungsannahme werden die Standardfehler unterschätzt. Es resultieren zu liberale statistische Tests.

  Schätzungen der festen Effekte sind jedoch sehr robust gegenüber einer Verletzung der Normalverteilungsannahme.
- Wenn zumindest 30 Gruppen auf Ebene 2 vorliegen, dann sind auch die Schätzungen der zufälligen Effekte sehr robust gegenüber einer Verletzung der Normalverteilungsannahme (Maas & Hox, 2004).

Modellvergleiche

- Die Likelihood ist einerseits ein Optimierungskriterium zur Bestimmung der Modellparameter. Andererseits sagt sie etwas darüber aus, ob das Modell auf die Daten passt: Ist die Likelihood groß, dann ist es wahrscheinlich, dass das Modell für die beobachteten Daten gültig ist.
- Die Beurteilung der Modellpassung erfolgt auf Basis der **Devianz**:

$$Dev = -2 \cdot \ln(L)$$

mit ln = natürlicher Logarithmus zur Basis e (= Euler'sche Zahl e = 2.71828…)

- Je <u>größer</u> Dev, desto <u>schlechter</u> passt das Modell auf die Daten.
- Die Devianz kann nicht direkt für ein Modell interpretiert werden. Aber anhand von Devianztests können zwei Modelle verglichen werden, um das <u>bessere</u> Modell zu identifizieren:

$$\chi^2(df) = Dev_1 - Dev_2$$

mit $df = t_2 - t_1$ Freiheitsgraden
(t = Anzahl zu schätzender Modellparameter)

Prüft die Nullhypothese, dass beide Modelle gleich gut zu den Daten passen,

Wird die H_0 verworfen, wird das allgemeinere Modell mit mehr Parametern angenommen

Ausblick

1. Hierarchische Datenstrukturen
2. Modell der hierarchischen Regression
3. Forschungsfragen
4. Inferenzstatistische Tests
5. **Anwendungsbeispiel**

Anhänge

- R-Funktionen

Anwendungsbeispiel

- **Fragestellung**: Wie beeinflusst die Anzahl der wöchentlichen Arbeitsstunden das subjektive Wohlbefinden von 7,382 Mitarbeiterinnen und Mitarbeitern in 99 mittelständischen Unternehmen?

- **Variablen**:

  Abhängige Variable: subjektives Wohlbefinden der Mitarbeiterinnen und Mitarbeiter (z-standardisiert)

  Unabhängige Personenvariable: Anzahl der wöchentlichen Arbeitsstunden der Mitarbeiterinnen und Mitarbeiter

  Unabhängige Gruppenvariable: durchschnittliche Anzahl der wöchentlichen Arbeitsstunden der Mitarbeiterinnen und Mitarbeiter innerhalb eines Unternehmens

- **Hierarchische Struktur**:

  Ebene 1: Mitarbeiterinnen und Mitarbeiter

  Ebene 2: Unternehmen

Modellvergleiche

- Es werden unterschiedliche Modelle verglichen, um zu entscheiden welche unabhängigen Variablen und zufälligen Effekte die Daten optimal beschreiben:

  Modell 1:
  - Es werden keine UVs verwendet.
  - Wie viel Varianz im Wohlbefinden wird durch die Zugehörigkeit zu unterschiedlichen Firmen erklärt?

  Modell 2:
  - Es werden die individuellen Arbeitsstunden der Personen (ohne random Slope) als UV verwendet.
  - Erklären die Anzahl der Arbeitsstunden der Mitarbeiterinnen und Mitarbeiter deren Wohlbefinden?

  Modell 3:
  - Es wird zusätzlich ein random Slope für die individuellen Arbeitsstunden berücksichtigt.
  - Unterscheidet sich der Effekt der individuellen Arbeitsstunden auf das Wohlbefinden zwischen den Firmen?

  Modell 4:
  - Es werden die durchschnittlichen Arbeitsstunden in den Firmen als zusätzliche UV berücksichtigt.
  - Erklärt der Kompositionseffekt zusätzliche Varianz im Wohlbefinden?

