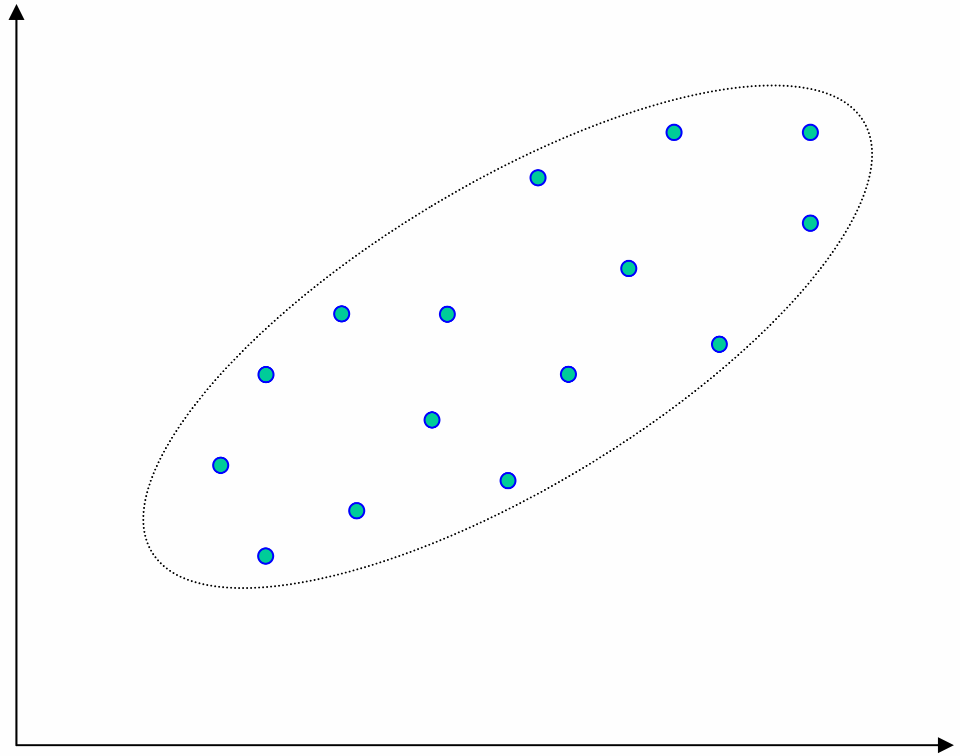
Ökologischer Fehlschluss und Cross-Level Inferenz

Inwieweit können wir von statistischen Beziehungen zwischen **Aggregat**merkmalen auf die Beziehungen zwischen den korrespondierenden **Individual**merkmalen schließen?

Kann z.B. aus einer Korrelation, die **„im Aggregat“** zwischen zwei Merkmalen besteht, auf die Korrelation zwischen den korrespondierenden **Individual**merkmalen geschlossen werden?

Anteil: im Ausland geborener Einwohner



Anteil: lese- und schreibkundig

W.S. Robinson‘s (1950) Beispiel

Analyseebenen

Aggregatebene(n)	**Kollektive** Eigenschaften	Aggregatdaten-analyse
	[Inklusion]	
Individualebene	**Individuelle** Eigenschaften	Individualdaten-analyse

Lazarsfeld – Menzel – Typologie von Variablen

… für Mehrebenenanalysen

	Absolute Eigenschaften	Eigenschaften, die basiert sind auf ..		
		..Verteilung	..Struktur	..Inklusion
Einheit, Level $n+1$	globale	analytische	strukturelle	
Einheit, Level n	absolute	komparative	relationale	kontextuelle

Beispiel	Anteil 1 [z.B.Arbeits-losenrate]	Anteil 2 [z.B. Anteil pol. extremer Wahlen]
Aggregat-Einheit 1	0,20	0,30
Aggregat-Einheit 2	0,30	0,40
Aggregat-Einheit 3	0,40	0,50
Aggregat-Einheit 4	0,50	0,60
.	.	.
.	.	.
.	.	.
Aggregat-Einheit k	.	.

[Fiktive Zahlen]

[Fiktive Zahlen]

Agg.-Einheit 1

	X = 1	X = 2	
Y = 1	?	?	0,30
Y = 2	?	?	0,70
	0,20	0,80	

Agg.-Einheit 2

	X = 1	X = 2	
Y = 1	?	?	0,40
Y = 2	?	?	0,60
	0,30	0,70	

Agg.-Einheit 3

	X = 1	X = 2	
Y = 1	?	?	0,50
Y = 2	?	?	0,50
	0,40	0,60	

Agg.-Einheit 4

	X = 1	X = 2	
Y = 1	?	?	0,60
Y = 2	?	?	0,40
	0,50	0,50	

		X 1	X 2	
Y	1	$p_{1\|1}$	$p_{1\|2}$	p_{1+}
	2	$p_{2\|1}$	$p_{2\|2}$	p_{2+}
		p_{+1}	p_{+2}	1,0

		X 1	X 2	
Y	1	0,30	0,40	0,38
	2	0,70	0,60	0,62
		0,20	0,80	1,0

$$p_{1+} = p_{1|1} \cdot p_{+1} + p_{1|2} \cdot p_{+2}$$

$$p_{1+} = 0{,}3 \cdot 0{,}2 + 0{,}4 \cdot 0{,}8$$

Struktur einer 4-Felder-Tafel (mit Spaltenprozenten)

[Fiktive Zahlen]

		X		
		1	2	
Y	1	$p_{1\|1}$	$p_{1\|2}$	p_{1+}
	2	$p_{2\|1}$	$p_{2\|2}$	p_{2+}
		p_{+1}	p_{+2}	1,0

