Mixture modelling – Ansätze und Anwendungsbeispiele aus Latent Class Analysen und Growth Mixture Modellen

Ringseminar: Statistische Methoden der Epidemiologie, Medizinischen Biometrie und Versorgungsforschung; Ulm, 14. Dezember 2015

Ferdinand Keller
Klinik für Kinder- und Jugendpsychiatrie / Psychotherapie
Universitätsklinikum Ulm

Zwei Zielrichtungen in der Datenanalyse:

Variable-centered vs. person-centered

Zusammenhänge zwischen Variablen, z.B. Faktorenanalyse, Korrelationen/Kovarianzen im Rahmen von Strukturgleichungsmodellen (SEM)

Klassifikation von Einheiten (Personen): Clusteranalyse, LCA etc.

Aber: zwei Seiten einer Medaille (= "Datenmatrix")

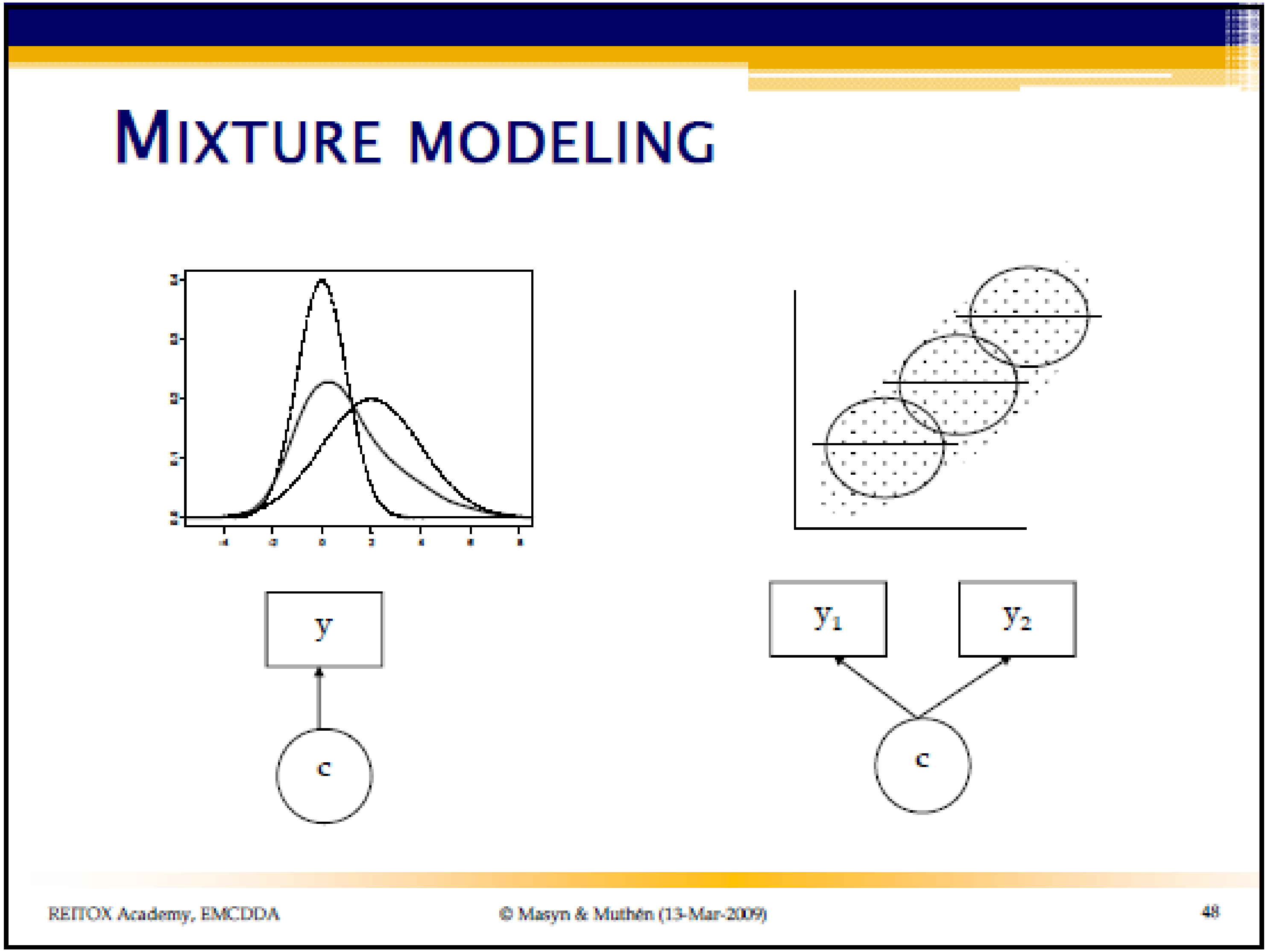
Allgemeine Idee:

Es wird angenommen, dass die beobachteten Summenstatistiken (z.B. Mittelwerte, Korrelationen) nicht von einer homogenen Gruppe stammen, sondern eine Mischung darstellen von zugrunde liegenden (unbekannten) Subgruppen.

Daher der globale Begriff:

“Mischverteilungsmodelle” (mixture models)

General idea of mixture modeling



Anwendungen in verschiedenen Disziplinen

Biologie: Mischung von Normalverteilungen

- Krabben aus der Bucht von Neapel (Pearson,1894)
- Alterskohorten von Fischen in einem Fluss
- Geysir Old Faithful (Yellowstone)

Sozialwissenschaften: Typologien, Abhängigkeiten bei kategorialen Variablen, psychometrische Meßmodelle etc.

- Lazarsfeld, 1950; Lazarsfeld & Henry, 1968

Medizin: “Typologie” hinsichtlich Kombination von Symptomen, Diagnostik

Marktforschung: Konsumer-Profile

Überblick und Beispiele dazu:

1) Normalverteilungen (Pearson)

2) LCA (dichotome / polytome Variablen)
  - Schmerzsymptome
  - Allg. Depressionsskala (Gruppe „unskalierbarer“)
  - Phänotypen von Asthma

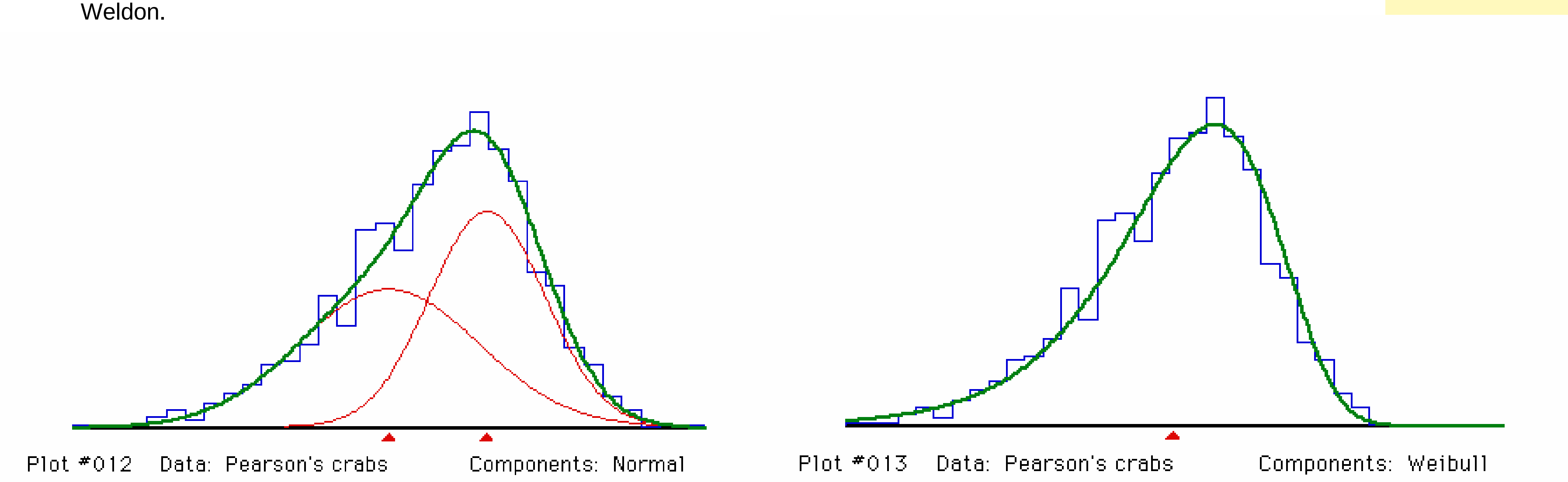
3) Growth mixture („trajectories“)
  - count-Variable (Delinquenz im Jugendalter)
  - latent class growth model (Depressionsstudie)
  - growth mixture
    - Depressionsstudie
    - BMI-Verläufe in Vorarlberg-Studie

4) „limitations“ und Ausblick

Pearson: Krabbenbeispiel

Pearson, K. (1894). Contributions to the mathematical theory of evolution. Phil. Trans. Roy. Soc. London A 185 , 71-110.

The data give the ratio of "forehead" breadth to body length for 1000 crabs sampled at Naples by Professor W.F.R. Weldon.



Pearson (1894) analysed this histogram by the method of moments. The calculation was formidable and done without the aid of computing machinery of any kind. He found two solutions, one with 41.45% of the population in the first component and 58.55% in the second, the other with 53.28% in the first component and 46.72% in the second. He preferred the first solution on the basis of agreement with the sixth moment. MIX does not converge to a unique solution if all parameters are unconstrained. The iterations wander between a 6:4 and 4:6 ratio for the two components, with no fit being significantly better than any other. The standard errors of the proportions are quite large. The fit shown here resolves this uncertainty by constraining the proportions to be equal. The presence of two components was interpreted by Pearson as evidence that there were two species of crabs. I know of no biological justification for constraining the proportions to be equal, but the fit obtained is excellent. Constraining the standard deviations to be equal does not give an acceptable fit.

Aus: Peter Macdonald, McMaster University, R package mixdist

Grundannahmen der latent class Analyse:

1) Lösungswahrscheinlichkeit eines Items ist für alle Personen einer Klasse konstant

2) jede Person gehört nur einer Klasse an (Summe der Wk = 1)

3) alle Items messen dieselbe (Personen)-Variable

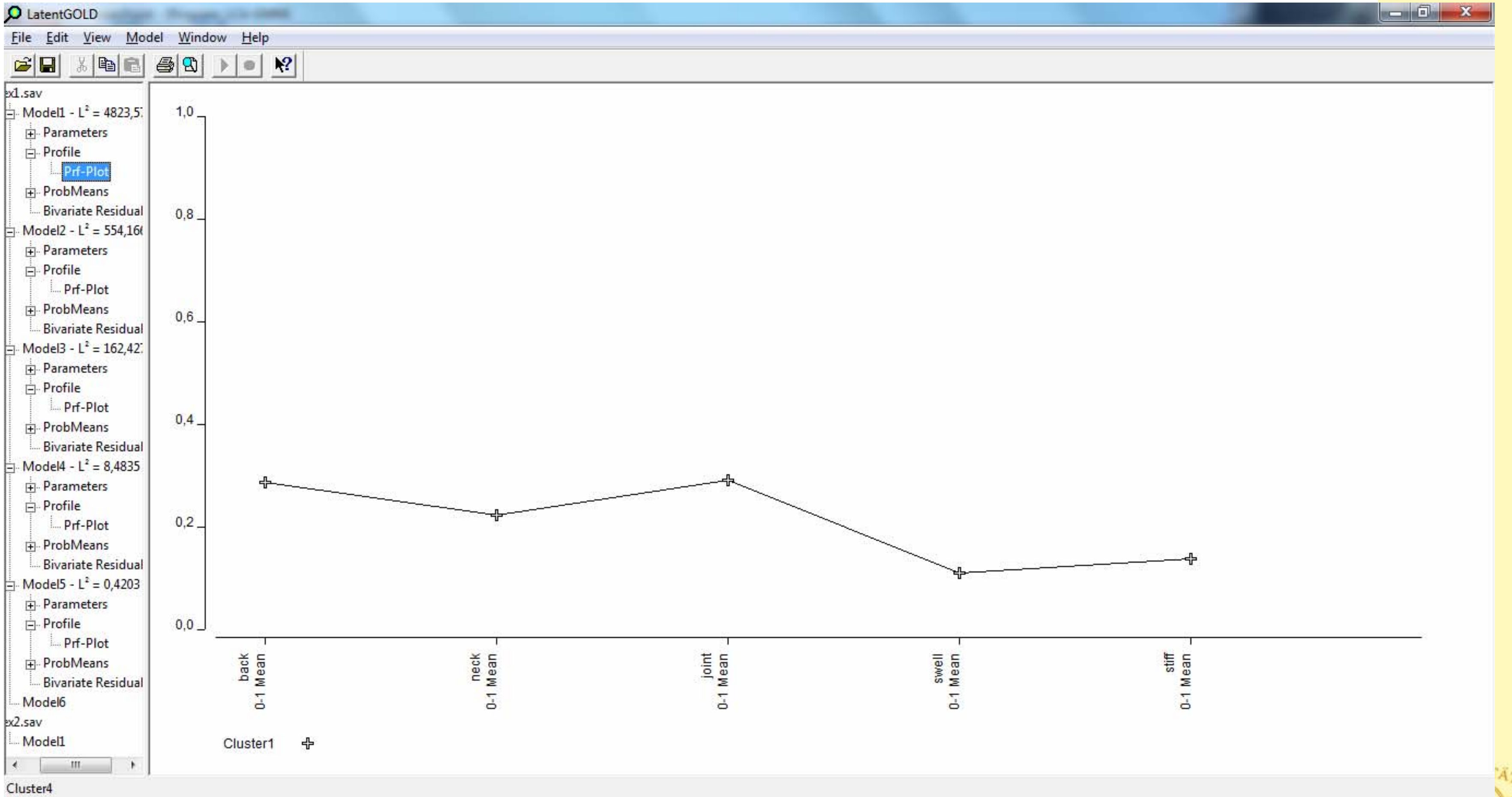
4) lokale stochastische Unabhängigkeit (Abhängigkeiten zwischen Items entstehen nur durch die Unkenntnis der latenten Variablen)

$$p(\underline{x}) = \sum_{g=1}^{G} \pi_g \prod_{i=1}^{k} \pi_{ig}^{x_i} \left(1 - \pi_{ig}\right)^{1-x_i}$$