  Modell 5:
  - Es wird zusätzlich die Cross-level Interaktion zwischen den individuellen Arbeitsstunden und den durchschnittlichen Arbeitsstunden berücksichtigt.
  - Hängt der Effekt der individuellen Arbeitsstunden auf das Wohlbefinden von den durchschnittlichen Arbeitsstunden in einer Firma ab?

Modellvergleiche

Likelihood

Devianztest

Modell	$\ln(L)$	Anzahl an Parametern	Verglichene Modelle	$\chi^2(df)$	p
Modell 1: Intercept-only	-10380	1	-	-	-
				Ja, Modell 2 passt besser als 1	
Modell 2: Personeneffekt von Arbeitsstunden	-10335	2	2 vs. 1	9.00 (1)	< .001
	Modell 3 und Modell 2 werden verglichen, kein sig. Untersch. Untersch. zwischen Arbeitszeit u Wohlb. bei allen				
Modell 3: Personeneffekt mit Random Slope	-10334	4	3 vs. 2	2.19 (2)	.225
Modell 4: Personen- & Gruppeneffekte von Arbeitsstunden	-10305	3	4 vs. 2	8.22 (1)	< .001
		Modell 2 und 4. werden verglichen weil 2 bisher das beste war. Ist sig.			
				Vergleich Modell 5 und 4, nicht sig.	
Modell 5: Cross-Level Interaktion	-10304	4	5 vs. 4	0.25 (1)	.348

- Es liegt ein signifikanter Personeneffekt für die Arbeitsstunden vor (Vergleich Modell 1 und 2).
- Es liegt ein signifikanter Gruppeneffekt für die Arbeitsstunden vor (Vergleich Modell 2 und 4).
- Es liegt kein random Slope für die Arbeitsstunden vor (Vergleich Modell 2 und 3).
- Es liegt keine Cross-level Interaktion vor (Vergleich Modell 4 und 5).

- **Umsetzung in R**: Hierarchische lineare Modelle können in R mit dem *nlme* Paket (Pinheiro et al., 2019) geschätzt werden (siehe Anhang).

Personen und Gruppeneffekte

- Die Intraklassenkorrelation (aus Modell 1) ist $ICC = .04$.
- Ergebnisse von Modell 4, welches die Daten am Besten beschreibt:

Parameter	Schätzung	Standardfehler
Fixe Effekte:		
γ_{00}	2.16^*	0.23
γ_{01}	-0.05^*	0.01
γ_{10}	-0.19^*	0.02
Zufällige Effekte:		
$SD(\upsilon_0)$	0.13^*	
Residuum:	Abhängige Variable war z-Standardisiert	
$SD(\varepsilon)$	0.97	

$^*\ p < .05$

Ebene 1:

$$WB_{ij} = \beta_{0j} + \beta_{1j} \quad HRS_{ij} + \varepsilon_{ij}$$

Ebene 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} \quad G.HRS_j + \upsilon_{0j}$$ (random intercept)

$$\beta_{1j} = \gamma_{10}$$ (fixed slope)

Ebene 1 + 2:

Arbeitszeit der Personen

$$WB_{ij} = \gamma_{00} + \gamma_{01} \quad G.HRS_j + \upsilon_{0j} + \gamma_{10} \quad HRS_{ij} + \varepsilon_{ij}$$

mit `WB` = subjektives Wohlbefinden der Personen
`HRS` = individuelle Arbeitsstunden der Personen
`G.HRS` = durchschnittliche Arbeitsstunden pro Firma

Zitierte Quellen

- Maas, C. J. M., & Hox, J. J. (2004). Robustness issues in multilevel regression analysis. *Statistica Netherlandica, 58*, 127-137.
- Marsh, H. W., Köller, O., & Baumert, J. (2001). Reunification of East and West German school systems: Longitudinal multilevel modeling study of the big-fish-little-pond effect on academic self-concept. *American Educational Research Journal, 38*, 321-350.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team (2019*). nlme: Linear and Nonlinear Mixed Effects Models*. URL: https://CRAN.R-project.org/package=nlme
- Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. London, United Kingdom: Sage.