		X		
		1(AK)	2 (MK)	
Y	1 (D)	$p_{1\|1} = ?$	$p_{1\|2} = ?$	0,6
	2 (R)	$p_{2\|1} = ?$	$p_{2\|2} = ?$	0,4
		0,9	0,1	1,0

$$p_{1+} = p_{1|1} \cdot p_{+1} + p_{1|2} \cdot p_{+2}$$

$$p_{1+} = p_{1|1} \cdot 0{,}9 + p_{1|2} \cdot 0{,}1$$

$$p_{1|1} = \frac{p_{1+}}{p_{+1}} - \frac{p_{+2}}{p_{+1}} \cdot p_{1|2}$$

$$p_{1|1} = \frac{0{,}6}{0{,}9} - \frac{0{,}1}{0{,}9} \cdot p_{1|2}$$

$$p_{1|1} = \frac{0{,}6}{0{,}9} - \frac{0{,}1}{0{,}9} \cdot p_{1|2}$$

$$wenn: p_{1|2} = 1{,}0$$

$$p_{1|1} = \frac{0{,}6}{0{,}9} - \frac{0{,}1}{0{,}9} \cdot 1{,}0 = 0{,}56$$

$$wenn: p_{1|2} = 0{,}0$$

$$p_{1|1} = \frac{0{,}6}{0{,}9} - \frac{0{,}1}{0{,}9} \cdot 0{,}0 = 0{,}67$$

	X: AK	X: MK	
Y: D	$p_{1\|1} = ?$	$p_{1\|2} = ?$	0,6
Y: R	$p_{2\|1} = ?$	$p_{2\|2} = ?$	0,4
	0,9	0,1	1,0

AK Arbeiterklasse

MK Mittelklasse

D Wahl der Demokraten

R Wahl der Republikaner

$$p_{1|2} = \frac{0{,}6}{0{,}1} - \frac{0{,}9}{0{,}1} \cdot p_{1|1}$$

$$wenn: p_{1|1} = 1{,}0$$

$$p_{1|2} = \frac{0{,}6}{0{,}1} - \frac{0{,}9}{0{,}1} \cdot 1{,}0 = 0{,}0$$

$$wenn: p_{1|1} = 0{,}0$$

$$p_{1|2} = \frac{0{,}6}{0{,}1} - \frac{0{,}9}{0{,}1} \cdot 0{,}0 = 1{,}0$$

X

Y		AK	MK	
	D	$p_{1\|1} = ?$	$p_{1\|2} = ?$	0,6
	R	$p_{2\|1} = ?$	$p_{2\|2} = ?$	0,4
		0,9	0,1	1,0

$$p_{1+} = p_{1|1} \cdot p_{+1} + p_{1|2} \cdot p_{+2}$$

$$p_{1|2} = \frac{p_{1+}}{p_{+2}} - \frac{p_{+1}}{p_{+2}} \cdot p_{1|1}$$

Goodman‘s „ökologische Regression“

Annahmen: $p_{1|1(j)} = p_{1|1(.)}$ $p_{1|2(j)} = p_{1|2(.)}$

		X		
		1	2	
Y	1	$p_{1\|1(j)}$	$p_{1\|2(j)}$	$p_{1+(j)}$
	2	$p_{2\|1(j)}$	$p_{2\|2(j)}$	$p_{2+(j)}$
		$p_{+1(j)}$	$p_{+2(j)}$	1,0

$$\boxed{p_{1+(j)} = p_{1|1(j)} \cdot p_{+1(j)} + p_{1|2(j)} \cdot p_{+2(j)}}$$

$$p_{1+(j)} = p_{1|1(j)} \cdot p_{+1(j)} + p_{1|2(j)} \cdot \left(1 - p_{+1(j)}\right)$$

$$p_{1+(j)} = p_{1|1(j)} \cdot p_{+1(j)} + p_{1|2(j)} - p_{1|2(j)} \cdot p_{+1(j)}$$

$$p_{1+(j)} = p_{1|2(j)} + \left(p_{1|1(j)} - p_{1|2(j)}\right) p_{+1(j)}$$

$$= \underbrace{p_{1|2(.)}}_{} + \underbrace{\left(p_{1|1(.)} - p_{1|2(.)}\right)}_{} p_{+1(j)} + U_j$$

$$Y = a + b \times X + U$$

[Konstante][Steigung]

Nach: Achen, Ch. H.; W. Ph. Shively (1995) Cross-Level Inference. Chicago, p. 13

Klein und Friedman‘s „Nachbarschaftsmodell“

		X 1	2	
Y	1	$p_{1\|1(j)}$	$p_{1\|2(j)}$	$p_{1+(j)}$
	2	$p_{2\|1(j)}$	$p_{2\|2(j)}$	$p_{2+(j)}$
		$p_{+1(j)}$	$p_{+2(j)}$	1,0

Annahmen: $p_{1|1(j)} = p_{1|2(j)}$

$$p_{1+(j)} = p_{1|1(j)} \cdot p_{+1(j)} + p_{1|2(j)} \cdot p_{+2(j)}$$

$$p_{1+(j)} = p_{1|1(j)} \cdot p_{+1(j)} + p_{1|1(j)} \cdot p_{+2(j)}$$

mit $p_{1|1(j)} = a + bp_{+1(j)}$

(Goodman‘s und Klein/Friedman‘s Modell empirisch nicht unterscheidbar)

Nach: Achen, Ch. H.; W. Ph. Shively (1995) Cross-Level Inference. Chicago, p. 14