

# Mehr-Ebenen-Analyse II

## Regressionsmodelle für Politikwissenschaftler

Wiederholung  
Konzepte und Begriffe  
Längsschnittmodelle/Panelmodelle  
Binäre abhängige Variable und verwandte Probleme  
Fazit

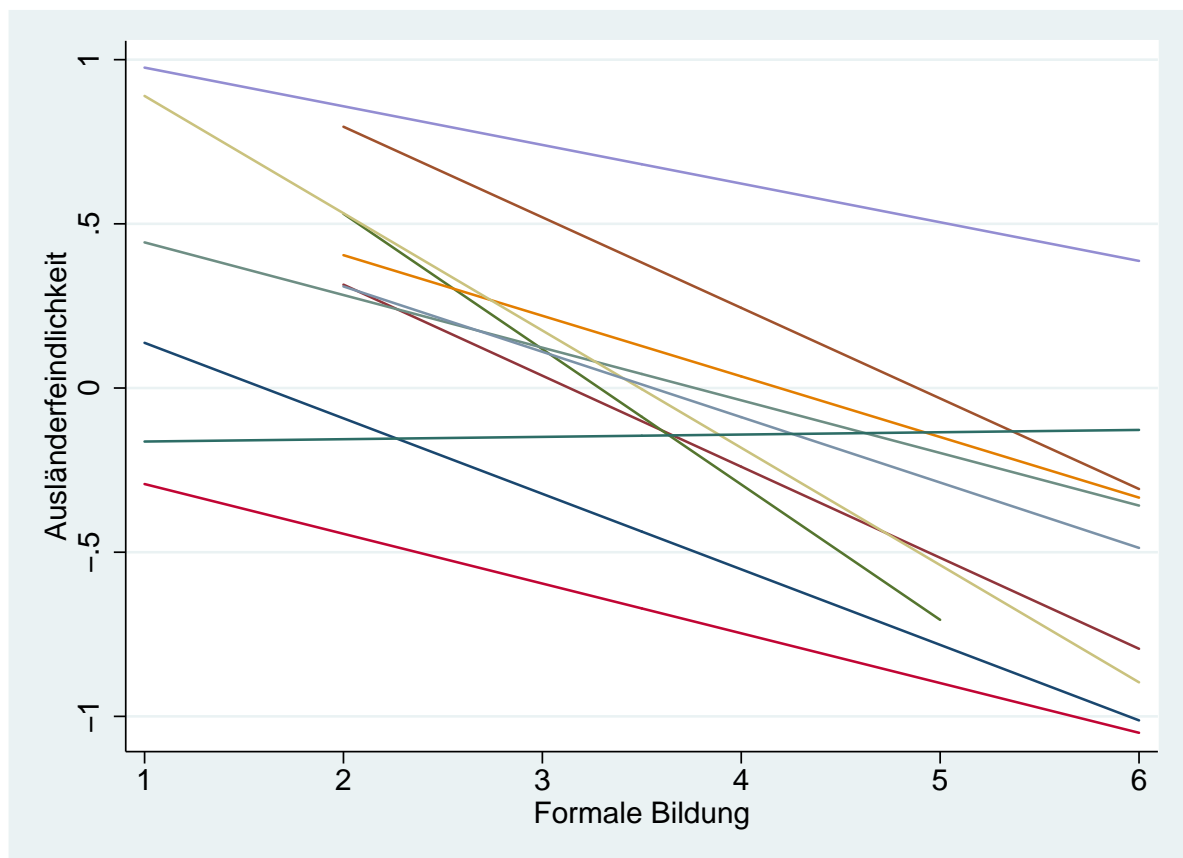
### Wie war das mit der Übung?

Sie untersuchen den Zusammenhang zwischen der individuellen formalen Bildung ( $x$ ), dem Ausländeranteil auf Kreisebene ( $z$ ) und der individuellen Ausländerfeindlichkeit ( $y$ ) in Ost- und Westdeutschland

1. Was spricht dagegen, ein Drei-Ebenen-Modell (Region – Kreis – Person) zu konstruieren? Welche einfachere Lösung ist sinnvoller?
2. Bitte schreiben Sie diese einfachere Lösung als Formel nieder. Gehen Sie dabei von fixen Effekten für Bildung und Ausländeranteil aus.
3. Wären random effects für diese beiden Variablen möglich/sinnvoll?
4. Eine Kommilitonin macht Sie darauf aufmerksam, daß der Ausländeranteil in Ost- beziehungsweise Westdeutschland einen unterschiedlichen Effekt haben dürfte (warum?). Wie läßt sich dies im Modell berücksichtigen?