Rost, J. (2004). Lehrbuch Testtheorie Testkonstruktion (2. Aufl.). Bern: Hans Huber

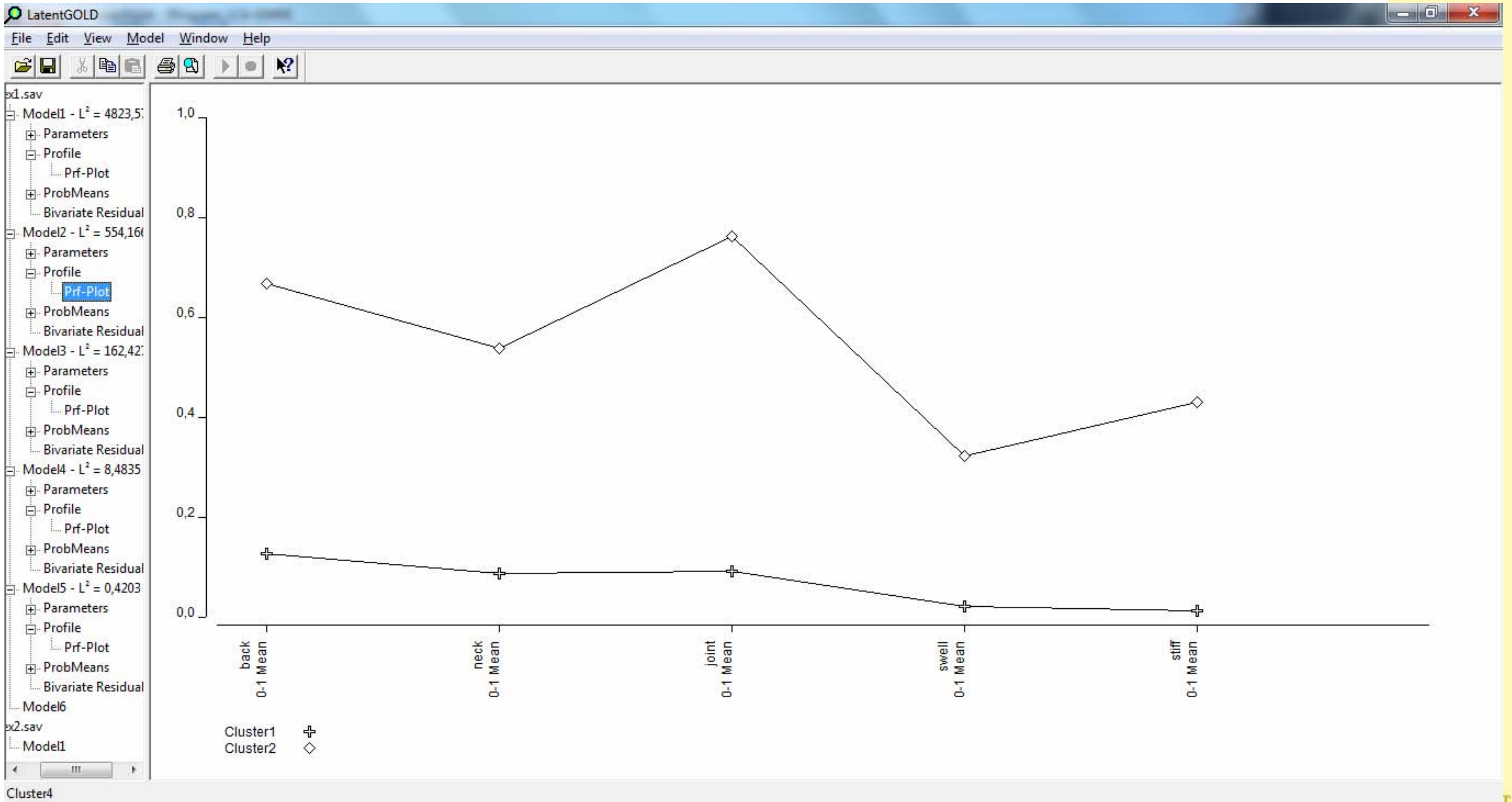
Beispiel zu fünf Schmerzsymptomen aus LatentGOLD

-- eine latente Klasse (=beobachtete Prozentanteile) --



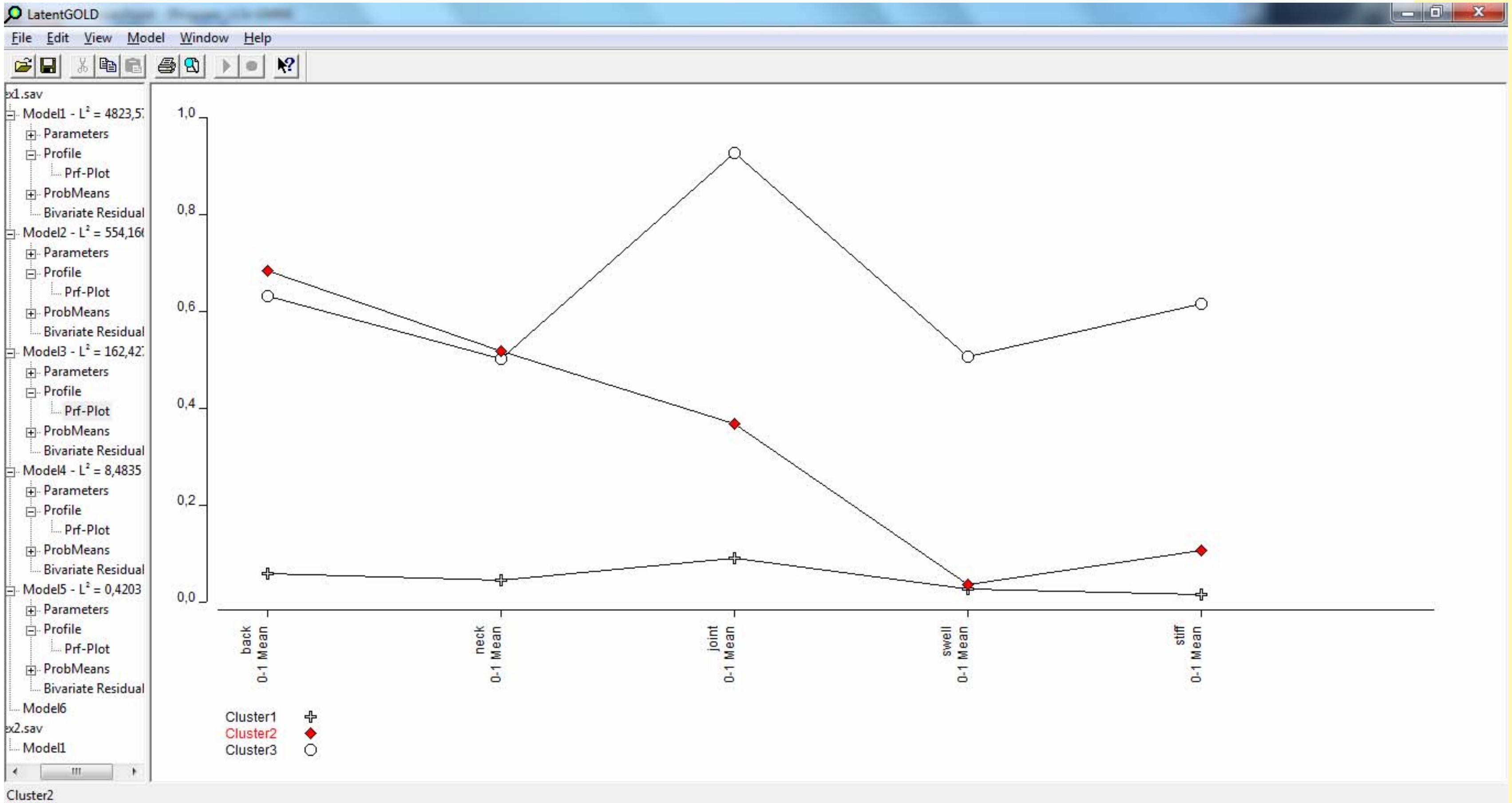
Beispiel zu fünf Schmerzsymptomen aus LatentGOLD

-- zwei latente Klassen --



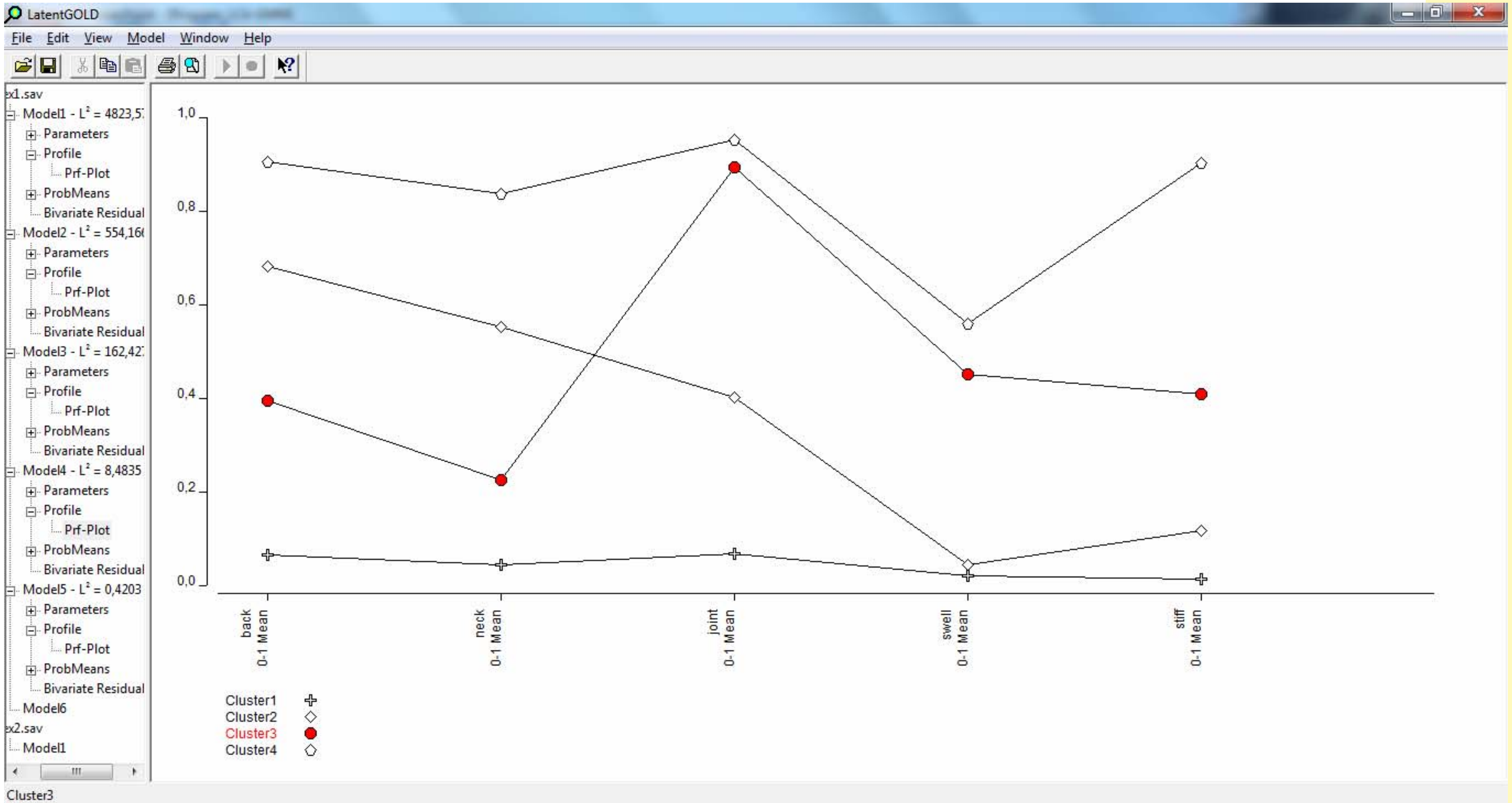
Beispiel zu fünf Schmerzsymptomen aus LatentGOLD

-- drei latente Klassen --



Beispiel zu fünf Schmerzsymptomen aus LatentGOLD

-- vier latente Klassen --



Vorteile gegenüber Clusteranalyse:

1) LCA benötigt (im Unterschied zur CA)
   - keine Festlegung eines Ähnlichkeitsmaßes
   - kein Algorithmus für Clusterung
   - keine Definition der Distanz zwischen Clustern

2) probabilistische Zuordnung (Zuordnungs-wk.)