Anhang: Beispielanalysen in R

Hierarchische lineare Modelle

Umsetzung in R

- **Variablen aus Anwendungsbeispiel**:

  `WB` = subjektives Wohlbefinden der Mitarbeiterinnen und Mitarbeiter (z-standardisiert)

  `HRS` = Anzahl der wöchentlichen Arbeitsstunden der Mitarbeiterinnen und Mitarbeiter

  `G.HRS` = durchschnittliche Anzahl der wöchentlichen Arbeitsstunden der Mitarbeiterinnen und Mitarbeiter innerhalbe einer Firma

  `GRP` = Gruppenindikator, der unterschiedliche Firmen kennzeichnet

- **Umsetzung von Modell 4 in R**:

```
> library(nlme)
> mod <- lme(fixed = WB ~ 1 + HRS + G.HRS,
             random = ~ 1 | GRP,
             data = data)
> summary(mod)
```

Paket laden

abh. Variable ist *WB*

Zufallsvarianz für den Intercept

Daten liegen im Objekt `data`

Ergebnisse ausgeben

Mehrebenenanalysen in R: nlme

Befehl:

`lme(fixed, random, data)`

- `fixed`: definiert die fixen Effekte, z.B.: `fixed = av ~ 1 + v1 + v1 * v2`

  Intercept

  abh. Variable

  unabh. Variable `v1`

  Haupteffekte und Interaktion der unabh. Variablen `v1` und `v2`

- `Random`: definiert die zufälligen Effekte, z.B.:
  `random = ~ 1 | ID`

  Zufallseffekt für den Intercept

  Legt die Gruppen fest

- `data`: zu analysierende Rohdaten

Modellvergleiche

```
# 1. Varianz innerhalb und zwischen Gruppen
> mod1 <- lme(fixed = WB ~ 1, random = ~ 1 | GRP, data = data)

# 2. Personeneffekte
> mod2 <- lme(fixed = WB ~ 1 + HRS, random = ~ 1 | GRP, data = data)

# 3. Gruppeneffekte
> mod3 <- lme(fixed = WB ~ 1 + HRS + W.HRS, random = ~ 1 | GRP, …)

# 4. Unterschiedliche Personeneffekte
> mod4 <- lme(fixed = WB ~ 1 + HRS, random = ~ 1 | GRP + HRS, …)

# 5. Cross-level Interaktion
> mod5 <- lme(fixed = WB ~ 1 + HRS * W.HRS, random = ~ 1 | GRP, …)
```

Ergebnisse

```
> summary(mod4)
Random effects:
  Formula: ~1 | GRP
           (Intercept) Residual
  StdDev: 0.1255254    0.9722988
Fixed effects:
  WBEING ~ 1 + W.HRS + G.HRS
             Value      Std.Error  DF    t-value  p-value
(Intercept) 2.1594440 0.23260785 7282  9.283625 0
W.HRS      -0.0511486 0.00538196 7282 -9.503706 0
G.HRS      -0.1909440 0.02044098 97   -9.341236 0
```

(Residual-)Varianz auf Ebene 1 (ε)

(Residual-)Varianz auf Ebene 2 (υ_0)

Fixe Effekte γ_{00}, γ_{01}, γ_{10}.

Anmerkung: *nlme* verwendet t-Tests zur Signifikanzprüfung von Regressionsgewichten, während andere Software z-Tests berichten.

Anhang: R-Funktionen

Hierarchische lineare Modelle

R-Funktionen

Funktion	Paket	Beschreibung
`lme(fixed, random, data, subset)`	`nlme`	Berechnet für die mit `subset` ausgewählten Fälle aus den Daten `data` die in `fixed` spezifizierte lineare Regression mit den in `random` angegebenen Zufallseffekten
`summary(x)`	`nlme`	Gibt die Ergebnisse des in `x` geschätzten hierarchischen linearen Modells aus
`anova(object, …)`	`nlme`	Devianztests für die in `object` und … geschätzten hierarchischen linearen Modelle