## Welche Alternativen zum Mehr-Ebenen-Modell gibt es?

- ▶ 108 einzelne Regressionen → Ausländeranteil und Ost-West kann nicht systematisch berücksichtigt werden
- ▶ Gemeinsame Regression mit Bildung, Ost und gegebenenfalls Interaktion, mit und ohne robuste Standardfehler



## Welche Alternativen zum Mehr-Ebenen-Modell gibt es?

- ▶ 108 einzelne Regressionen → Ausländeranteil und Ost-West kann nicht systematisch berücksichtigt werden
- ▶ Gemeinsame Regression mit Bildung, Ost und gegebenenfalls Interaktion, mit und ohne robuste Standardfehler

## Wie sieht die Lösung per ML aus?

- ▶ Gegen ein Drei-Ebenen-Modell spricht, daß die Regionen „Ost“ und „West“ nicht als Elemente einer Stichprobe aufzufassen sind, deren Effekt irgendeiner Verteilung folgen könnte
- ▶ Die einfache Lösung: Ost/West als Dummy auf der Kreisebene (weitere Eigenschaft des Kreises), random intercept

$$y_{ij} = \beta_{0ij} + \beta_1 x_{1ij} + \beta_2 x_{2j} + \beta_3 x_{3ij} \quad (1)$$

$$\beta_{0ij} = \beta_0 + u_{0j} + e_{0ij} \quad (2)$$

mit

$$(u_{0j}) \sim N(0, \mathbf{\Omega}_u) : \mathbf{\Omega}_u = (\sigma_{u0}^2) \quad (3)$$

$$(e_{0ij}) \sim N(0, \mathbf{\Omega}_e) : \mathbf{\Omega}_e = (\sigma_{e0}^2) \quad (4)$$

## Während random effects für Bildung/Ausländeranteil möglich/sinnvoll?

- ▶ Random Effect für Ausländeranteil hieße, daß der Ausländeranteil im Kreis in jedem Kreis einen anderen Effekt hätte → nicht schätzbar und eigentlich auch nicht sinnvoll
- ▶ Random Effect für Bildung → In Grafik oben scheinbar stark variierender Effekt der Bildung → empirisch unbedeutend (Fallzahlen auf Kontextebene zwischen 8 und 99)

## Unterschiedlicher Wirkung des Ausländeranteils in Ost und West?

- ▶ Ausländeranteil im Osten 1,5 Prozent, im Westen 10 Prozent, trotzdem Ausländerfeindlichkeit im Osten viel höher
- ▶ Unterschiede in Sozialisation und Strukturen etc.
- ▶ Interaktion zwischen Ausländeranteil und Ost/West

## Wie lassen sich die Ergebnisse zusammenfassen?

- ▶ Modellierung mit ML sinnvoll
- ▶ Struktur hat aber keine dramatischen Effekte
  - ▶ Standardfehler
  - ▶ Punktschätzungen sehr ähnlich, aber nicht identisch (shrinkage estimates)
  - ▶ Geringer Effekt der Kontexte – Kontexte bereits zu groß?
- ▶ Konventionelle Methoden führen im wesentlichen zum selben Ergebnis
- ▶ Das weiß man naturgemäß aber erst hinterher

## Was bedeuten Random & Fixed?

- ▶ In der Methodenliteratur drei verschiedene Bedeutungen, die zusammenhängen, aber nicht identisch sind
  1. Random/Fixed Variables
  2. Random/Fixed Effects
  3. Random/Fixed Coefficients
- ▶ Gute Darstellung hier: [http://www.ioa.pdx.edu/newsom/mlrclass/ho\\_randfixd.doc](http://www.ioa.pdx.edu/newsom/mlrclass/ho_randfixd.doc)

## Was sind Random/Fixed Variables?

- ▶ Fixed Variable: Ohne Fehler gemessen, gleicher Meßwerte, wenn Stichprobenziehung/Untersuchung wiederholt wird
- ▶ Random Variable: Meßwerte sind Stichprobe aus Grundgesamtheit möglicher Werte und würden in nächster Stichprobe etwas anders aussehen (z. B. wegen zufälligen Meßfehlern)
- ▶ Unsere Standardmodelle gehen immer von fixed variables aus, viele sind aber tatsächlich random (Einstellungen!) → Standardfehler vermutlich häufig zu optimistisch

## Was sind Random & Fixed Effects?

- ▶ Wenn sich Generalisierung auf die in einer Studie tatsächlich vorkommenden Werte der unabhängigen Variablen beschränkt, fixed effects
- ▶ Wenn Werte als Stichprobe aus einem Universum von möglichen Werten betrachtet werden, random effects (größere Standardfehler)
- ▶ Vor allem im Zusammenhang mit experimentellen Studien (ANOVA) diskutiert: Sind die Stimuli fixed oder random? (Beispiel: Medikamentendosen vs. Kunstwerke)