3) höherdimensionale Zusammenhänge werden abgebildet (CA: nur bivariat)

How many classes?

„Problem“:

number of classes is not a model parameter.

Several conceptual approaches are helpful / necessary

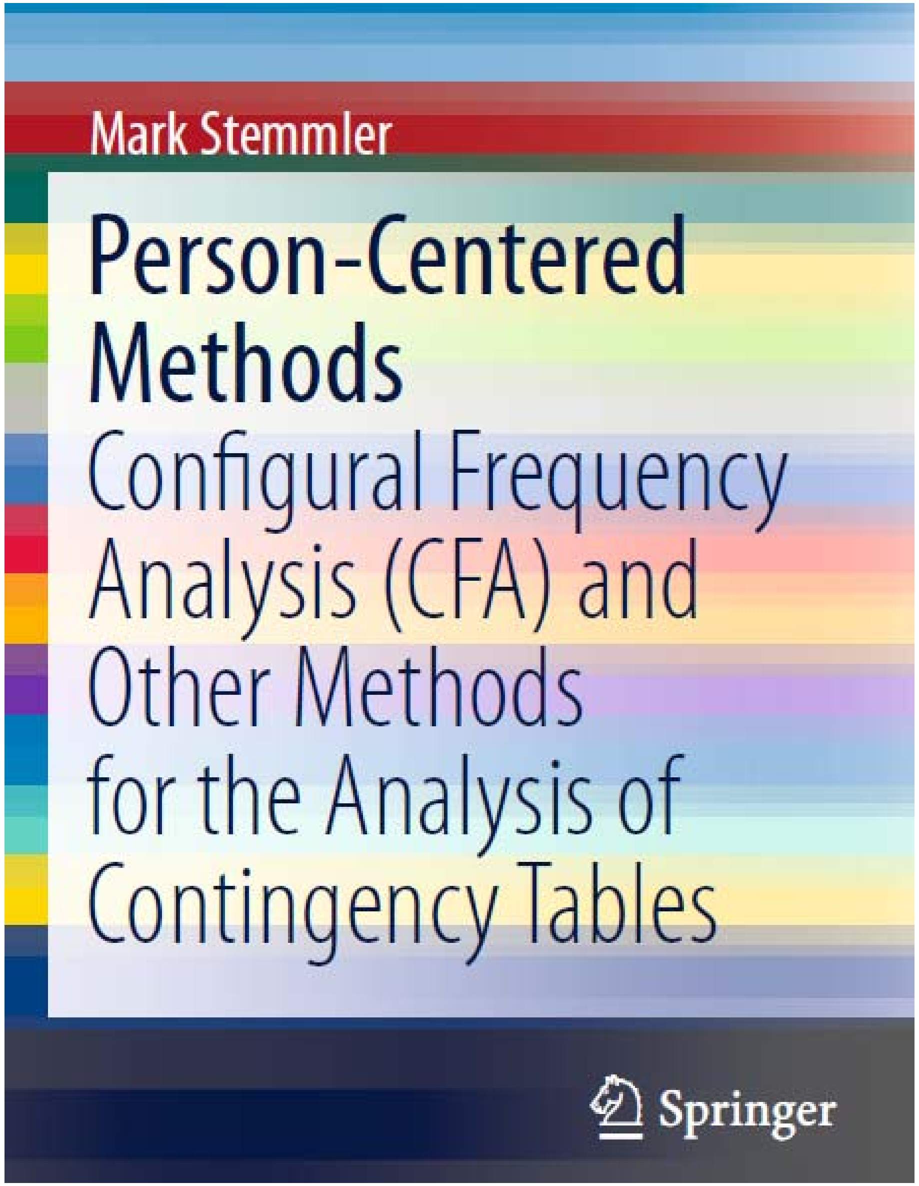
- fit criteria
- information criteria (IC), e.g. Bayesian IC (BIC), AIC
- parsimony, theoretical justification, clinical interpretability
- high membership probabilities, classes not too small
- (bootstrap) likelihood ratio tests

Beispiel zu fünf Schmerzsymptomen aus LatentGOLD

-- Überblick zu Modell-Fit von 1-5 latenten Klassen --

File name:	C:\Program Files\LatentGOLD4.5\DemoData\ex1.sav								
File size:	884 bytes								
File date:	2000-Feb-25	19:21:36							
		LL	**BIC(LL)**	**CAIC(LL)**	**Npar**	**L²**	**df**	**p-value**	**Class.Err.**
Model1	1-Cluster	-17729,2147	35502,8120	35507,8120	5	4823,5753	26	3,0e-1016	0,0000
Model2	2-Cluster	-15594,5103	31286,6627	31297,6627	11	554,1667	20	1,3e-104	0,0698
Model3	3-Cluster	-15398,6409	30948,1831	30965,1831	17	162,4278	14	2,3e-27	0,1417
Model4	4-Cluster	-15321,6688	30847,4980	30870,4980	23	8,4835	8	0,39	0,1727
Model5	5-Cluster	-15317,6372	30892,6941	30921,6941	29	0,4203	2	0,81	0,1650
Model6	5-Cluster								

Betrachtung von Antwortmustern auf der manifesten Ebene
-- Konfigurationsfrequenzanalyse (Gustav A. Lienert) --



Berechnung über:
R package ***confreq***
(Jörg Heine und Mark Stemmler)

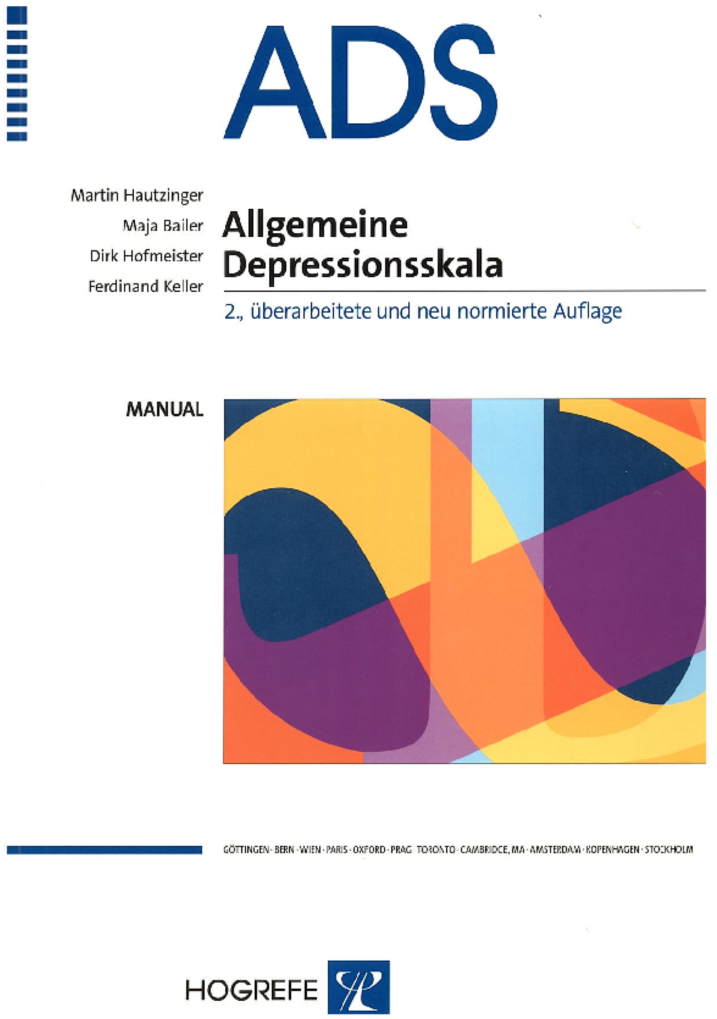
LCA zur Identifizierung einer „unskalierbaren“ Gruppe

Center of Epidemiological Studies-Depression (CES-D) (in zahlreichen Sprachen verfügbar)

Dt. Version: Allgemeine Depressionsskala

enthält negativ und positiv formulierte Items

Fragestellung: gibt es eine Gruppe von Person die die Polung übersieht (aus Unaufmerksamkeit, kognitiver Beeinträchtigung o.ä.)



ADS (Allgemeine Depressionsskala), insgesamt 20 Fragen

Bitte kreuzen Sie bei den folgenden Aussagen die Antwort an, die Ihrem Befinden während der letzten Woche am besten entspricht/entsprochen hat.

Antwortmöglichkeiten:

0	= selten oder überhaupt nicht	(weniger als 1 Tag)
1	= manchmal	(1 bis 2 Tage lang)
2	= öfter	(3 bis 4 Tage lang)
3	= meistens, die ganze Zeit	(5 bis 7 Tage lang)

Achtung: bei Summenbildung Fragen 4, 8, 12 und 16 umpolen

Während der letzten Woche ...

	selten oder überhaupt nicht 0	manchmal 1	öfter 2	meistens, die ganze Zeit 3
1. haben mich Dinge beunruhigt, die mir sonst nichts ausmachen.	◯	◯	◯	◯
2. hatte ich kaum Appetit.	◯	◯	◯	◯
3. konnte ich meine trübsinnige Laune nicht loswerden, obwohl mich meine Freunde/Familie versuchten, aufzumuntern.	◯	◯	◯	◯
4. kam ich mir genau so gut vor wie andere.	◯	◯	◯	◯
5. hatte ich Mühe, mich zu konzentrieren.	◯	◯	◯	◯
6. war ich deprimiert/niedergeschlagen.	◯	◯	◯	◯
7. war alles anstrengend für mich.	◯	◯	◯	◯
8. dachte ich voller Hoffnung an die Zukunft.	◯	◯	◯	◯
9. dachte ich, mein Leben ist ein einziger Fehlschlag.	◯	◯	◯	◯

ADS: 5-cluster Lösung bei Normstichprobe Erwachsener

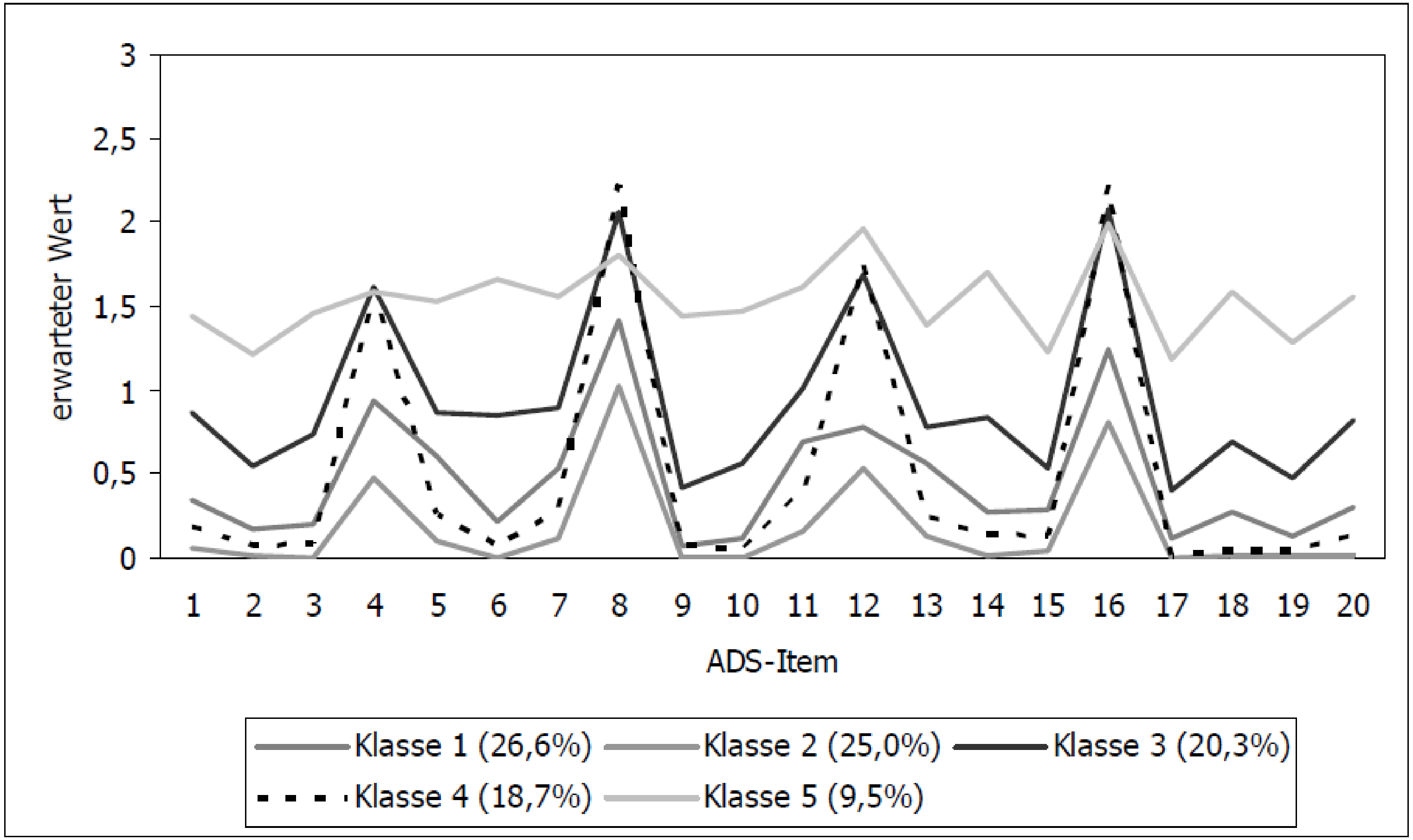


Abbildung 2: Profil der Erwartungswerte für die fünf Klassen in der Normstichprobe (n = 2343)

Aus: Keller, F. (2012). Latent-Class- und Mixed-Rasch-Modelle zur Identifizierung skalierbarer und unskalierbarer Personengruppen in der Allgemeinen Depressionsskala. In: W. Kempf & R. Langeheine (Hrsg.): *Item-Response-Modelle in der sozialwissenschaftlichen Forschung* (S. 171-188). Berlin: Verlag irena regener.

LCA zur Identifizierung einer „unskalierbaren“ Gruppe

Allgemeine Depressionsskala

enthält negativ und positiv formulierte Items

Fragestellung: gibt es eine Gruppe von Personen, die die Polung übersieht (aus Unaufmerksamkeit, kognitiver Beeinträchtigung .o.ä.)?

Ergebnis:

Ja, eine Klasse (auch in anderen Studien replizierbar)

Aber: nur Tendenz

Alternative Erklärungen:

- Unterscheidung in positiv- und negativ-Faktor
- ‚realistische‘ Klasse: wenig Negativsymptomatik, aber auch wenig positives Erleben

ISAAC LCA - Analysis

Asthma phenotypes identified by latent class analysis in the ISAAC phase II Spain study

Clinical & Experimental Allergy, 43, 223–232

G. Weinmayr1*, F. Keller2*, A. Kleiner1, J. B. du Prel1, L. Garcia-Marcos3, J. Batlle´s-Garrido4, G. Garcia-Hernandez5, M. M. Suarez-Varela6, D. P. Strachan7 and G. Nagel1

Spanish Phase Two data

Objectives

- Classify potential phenotypes based on cross-sectional symptoms data, only
- Identification with LCA and ISAAC questions
- Explore associations of these phenotypes with
  - Objective markers
  - Other allergic disease
  - Parental history
  - Medication use

Variables used

Cough and phlegm

1. In the last 12 months, has your child usually seemed congested in the chest or coughed up phlegm (mucus) with colds? Yes ❑ No ❑

2. In the last 12 months, has your child usually seemed congested in the chest or coughed up phlegm (mucus) when he/she did not have a cold? Yes ❑ No ❑

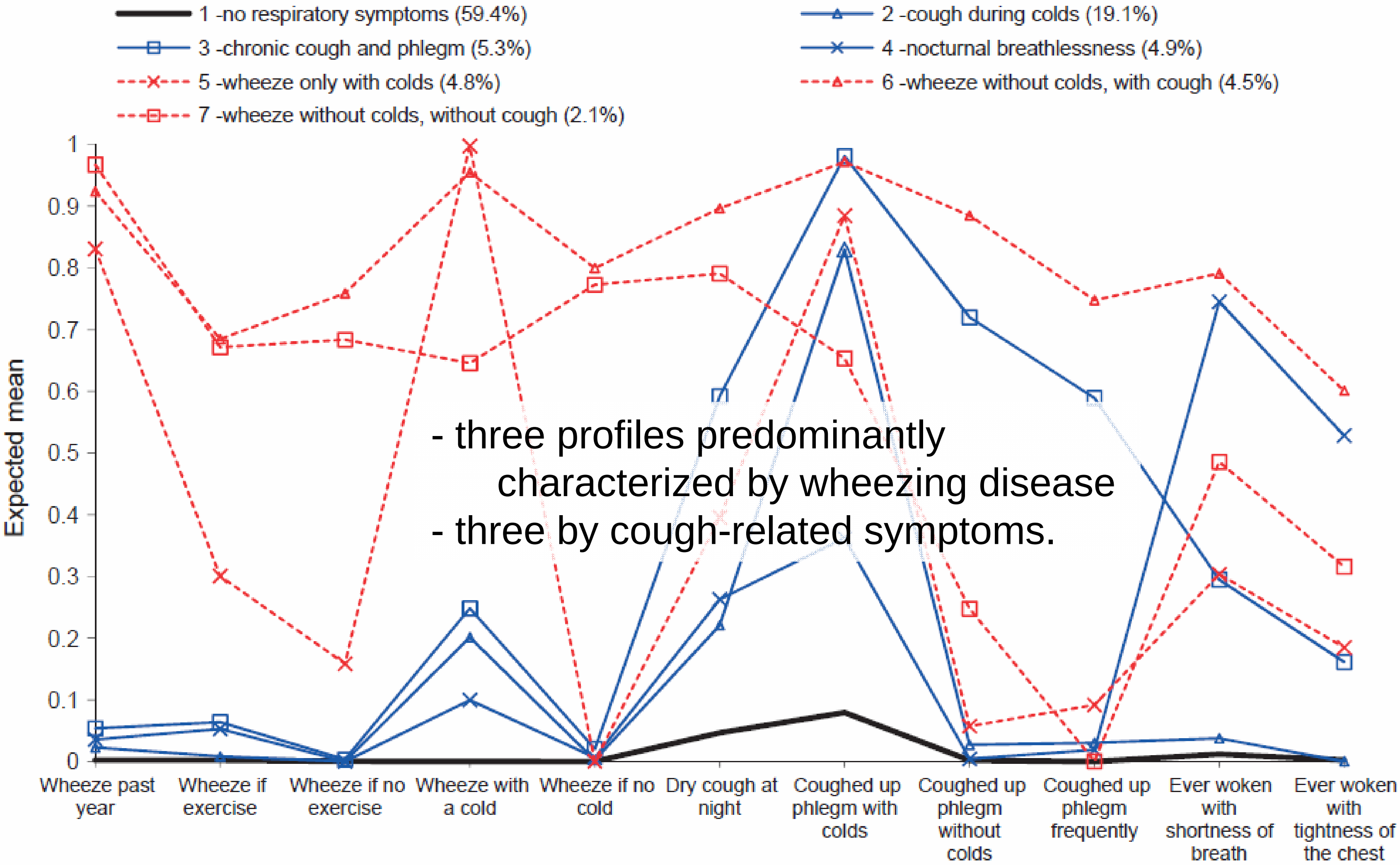
IF YOU HAVE ANSWERED "NO" TO BOTH OF THESE QUESTIC PLEASE SKIP QUESTIONS 3 & 4.

---

3. Does your child seem congested in the chest or cough up phlegm (mucus) on most days (4 or more days a week) for as much as 3 months of the year? Yes ❑ No ❑

IF YOU HAVE ANSWERED "NO", PLEASE SKIP QUESTION 4.

Example: Latent class analysis (cross-sect.)



- three profiles predominantly characterized by wheezing disease
- three by cough-related symptoms.

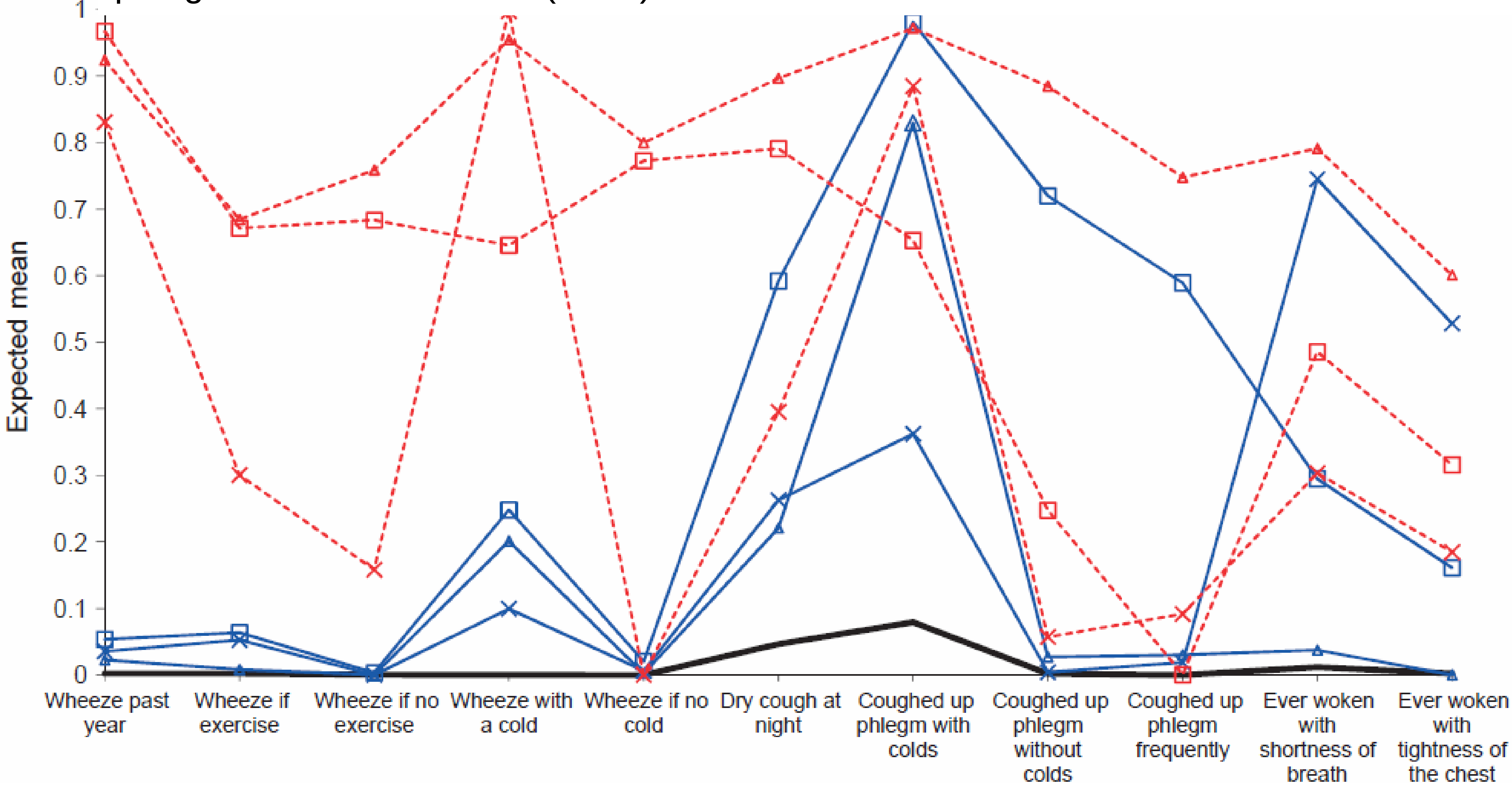
Fig. 1. Estimated prevalences (expected mean) of the respiratory symptoms in, and estimated class sizes (in brackets) of the seven phenotypes identified by latent class analysis.

Weinmayr/Keller/Kleiner,….., Nagel (2013). Asthma phenotypes identified by latent class analysis in the ISAAC phase II Spain Study. *Clinical & Experimental Allergy, 43,* 223-232.

• **Wheeze only with colds** (4.8%): high frequency of children wheezing and coughing in presence of a cold (100% and 87.6% respectively), but not if no cold (0% and 4.8%)

• **Wheeze without colds, with cough** – Chronic wheeze with concomitant cough and phlegm (4.4%): all symptoms are frequently reported (from 59.4% for 'ever woken with tightness of the chest' to 97.7% for 'coughed up phlegm with colds')

• **Wheeze without colds, without cough** – chronic wheeze with no cough and phlegm unless with a cold (2.0%).