## Was sind Random & Fixed Coefficients?

- ▶ Im Mehrebenenmodell geht man davon aus, daß die unabhängigen Variablen fixed sind
- ▶ Die Achsenabschnitte/Steigungen können aber auf der unteren Ebene variieren → random coefficients → Gruppen sind Stichprobe aus Grundgesamtheit von Gruppen (in denen jeweils unterschiedliche Zusammenhänge gelten)
- ▶ Random Coefficient ist mit random variable vergleichbar
- ▶ Bestimmte Mehr-Ebenen-Modelle sind strukturell mit bestimmten Typen von ANOVA für random effects identisch

## Welche Grundtypen von ML-Modellen gibt es?

$$y_{ij} = \beta_0 + u_{0j} + e_{0ij} \quad (5)$$

- ▶ Dieses Modell wird als „empty model“ oder „intercept-only“ model bezeichnet
- ▶ Relative Bedeutung von Kontext-/Individualeffekten kann bestimmt werden (mehr dazu gleich)
- ▶ Sehr nützlich als Ausgangspunkt für komplexere Modelle

## Welche Grundtypen von ML-Modellen gibt es?

$$y_{ij} = \beta_0 + \beta_1 x_1 + \dots + u_{0j} + e_{0ij} \quad (6)$$

- ▶ Wenn man das empty model um Variablen auf der Individual- und gegebenenfalls Kontextebene (fixed coefficients) erweitert, erhält man ein „random intercept“ model
- ▶ „Normale“ Regression, aber zufällige, kontextspezifische Einflüsse, die Achsenabschnitt variieren lassen, Varianzkomponentenmodelle

## Welche Grundtypen von ML-Modellen gibt es?

$$y_{ij} = \beta_0 + \beta_{1ij} x_1 + \dots + u_{0j} + e_{0ij} \quad (7)$$

$$\beta_{1ij} = \beta_1 + u_{1j} \quad (8)$$

$$(9)$$

- ▶ Wenn man den Effekt von Individualvariablen zufällig über die Gruppen/Kontexte variieren läßt, erhält man ein „random coefficient“ model
- ▶ Weitere Komplikationen möglich
  - ▶ Cross-Level-Interaktionen
  - ▶ Subgruppenspezifische Kontextvarianzen etc.
  - ▶ Sinn? Test von Koeffizienten/Varianzen per Wald-Test, Veränderung der Devianz, BIC

## Wie war das mit den Varianzkomponenten?

$$\text{VPC} = \rho = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_{e0ij}^2} \quad (10)$$

- ▶ Im Empty Model wird die Gesamtvarianz als zufällig betrachtet . . .
- ▶ Und auf die Ebenen der Struktur aufgeteilt ( $u_{0j} + e_{0ij}$ )
- ▶ Der Anteil der Kontextvarianz an der Gesamtvarianz wird als Variance Partition Coefficient oder auch als Intra-Class-Correlation bezeichnet
- ▶ Intra-Class-Correlation ist die mittlere Korrelation zwischen den Werten von  $y$ , die sich errechnet, wenn ich Paare aus derselben Gruppe betrachte
- ▶ Im Zwei-Ebenen-Modell fallen IPC und  $\rho$  zusammen, in Modellen mit mehr als zwei Ebenen können sie sich unterscheiden

## Was sind Längsschnittmodelle?

- ▶ Dieselben Objekte(Personen) werden mehrmals untersucht
- ▶ Invarianz der Instrumente/Memory-Effekte
- ▶ Betrachtung als Mehr-Ebenenmodell: Person ist Kontext/Ebene 2, Messungen sind Untersuchungsobjekte/Ebene 1
- ▶ Fixed/varying occasions

## Ein Beispiel

- ▶ Ein sechs aufeinanderfolgenden Semestern werden jeweils zum selben Zeitpunkt die akademischen Leistungen, die Zeit fürs Jobben, das Geschlecht und der GPA aus der Schule erhoben
- ▶ Bis zu sechs Messungen pro Person, möglicherweise aber missings für Meßzeitpunkte, eine unabhängige Variable variiert über die Zeit
- ▶ Zwei Aspekte:
  1. Wie sieht der Zusammenhang zwischen unabhängigen und abhängiger Variablen aus?
  2. Gibt es bei der abhängigen Variablen einen Trend über die Zeit (Autoregression)

## Was macht man hier mit ML?

- ▶ Standardmodell (random intercept)
- ▶ Keine besonderen Komplikationen, korrekte Schätzungen für Zusammenhang zwischen unabhängigen Variablen und akademischer Leistung, korrekte Standardfehler

## Was wird traditionell gemacht?

- ▶ Mehr oder minder komplizierte Formen von Varianzanalyse für wiederholte Messungen
- ▶ Einbeziehung von Variablen, die über die Zeit variieren problematisch
- ▶ Ergebnis hier: Leistungen steigen über die Zeit linear an, signifikanter Effekt des Geschlechts und der Leistungen in der Schule

## Wie kann man die Entwicklung im ML-Modell berücksichtigen?

- ▶ Auch ML kann Trend inkorporieren
- ▶ Ergänzung des Modell um Variable für Meßzeitpunkt
- ▶ Effekt kann über Personen variieren → individuelle (random) Lernkurven
- ▶ Für verschiedene Modellvarianten vgl. Tabelle 5.3 auf Seite 81
- ▶ Entwicklung der Varianzen auf beiden Ebenen über Modelle hinweg → „Varianz explained“ in Mehr-Ebenen-Modellen noch problematischer als sonst
- ▶ VPC hängt von Modellierung ab

## Was sind die Hauptergebnisse?

- ▶ Trend macht die Unterschiede zwischen den Personen etwas besser sichtbar
- ▶ Arbeitszeit erklärt einen größeren Teil der Unterschiede zwischen den Personen, aber nur einen geringen Teil der Unterschiede „innerhalb“ einer Person (also zwischen den Zeitpunkten) → wie stark variiert Variable über die Zeit?
- ▶ Vorleistungen und Geschlecht erklären einen Teil der verbleibenden Unterschiede zwischen den Personen

## Wie war das mit den Lernkurven?

- ▶ In Tabelle 5.4 kann der Koeffizient für die Lernkurve frei variieren → Varianz von 0,004
- ▶ Außerdem wird in M7 noch eine Cross-Level-Interaktion zugelassen (Geschlecht und Zeit)
- ▶ Achtung: Ich glaube, im Text ist ein Interpretationsfehler (0,038 ist die Varianz des Achsenabschnitts, nicht der Steigung)
- ▶ Verschiedene Kodierungen können für weitere Konfusion sorgen

## Welche weiteren Weiterungen gibt es?

- ▶ Das nächste Beispiel kombiniert Daten aus zwei Entwicklungsstudien und ungleiche Abstände zwischen Messungen
- ▶ Außerdem werden jetzt nicht-lineare Entwicklungstrends zugelassen
- ▶ Konvergenzprobleme möglich/Unterschiede durch Algorithmen

## Welche Vorteile bringt ML hier?

1. Individuelle Wachstums-/Lernkurven (für uns relevant?)
2. Anzahl und Abstand zwischen Messungen kann variieren
3. Korrelierte Meßfehler können modelliert werden
4. Bei Verwendung bestimmter Schätzverfahren äquivalent zu traditionellen Verfahren
5. Weitere Ebenen können hinzukommen
6. Über die Zeit variierende Variablen möglich

## Was gibt es sonst noch?

- ▶ Entwicklung über die Zeit kann auch durch andere Funktionen oder einfach durch Dummies modelliert werden
- ▶ Fehlende Werte werden quasi-automatisch berücksichtigt, MAR ist einigermaßen plausibel
- ▶ Korrelierte Fehler/variierende Fehlervarianzen möglich, wenn entsprechende Matrizen (mit Restriktionen) formuliert werden

## Welche Komplikationen gibt es für uns?

- ▶ Binäre abhängige Variablen beziehungsweise Anteilswerte oder Häufigkeiten machen im Mehr-Ebenen-Modell die gewohnten Schwierigkeiten
- ▶ Generalisierte Lineare Modelle sind uns ja schon bekannt
- ▶ Das Mehr-Ebenen-Modell kann entsprechend erweitert werden

## Wie sieht das ML-Modell für die logistische Regression aus?

$$y_{ij} \sim \text{Binomial}(1, \pi_{ij}) \quad (11)$$

$$\text{logit}(\pi_{ij}) = \beta_{0j} + \beta_1 x_{1ij} + \beta_2 x_{2j} \quad (12)$$

$$\beta_{0j} = \beta_0 + u_{0j} \quad (13)$$

mit

$$(u_{0j}) \sim N(0, \mathbf{\Omega}_u) : \mathbf{\Omega}_u = (\sigma_{u0}^2) \quad (14)$$