Differences in severity of LCA asthma phenotypes

ISAAC

ulm

Class	5		6		7	
Labels	Wheeze only with colds		Wheeze without colds, with cough		Wheeze without colds, without cough	
	n	% (95% CI)	*n*	% (95% CI)	*n*	% (95% CI)
Age (years) (mean)	186	9.80 (9.69–9.91)	170	9.65 (9.54–9.77)	76	9.67 (9.49–9.86)
Sex (males)	186	51.1 (43.8–58.3)	170	60.0 (52.6–67.4)	76	61.8 (50.7–73.0)
Asthma ever	184	36.4 (29.4–43.4)	168	70.2 (63.3–77.2)	75	64.0 (52.9–75.1)
Inhaled corticosteroids	103	29.1 (20.2–38.1)	124	41.1 (32.4–49.9)	65	36.9 (24.9–49.0)
Wheeze severity§						
At least four wheezing attacks	185	4.9 (1.7–8.0)	167	35.9 (28.6–43.3)	74	31.1 (20.3–41.9)
Wheeze disturbing sleep	185	3.2 (0.7–5.8)	167	18.0 (12.1–23.9)	74	6.8 (0.9–12.6)
Wheeze limiting speech	181	4.4 (1.4–7.4)	165	27.3 (20.4–34.1)	74	12.2 (4.5–19.8)

ISAAC

Associations

Table 3. Association

	Class 5 vs. 1 (186 vs. 2204) OR (95% CI)	Class 6 vs. 1 (170 vs. 2204) OR (95% CI)	Class 7 vs. 1 (76 vs. 2204) OR (95% CI)	Wheeze past year vs. no wheeze OR (95% CI)	Severe wheeze vs. no wheeze OR (95% CI)
Age (years) (mean)	1.15 (0.92–1.43)	0.95 (0.75–1.19)	0.85 (0.60–1.20)	0.91 (0.80–1.04)	0.98 (0.79–1.21)
Sex (males)	1.11 (0.82–1.50)	1.61 (1.17–2.21)	1.72 (1.08–2.75)	1.33 (1.11–1.58)	1.30 (0.96–1.75)
Asthma ever	25.86 (17.06–39.19)	119.68 (74.51–192.23)	85.84 (48.73–151.22)	17.06 (13.80–21.08)	40.88 (28.65–58.32)
Inhaled corticosteroids	3.08 (0.60–15.65)	4.24 (0.81–22.19)	3.86 (0.56–26.86)	1.73 (1.08–2.76)	2.30 (1.33–3.96)
Other symptoms					
Rhinoconjunctivitis	4.43 (3.07–6.39)	20.00 (13.86–28.84)	9.99 (6.10–16.38)	5.12 (4.21–6.22)	8.33 (6.04–11.49)
Rhinitis without conjunctivitis	1.63 (1.02–2.61)	1.84 (1.15-2.94)	3.43 (1.95–6.01)	1.55 (1.21–2.00)	1.31 (0.85–2.02)
Eczema by examination	2.06 (0.84–5.04)	3.20 (1.47–6.95)	2.69 (0.90–8.05)	2.36 (1.43–3.91)	2.48 (1.12–5.49)
Parental asthma	1.44 (0.92–2.25)	2.99 (2.03–4.40)	3.17 (1.83–5.49)	2.12 (1.69–2.66)	2.44 (1.70–3.50)
Parental allergies	1.36 (1.00–1.84)	2.16 (1.57–2.97)	3.03 (1.89—4.86)	1.73 (1.45–2.06)	1.90 (1.41–2.56)
Skin-prick test positivity	2.52 (1.76–3.61)	5.00 (3.38–7.39)	10.90 (5.84–20.32)	3.60 (2.92–4.45)	5.47 (3.78–7.92)
Total IgE (kU/l) (geometric mean)	1.27 (1.02–1.58)	1.92 (1.55–2.38)	2.10 (1.57–2.82)	1.53 (1.36–1.73)	1.82 (1.51–2.19)
Spec. IgE ≥ 0.7 kU/l	2.54 (1.39–4.63)	6.38 (3.51–11.60)	13.11 (4.71–36.51)	3.60 (2.61–4.97)	4.74 (2.86–7.83)
Lung function and bro					
FEV1 predicted (5%*)	0.93 (0.82–1.05)	0.99 (0.87–1.12)	1.05 (0.91–1.21)	0.99 (0.93–1.05)	0.98 (0.89–1.09)
BHR yes/no†	4.52 (2.17–9.42)	6.40 (2.90–14.09)	13.73 (4.43–42.57)	5.95 (3.89–9.10)	7.02 (2.98–16.57)

*The effect is calculat

†Fall ≥ 15% compar

Growth Mixture Models: „Traditions“

From antisocial behaviour/delinquency research:

Latent class growth model **(LCGA)** (Nagin, Tremblay, Jones):

Individual growth trajectories within a class are homogeneous

From latent variable modelling:

Growth mixture modelling (**GMM**) (Muthen):

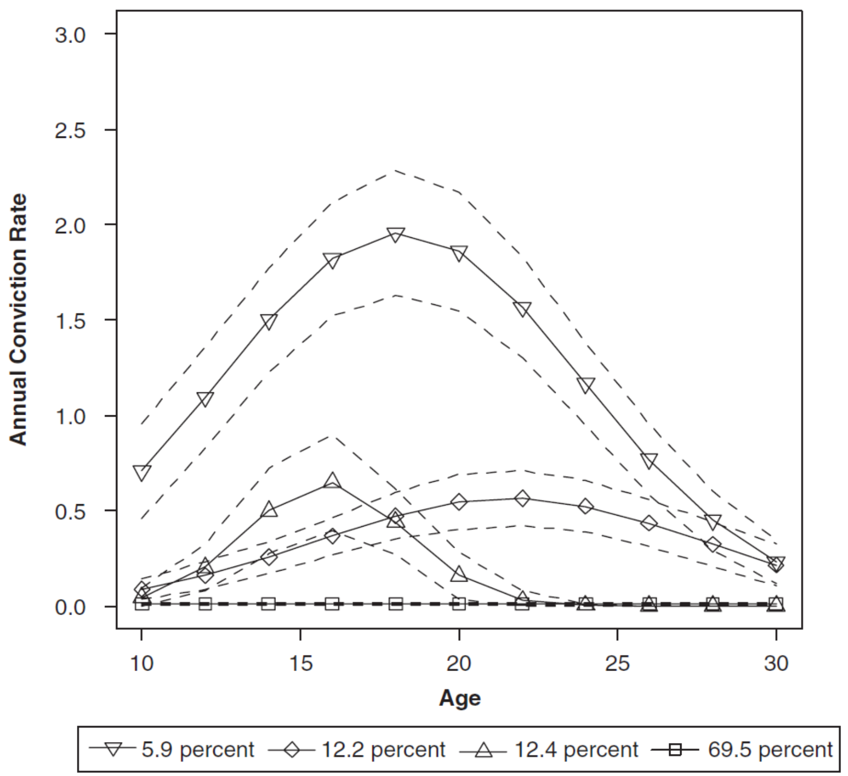
Allows for within-class variation (and freeing/fixing of other parameters)

- Heterogeneity in growth is captured through trajectory classes (categorical latent variable) **and** random effects *within* class.

Example: latent class growth model for delinquency

Jones, Nagin / Group-Based Trajectory Modeling 547

Figure 1
Annual Conviction Rate Versus Age: Four-Group Poisson Model

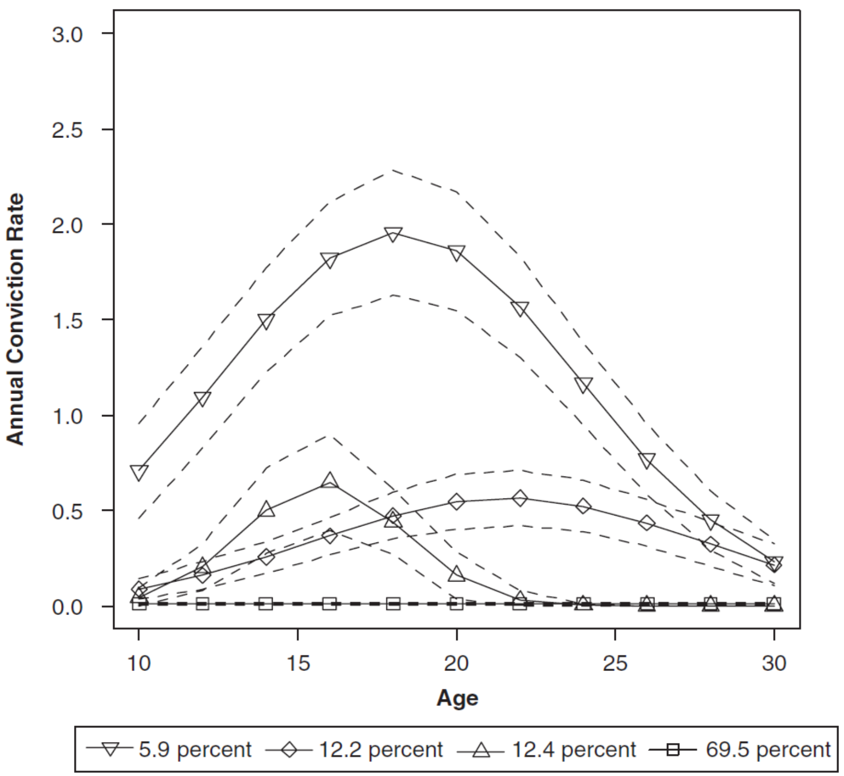


Advances in Group-Based Trajectory Modeling and an SAS Procedure for Estimating Them
Bobby L. Jones and Daniel S. Nagin
Sociological Methods Research 2007; 35; 542

Example: latent class growth model for delinquency

Jones, Nagin / Group-Based Trajectory Modeling 547

Figure 1
Annual Conviction Rate Versus Age: Four-Group Poisson Model

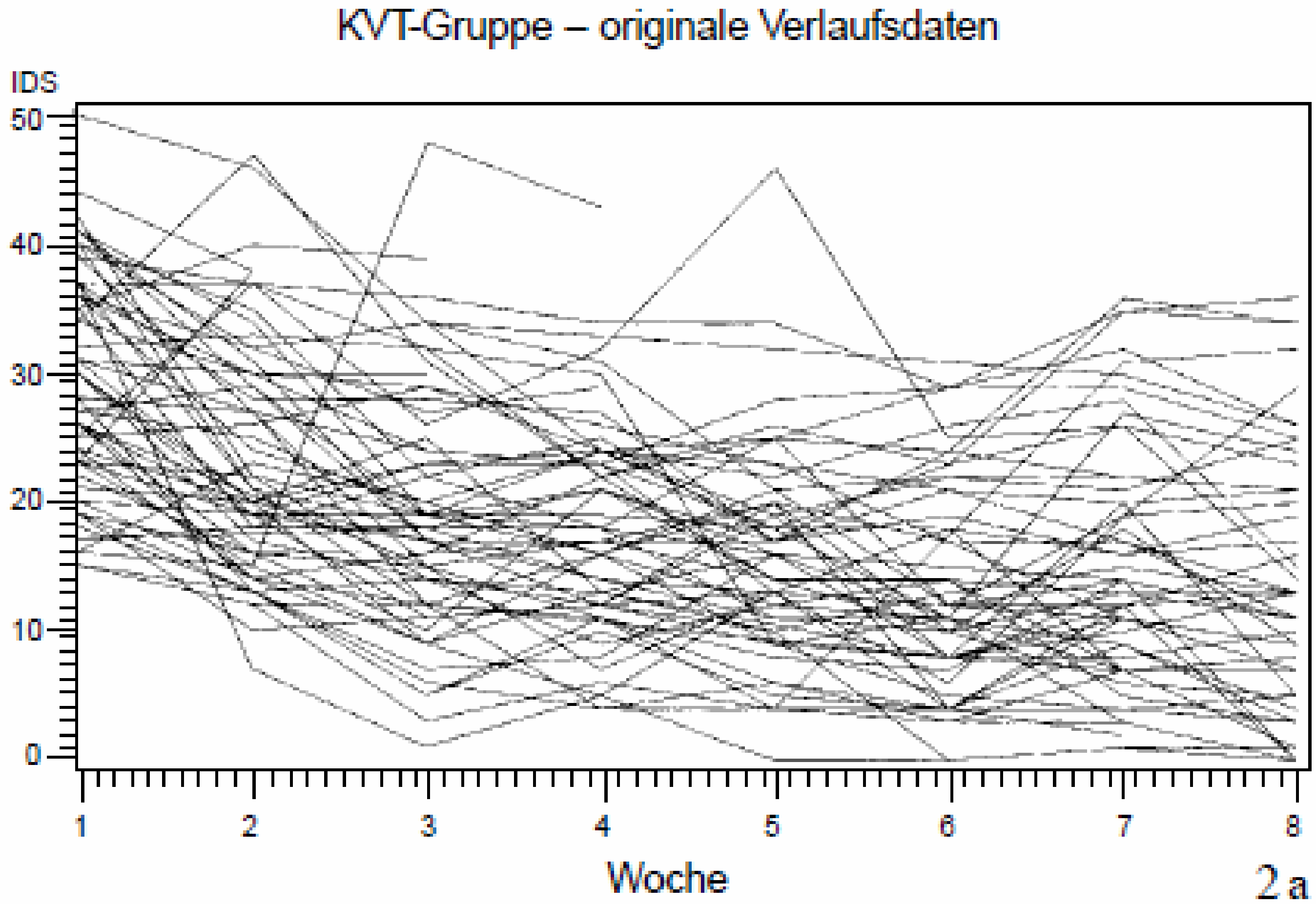


Methodisch: Zählvariable (counts)
Poisson-Modell

Bei Delinquenz: viele Nullen, daher meist zero-inflated Poisson (ZIP)

Advances in Group-Based Trajectory Modeling and an SAS Procedure for Estimating Them
Bobby L. Jones and Daniel S. Nagin
Sociological Methods Research 2007; 35; 542

Depression therapy study Hautzinger/de Jong (Raw data, CBT group only) „Spaghetti-Plot“



Aus: Keller, F. (2003): Analyse von Längsschnittdaten: Auswertungsmöglichkeiten mit hierarchischen linearen Modellen. *Zeitschrift für Klinische Psychologie und Psychotherapie, 32,* 51-61

Single population vs. unobserved subpopulations

Growth curves as random-effects models:

Growth curve parameters, e.g. intercept and slope, vary across individuals.

However,

- a **single population** with common parameters is assumed, or
- **subgroups are known**, e.g. treatment condition, gender…

Growth mixture modelling:

allows for differences in growth parameters across **unobserved subpopulations**, resulting in separate growth models for each subpopulation (latent class).

SAS Proc Traj - Home - Mozilla Firefox

Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe

http://www.contrib.andrew.cmu.edu/~bjones/index.htm

Kostenlose Hotmail | Links anpassen | Home - Universitätskli... | Windows Media | Windows

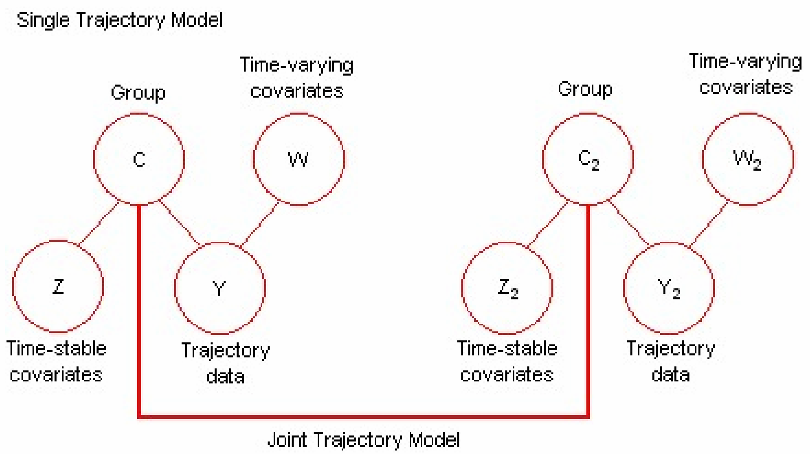
PROC TRAJ

- Home
- Download
- Examples
- Documentation

PROC TRAJ is a SAS procedure that fits a discrete mixture model to longitudinal data. The model performs data sequence grouping, with different parameter values for the groups' data distribution. Groupings may identify distinct subpopulations. Alternatively, groups may represent distribution components approximating an unknown (possibly complex) data distribution.

Supported distributions are: censored (or regular) normal, zero inflated (or regular) Poisson, and Bernoulli distributions (logistic model). The censored normal model is useful for psychometric scale data, the zero inflated Poisson model useful for count data with extra zeros, and the Bernoulli model useful for 0/1 data. The model is appropriate for data with average values changing smoothly as a function of the dependent variable (time, age, ...). Some sharp changes can be handled through the inclusion of time dependent covariates.

MODEL STRUCTURE: Data sequences, Y, with similar shapes are grouped in a model-based manner. The probability of group membership can be a function of time stable covariates (risk factors), Z. Time dependent covariates, W, can further influence trajectories with effects differing by group, C. A trajectory model for two sets of dependent variables (joint trajectory modeling) is also supported. The model is illustrated in the figure below.

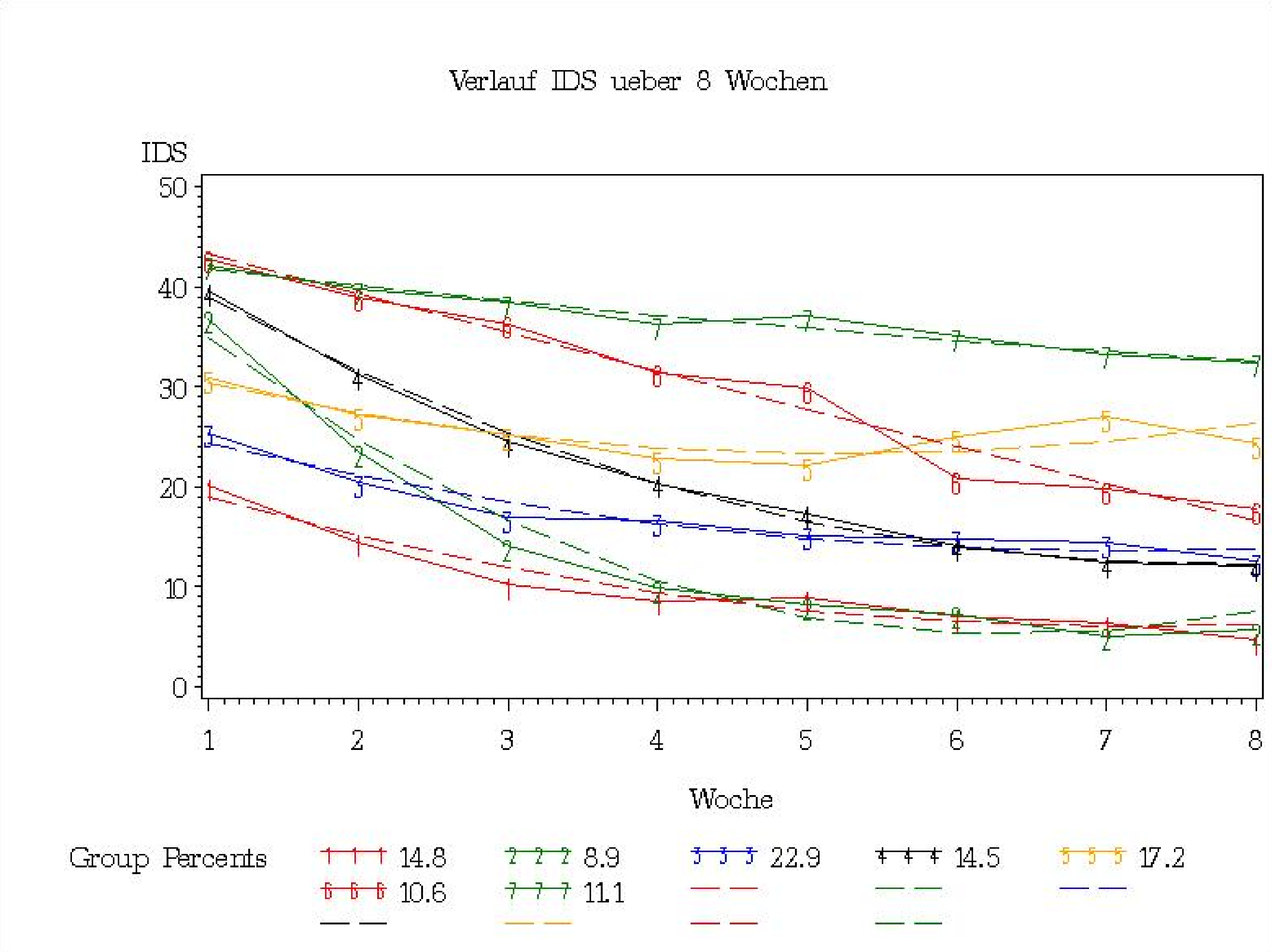


Downloads: Jones, B., Nagin, D., & Roeder, K. "A SAS Procedure Based on Mixture Models for Estimating Developmental Trajectories," Sociol Method Res, 2001, 29: 374-393.

Jones, B., & Nagin, D. "Advances in Group-Based Trajectory Modeling and a SAS Procedure for Estimating Them," submitted

Fertig

Trajectories for 7 classes (LCGA, estimated and sample means)



Note: Current version is 7.3, but Flash for presenting the version no longer allowed…

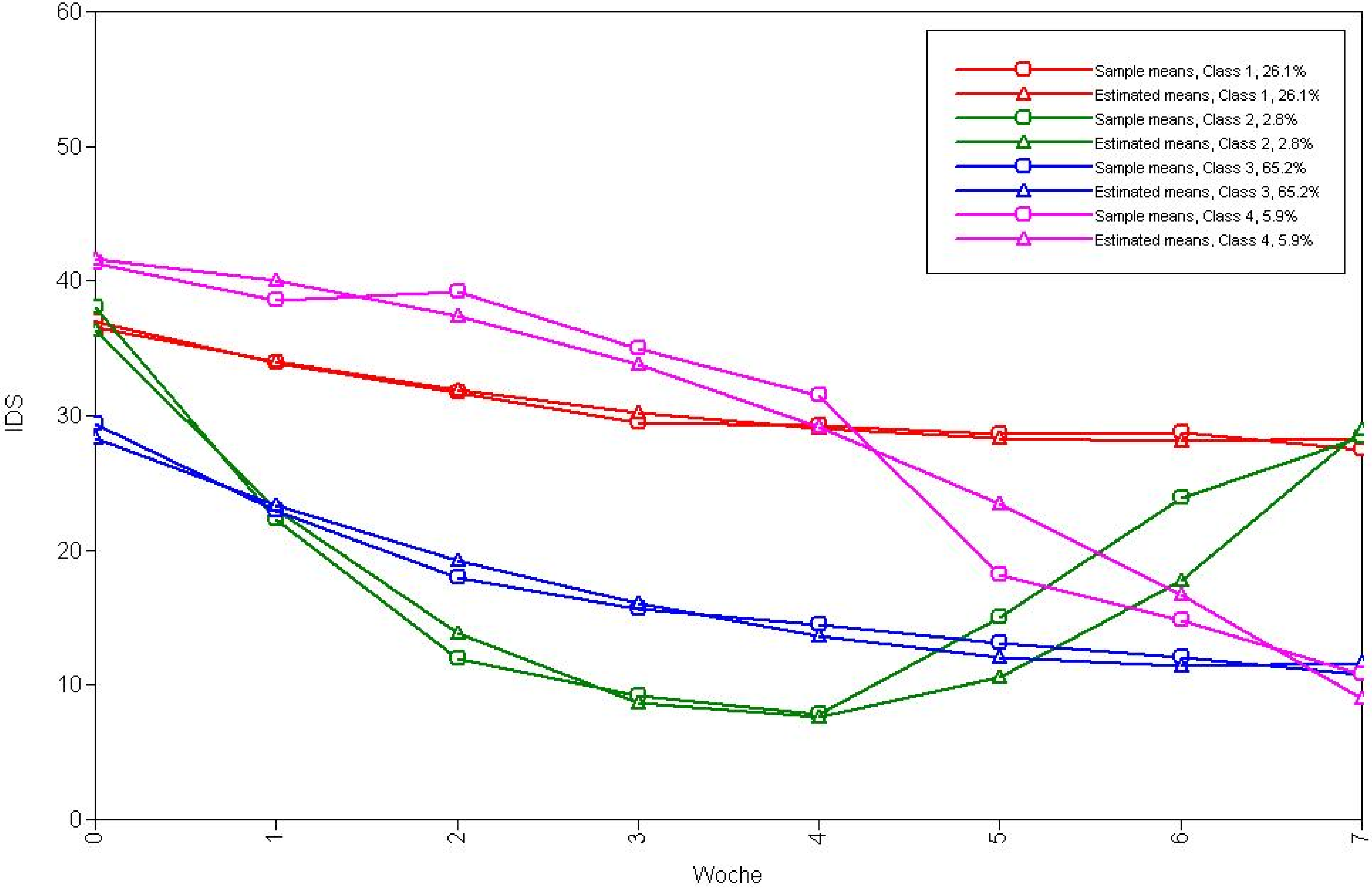
Latent Class Growth model: Example syntax in M*plus*

```
TITLE: course of IDS over 8 weeks, 7 classes, no random effects;
DATA: FILE IS ids.dat;

VARIABLE: NAMES ARE STUDIE THGRUPPE PATNR BDIT1SUM BDIT3SUM
           HA21SUT3 y1 y2 y3 y4 y5 y6 y7 y8 ;
           USEVAR = y1-y8;
           IDVARIABLE=patnr;
           MISSING ARE ALL (999);
           CLASSES=C(7);

ANALYSIS: TYPE=Mixture missing;
           estimator=ML;
MODEL:
  %OVERALL%
       i s q | y1@0 y2@1 y3@2 y4@3 y5@4 y6@5 y7@6 y8@7;
       i@0; s@0; q@0;    ! Set variances to zero, yields LCGA model ;

OUTPUT: RESIDUAL stand  tech8;
PLOT: Type is plot1 plot2 plot3;  series = y1-y8 (s);
```

How many classes?
Fit and information criteria for GMM with 3 – 5 classes

number of classes	Log Lik.	# of parameters	AIC	BIC	ssaBIC	entropy
3	-7441.8	25	14933.7	15026.6	14947.3	.804
4	-7424.8	29	14907.6	**15015.4**	14923.4	.786
5	-7415.4	33	**14896.8**	15019.5	**14914.8**	.788

GMM, solution with 4 classes (M*plus*)