- ▶ Bekannte Struktur
- ▶ Keine Varianz auf der untersten Ebene, da diese bei der logistischen Regression nicht geschätzt, sondern fixiert wird, um das Modell zu identifizieren

## Welche sonstigen Besonderheiten gibt es?

- ▶ Parameter für Mehr-Ebenen-Modelle und GLM werden per Maximum Likelihood bestimmt
- ▶ Aber: Funktion ist selbst für einfache Modelle so komplex, daß Schätzung praktisch unmöglich
- ▶ Linearisierung des Modells durch eine Taylor-Serie → Pseudo-Maximum-Likelihood
- ▶ Verschiedene Varianten (MQL1/MQL2 vs. PQL1/PQL2) → Verzerrungen; immer nur Pseudo-Likelihood, d. h. keine gültigen Devianz-Werte, kein LR-Test
- ▶ Alternative computerintensive Verfahren (bootstrapping, bayesianisch)
- ▶ Übertragung auf Poisson-Regression etc. möglich

## Ein Beispiel?

- ▶ Im Text relativ komplexes metanalytisches Beispiel
- ▶ Statt dessen ein einfacheres Beispiel?
- ▶ Eurobarometer-Daten von 1980-2002 aus 15 Ländern
- ▶ Struktur: Befragter – nationale Umfragewelle – Land (als fixed betrachtet, über Dummies modelliert, keine Konstante)
- ▶ Zwei-Ebenen-Modell mit ca. 300 Kontexten
- ▶ Wahl der extremen Rechten in Abhängigkeit von sozio-demographischen Variablen und Kontext

## Wie sieht das Modell aus?

$$y_{ij} \sim \text{Binomial}(1, \pi_{ij}) \quad (15)$$

$$\text{logit}(\pi_{ij}) = \beta_1 AT_j \cdots + \beta_{15} SE_j + \beta_{16} x_{16ij} \cdots \beta_{25} x_{25ij} + u_{0j} \quad (16)$$

mit

$x_{16ij} \dots x_{25ij}$  : Geschlecht, Alter, Bildung, Berufsgruppe

$$(u_{0j}) \sim N(0, \mathbf{\Omega}_u) : \mathbf{\Omega}_u = (\sigma_{u_0}^2) \quad (17)$$

## Was sind die Ergebnisse?

Parameter/Kennwerte	listenweises Löschen		multiple Imputation	
Geschlecht: männlich	0,508	(0,022)	0,486	(0,020)
18-29 Jahre	0,293	(0,031)	0,282	(0,038)
30-45 Jahre	0,055	(0,030)	0,051	(0,030)
älter als 65 Jahre	-0,041	(0,041)	-0,042	(0,041)
Bildung: mittel/hoch	0,111	(0,027)	0,101	(0,029)
Bildung: Hochschulabschluß petty bourgeoisie	-0,298	(0,033)	-0,267	(0,030)
Arbeiter	0,380	(0,036)	0,360	(0,035)
Rentner	0,325	(0,030)	0,338	(0,031)
Arbeitslose	0,121	(0,040)	0,120	(0,040)
	0,467	(0,042)	0,484	(0,040)
AT	-2,086	(0,224)	-2,096	(0,224)
BE	-3,517	(0,152)	-3,477	(0,153)
DE-O	-4,678	(0,176)	-4,680	(0,176)
...	...	...	...	...
SE	-6,339	(0,420)	-6,089	(0,438)
$\sigma_{u0}^2$	0,527	(0,049)	0,533	(0,051)
N(1)	235 320		307 869	
N(2)	314		314	

Regressionsmodelle für Politikwissenschaftler

Mehr-Ebenen-Analyse II (33/35)

## Was ist das Fazit für heute?

- ▶ Mehr-Ebenen-Modelle eröffnen viele neue Analysemöglichkeiten
- ▶ Außerdem bieten sie oft alternative Lösungen für altbekannte Probleme
- ▶ Wichtig ist, die Übersicht zu bewahren

# Was ist die Übungsaufgabe

- ▶ Betrachten Sie nochmals die Tabelle auf der vorletzten Folie
- ▶ Mit welcher Wahrscheinlichkeit wird ein 22 Jahre alter arbeitsloser Mann mit einfacher Bildung, der in Ostdeutschland lebt, die extreme Rechte wählen (multiple Imputation)
- ▶ Welche Wahrscheinlichkeit erwarten Sie für die gleiche Person, wenn diese statt dessen in Österreich lebt?
- ▶ Wie wirkt sich  $\sigma_{u0}^2$  auf diese Wahrscheinlichkeiten aus?
- ▶ Antwort bitte bis nächsten Mittwoch an die bekannte Adresse