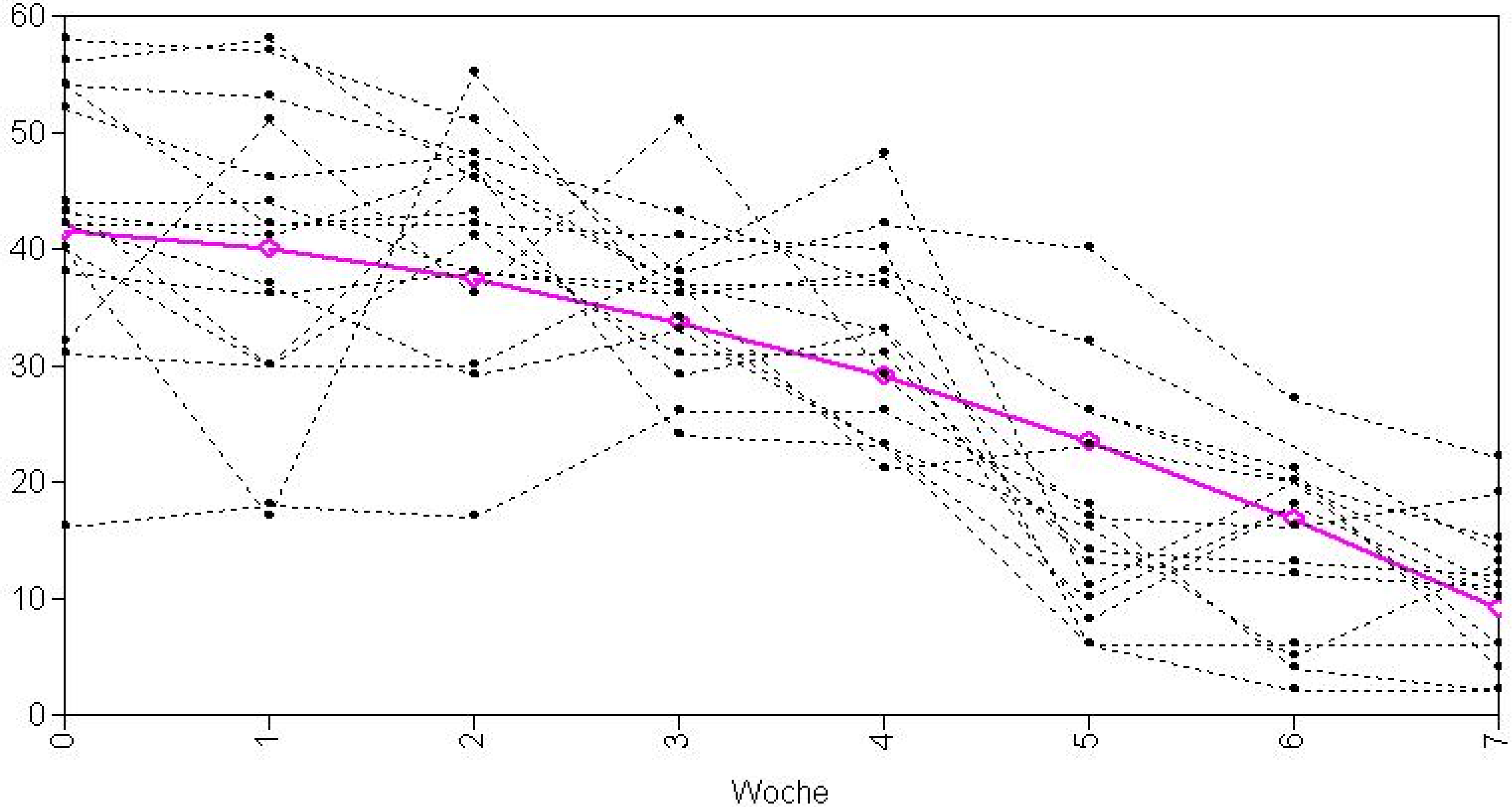
GMM, solution with 4 classes (M*plus*)
(class membership prob. and class attribution)

The SAS System

membership probabilities and modal class

Obs	patnr	prob1	prob2	prob3	prob4	class
1	1101	0.001	0.000	0.446	0.552	4
2	1102	0.006	0.000	0.991	0.002	3
3	1103	0.000	0.000	0.871	0.129	3
4	1104	0.000	0.000	0.994	0.006	3
5	1105	0.000	0.000	0.997	0.003	3
6	1106	0.049	0.000	0.624	0.327	3
7	1107	0.006	0.029	0.856	0.109	3
8	1108	0.021	0.000	0.974	0.005	3
9	1110	0.000	0.000	0.050	0.950	4
10	1113	0.000	0.004	0.973	0.022	3
11	1114	0.000	0.000	0.996	0.004	3
12	1115	0.000	0.974	0.022	0.004	2
13	1116	0.000	0.000	1.000	0.000	3
14	1118	0.026	0.000	0.540	0.434	3
15	1119	0.000	0.000	0.999	0.001	3
...	...	...				

Observed individual courses and estimated trajectory for class 4 („delayed response“)



Keller, F. & Hautzinger, M. (2007). Klassifikation von Verlaufskurven in der Depressionsbehandlung: Ein methodischer Beitrag. *Zeitschrift für Klinische Psychologie und Psychotherapie, 36,* 83-92

Example: Trajectories in BMI und Typ 2 Diabetes

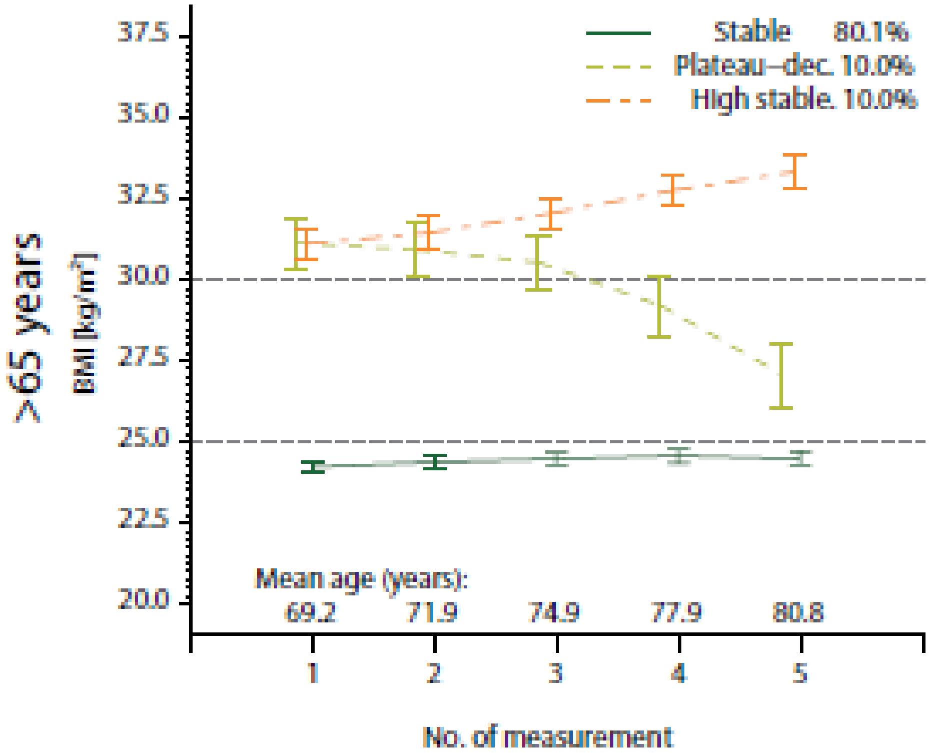
Vorarlberg Health Monitoring & Prevention Program (VHM&PP)

-a risk factor surveillance program in Vorarlberg
-included individuals (N=24,875) with at least five BMI measurements over a 12 year period.
-Trajectory classes were identified using growth mixture modeling (GMM) for:
- three predefined age groups (<50, 50-65, >65 years of age) - men and women separately.

Objectives:
-identify long-term BMI trajectories
-appraise them in different age groups regarding their impact on glucose impairment, diabetes related and overall mortality.

Example: Trajectories in BMI und Typ 2 Diabetes

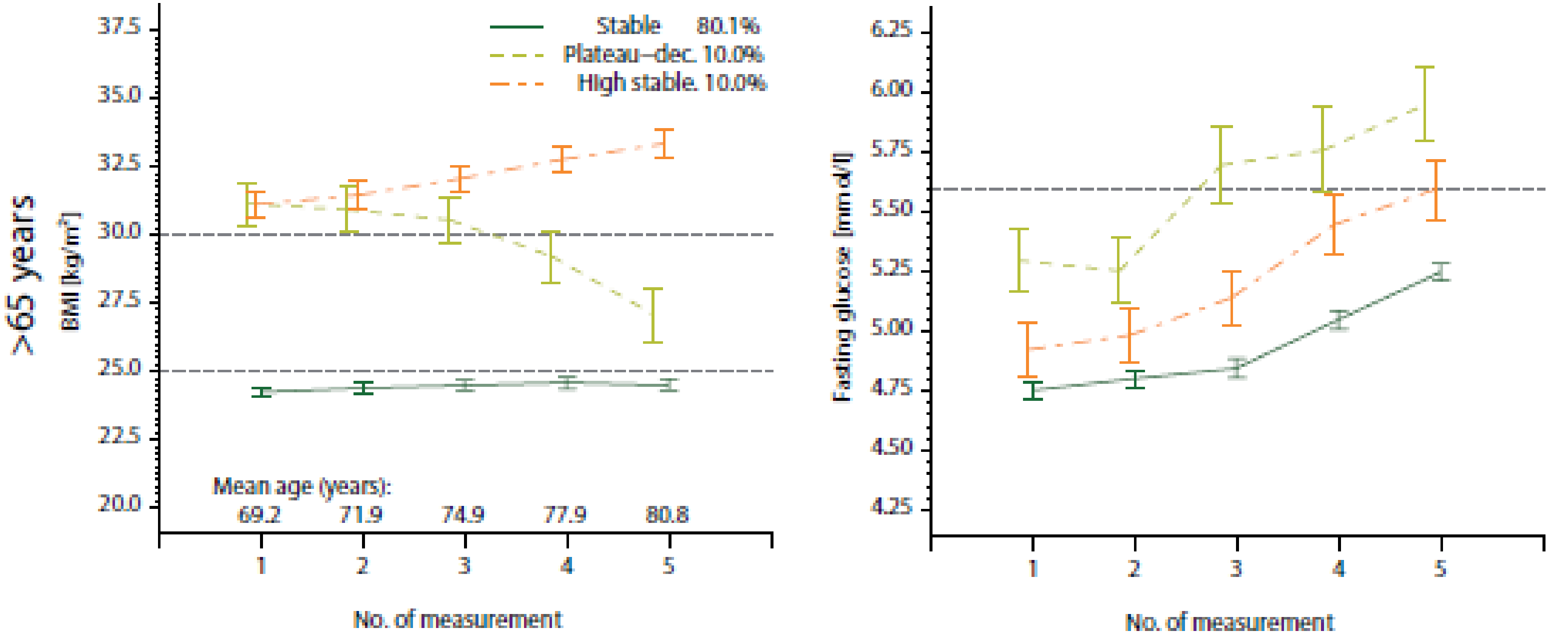
-- data from VHM&PP, age group > 65y and female --



Raphael Peter, Ferdinand Keller & Gabriele Nagel (in prep.): Body Mass Trajectories, Type 2 Diabetes Mellitus and Mortality in a Large Cohort of Adults – Results from the VHM&PP

Example: Trajectories in BMI und Typ 2 Diabetes

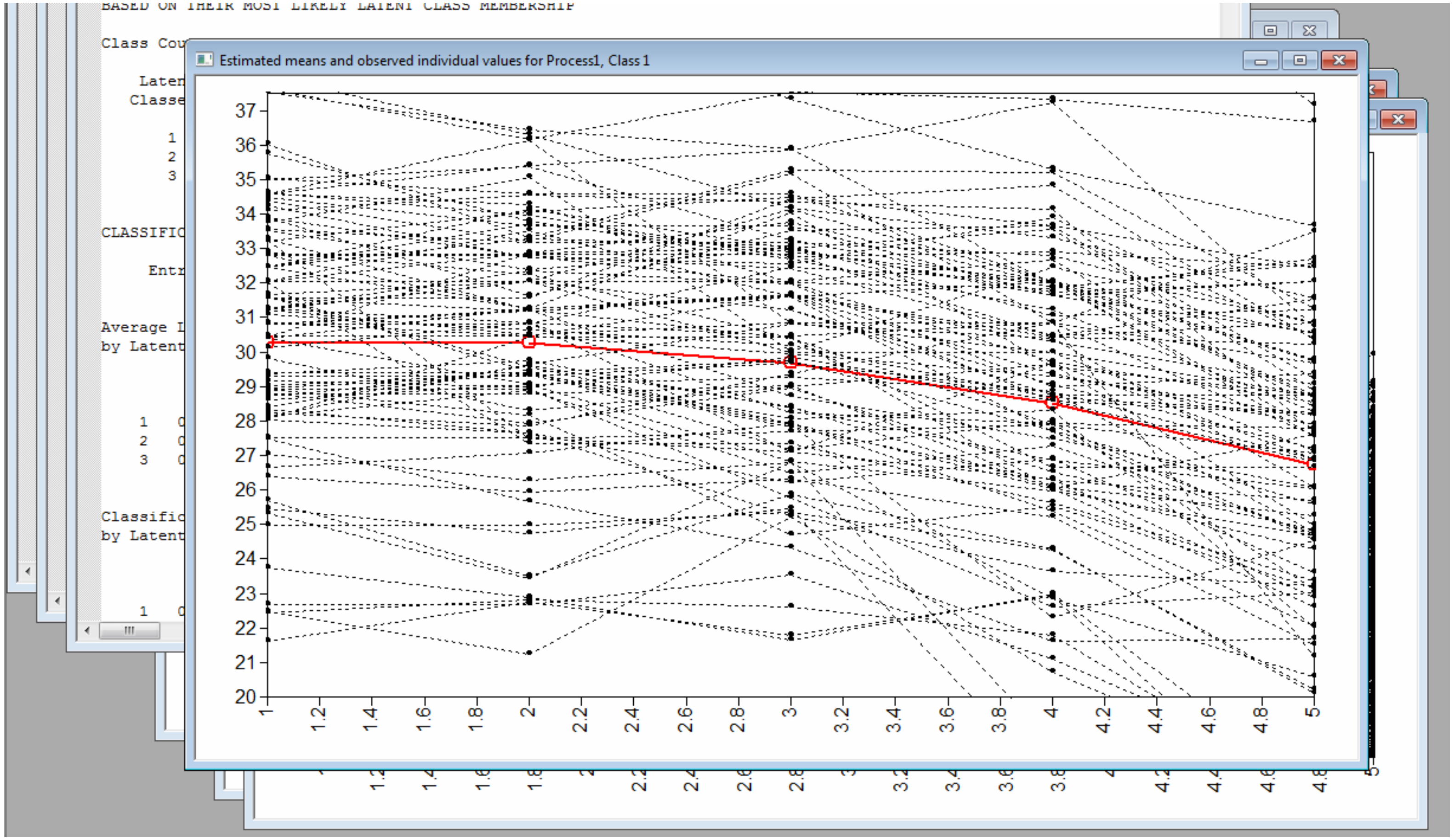
-- data from VHM&PP, age group > 65y and female --



Raphael Peter, Ferdinand Keller & Gabriele Nagel (in prep.): Body Mass Trajectories, Type 2 Diabetes Mellitus and Mortality in a Large Cohort of Adults – Results from the VHM&PP

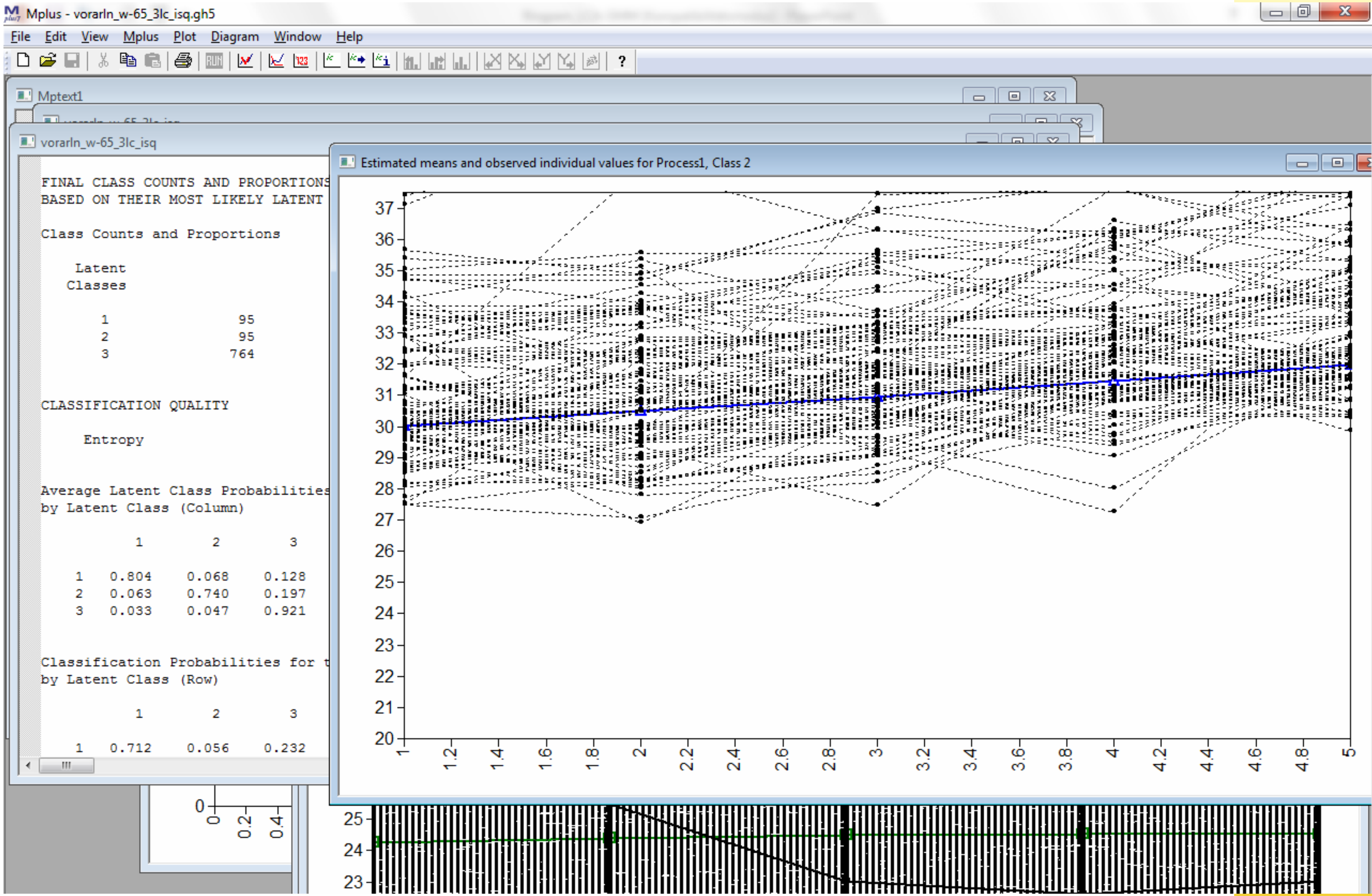
Example: Trajectories in BMI und Typ 2 Diabetes

-- estimated means and observed individual courses for class ‚plateau – decreasing' --



Example: Trajectories in BMI und Typ 2 Diabetes

-- estimated means and observed individual courses for class ‚high stable‘ --



Überblick und Beispiele dazu:

1) Normalverteilungen (Pearson)

2) LCA (dichotome / polytome Variablen)
   - Schmerzsymptome
   - Allg. Depressionsskala (Gruppe „unskalierbarer")
   - Phänotypen von Asthma

3) Growth mixture („trajectories")
   - count-Variable (Delinquenz im Jugendalter)
   - latent class growth model (Depressionsstudie)
   - growth mixture
     - Depressionsstudie
     - BMI-Verläufe in Vorarlberg-Studie

4) „limitations" und Ausblick

„Reality“ of latent classes

LATENT CLASS IDENTITY CRISIS

- Who are we?
- Are we “real”?
- How many are we?
- What defines us?
- What predicts us?
- What do we predict?

REITOX Academy, EMCDDA © Masyn & Muthén (13-Mar-2009) 60

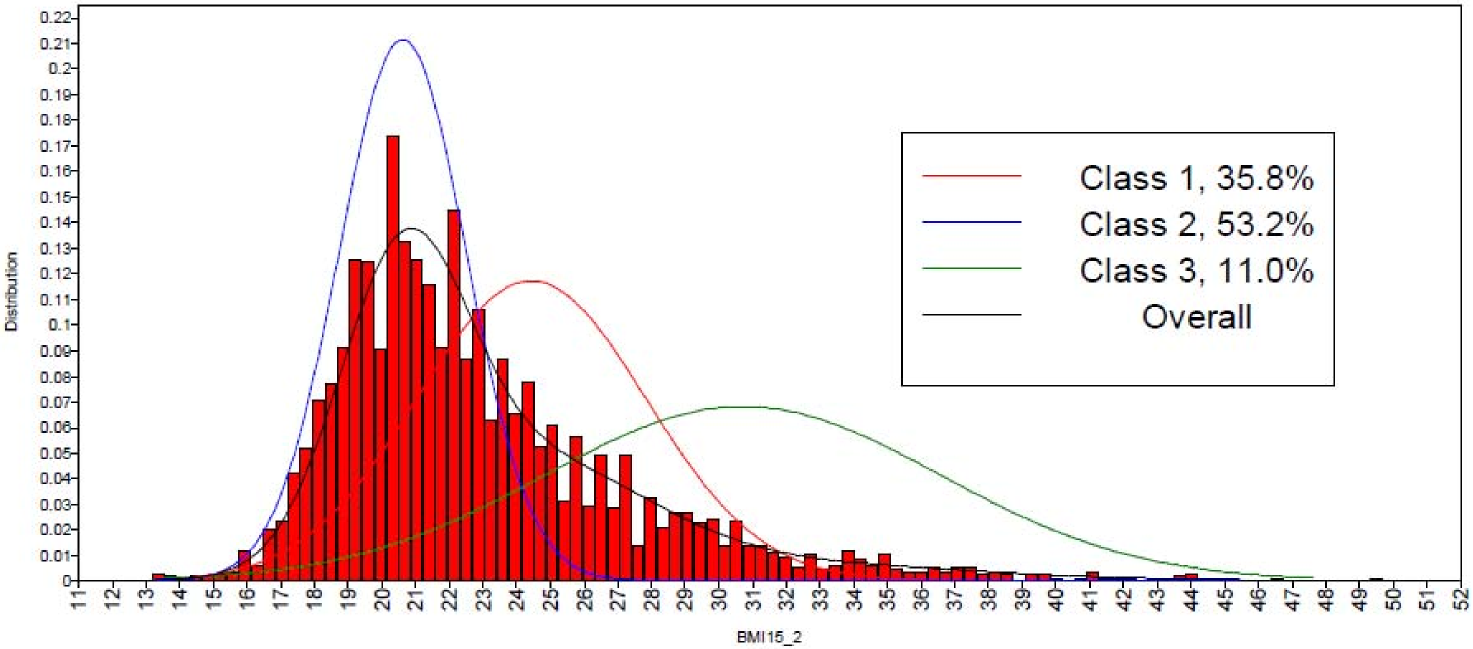
Mixture of several classes or „simply“ non-normal distribution?

- Pearson (1895)
- Hypertension debate:
  - Platt (1963): Hypertension is a "disease" (separate class)
  - Pickering (1968): Hypertension is merely the upper tail of a skewed distribution
- Schork et al (1990): Two-component mixture versus lognormal

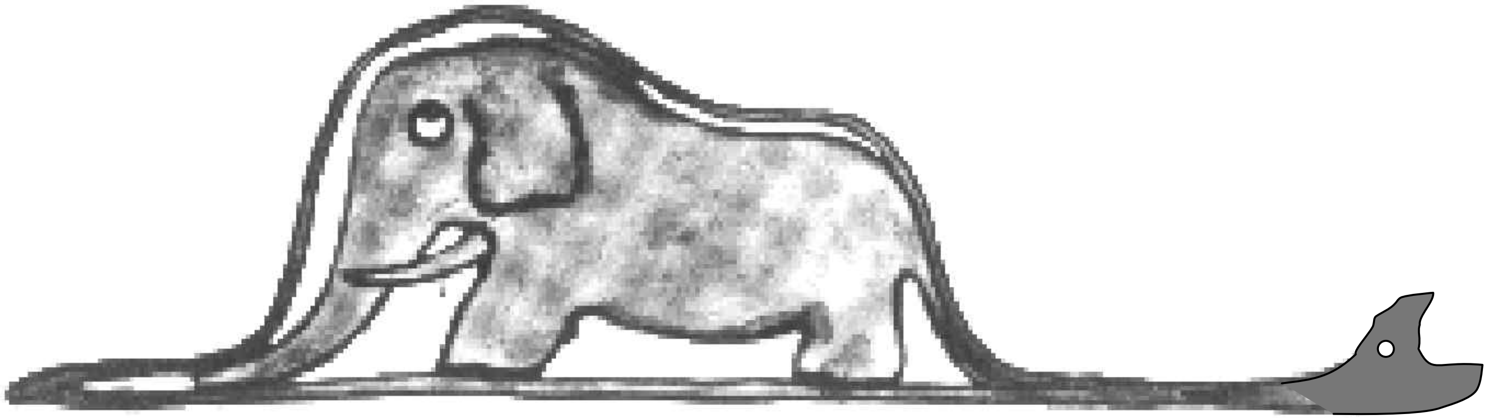
- Discussion on overextraction of latent trajectory classes in 2003 in *Psychological Methods* (Bauer & Curran; Rindskopf; Muthen;)

Explaining example: BMI of 15 year old boys

(from a presentation of Muthen, 2014, to PSMG (see at www.statmodel.com))



Outcome distributions as evidence, using the metaphor of Antoine de Saint-Exupéry‘s „Little prince“- story



Kinder- und Jugend-
psychiatrie / Psychotherapie

Universitätsklinikum Ulm

**Klinik für Kinder- und Jugendpsychiatrie /
Psychotherapie des Universitätsklinikums Ulm**

Steinhövelstraße 5
89075 Ulm

www.uniklinik-ulm.de/kjpp

Ärztlicher Direktor: Prof. Dr. Jörg M. Fegert

Beispiel zur Konfigurationsfrequenzanalyse

Datenbeispiel Nr. 2

Folgendes Datenbeispiel stammt aus der Erlanger Uni-Kinderklinik (siehe Tabelle 10). Epiletische Anfälle: '+' = ja; '-'= nein; Geburtsgewicht '+' = > 1000gr; '-' = < 1000gr.; Intelligenz im Alter von 5 Jahren: '+' = > Durchschnitt; '-' < Durchschnitt.

Tabelle: 10 Frühgeborene Kinder mit und ohne epiletische Anfälle und Geburtsgewicht sowie Intelligenz im Kindergartenalter

Epil. Anfälle	Geburtsgewicht	Intelligenz	f(o)
+	+	+	4
+	+	-	1
+	-	+	2
+	-	-	7
-	+	+	27
-	+	-	4
-	-	+	12
-	-	-	1