Longitudinal Mixed-Membership Models for Survey Data on Disability

Daniel Manrique –Vallier
Stephen E. Fienberg

Carnegie Mellon University

DATA ANALYSIS AND STATISTICAL FOUNDATIONS 3—DASF(3)
April 30-31, 2010

University of Toronto—1960-1964

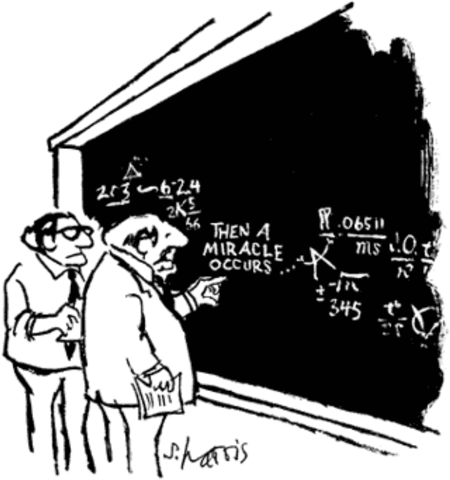
- My UofT Teachers (note overlap with DB and APD)
  - Coxeter (yrs. 1, 2, 3, 4)
  - DeLury (yrs. 2, 4)
  - Tutte (yr. 2)
  - DASF (yrs. 3, 4)
  - Wormleighton (yr. 3 lab)
  - Weber
- The Princeton Connection
  - The end of the line
- Other UofT Links
  - Dempster (Harvard)
  - Brillinger (Tukey's NBC Election Night ForecastingTeam)
    - Hierarchical Bayesian modeling
- Tukey-Mosteller collaboration

My First Course in Statistics—1962-1963

- Learning about inference and interval estimation.

My First Course in Statistics—1962-1963

- Learning about inference and interval estimation.
- Was DASF frequentist, fiducial, Bayesian, or structural?



"I think you should be more explicit here in step two."

DASF's Deely Award—ISBA 1995

- DASF was a late convert to the Bayesian school of inference.

DASF's Deely Award—ISBA 1995

- DASF was a late convert to the Bayesian school of inference.
- But only after he learned that Bayesians had better parties, especially after attending Valencia I in 1978.

DASF's Deely Award—ISBA 1995

- DASF was a late convert to the Bayesian school of inference.
- But only after he learned that Bayesians had better parties, especially after attending Valencia I in 1978.
- His crowning Bayesian achievement came with a Deely Award at the ISBA World meeting in Oaxaca, Mexico. 1996 for:

Best Talk in a Foreign Language

DASF's Deely Award—ISBA 1995

- DASF was a late convert to the Bayesian school of inference.
- But only after he learned that Bayesians had better parties, especially after attending Valencia I in 1978.
- His crowning Bayesian achievement came with a Deely Award at the ISBA World meeting in Oaxaca, Mexico. 1996 for:

Best Talk in a Foreign Language

- Even though the meeting was in Mexico, all the talks were in English!

Inference and Hierarchical Models

- Is DASF an objective or a subjective Bayesian?

Inference and Hierarchical Models

- Is DASF an objective or a subjective Bayesian?
- Hierarchical models pose interesting issues for the objective-subjective debate.

Inference and Hierarchical Models

- Is DASF an objective or a subjective Bayesian?
- Hierarchical models pose interesting issues for the objective-subjective debate.
- Morrie DeGroot often commented on the problem of deciding what was prior and what was likelihood, and the fact that the likelihood is often more subjective than the prior.

Inference and Hierarchical Models

- Is DASF an objective or a subjective Bayesian?
- Hierarchical models pose interesting issues for the objective-subjective debate.
- Morrie DeGroot often commented on the problem of deciding what was prior and what was likelihood, and the fact that the likelihood is often more subjective than the prior.
- As I will illustrate, these issues are exacerbated in hierarchical model settings, but at some level gentle priors are all we require.

Inference and Hierarchical Models

- Is DASF an objective or a subjective Bayesian?
- Hierarchical models pose interesting issues for the objective-subjective debate.
- Morrie DeGroot often commented on the problem of deciding what was prior and what was likelihood, and the fact that the likelihood is often more subjective than the prior.
- As I will illustrate, these issues are exacerbated in hierarchical model settings, but at some level gentle priors are all we require.
- *My Approach:* Pragmatic subjectivism or perhaps "constructive realistic" subjectivism.

My Only Collaboration With DASF

STATISTICS AT YORK

Prepared by:

Stephen E. Fienberg
D.A.S. Fraser
John Fox
Peter Peskun
Helene Massam

Draft: March 18, 1993

PREAMBLE

Several statisticians within the Department of Mathematics and Statistics continue to believe that the time has come for greater autonomy for statistics at York. The new divisional structure of the department has clearly allowed for some important dimensions of curricular control and has allowed the statisticians to organize activities in a more productive fashion. But the fact remains that the applied nature of statistics is far from flourishing within the department and that the formal link to mathematics is a deterrent to possible developments for statistics at York.

Many in the current department and elsewhere at York are unaware of the context in which statistics has developed as an independent discipline, albeit linked in various important ways to mathematics, and especially probability. In this document we have attempted to provide a brief overview of the development of the field of statistics and its institutionalization in Ontario universities. We have also identified some features that bind statistics with mathematics, e.g., computing and probability and others that draw statistics towards other disciplines and activities at York, especially in the social and behavioral sciences. We would also like to add an extensive discussion of the role of probability at the interface of Mathematics and Statistics.

Our intent is to begin a dialogue with colleagues in the Department of Mathematics and Statistics, as well as with interested colleagues outside the Department, about the future of statistics at York and how the aspirations for autonomy of statistics can be fulfilled. We are mindful that the probabilists in the current Department have strong academic interests in both mathematics and statistics and that any divided structure would need to have them in a bridging role.

We have explicitly avoided suggesting an administrative arrangement that would entail the necessary autonomy for the statisticians, although a separate department is obviously one possible model. We have not precluded other options, especially ones involving share resources and facilities. We have also not suggested any particular timetable for action.

We would welcome edits and additions to the following materials, and hope that some revised and expanded document might serve as the focus for a series of meetings on the development of statistics at York.

Outline for Remainder of My Talk

1. Motivation - Disability and long-term care in the U.S.
2. Data—The NLTCS
3. Proposed Approach—Trajectory GoM models
   1. General Construction
   2. Basic Model
   3. Estimation
4. Example Computations
5. Extensions
6. Discussion

Motivation - Disability and Long-Term Care in the U.S.

- The number and proportion of older Americans is increasing rapidly. Seniors are living longer than in previous generations.

Motivation - Disability and Long-Term Care in the U.S.

- The number and proportion of older Americans is increasing rapidly. Seniors are living longer than in previous generations.
- Older people often need long-term care, specially in the presence of disabilities.

Motivation - Disability and Long-Term Care in the U.S.

- The number and proportion of older Americans is increasing rapidly. Seniors are living longer than in previous generations.
- Older people often need long-term care, specially in the presence of disabilities.
- How are older Americans aging?

Motivation - Disability and Long-Term Care in the U.S.

- The number and proportion of older Americans is increasing rapidly. Seniors are living longer than in previous generations.
- Older people often need long-term care, specially in the presence of disabilities.
- How are older Americans aging?
  - Are disabilities compressed toward the end of life?

Motivation - Disability and Long-Term Care in the U.S.

- The number and proportion of older Americans is increasing rapidly. Seniors are living longer than in previous generations.
- Older people often need long-term care, specially in the presence of disabilities.
- How are older Americans aging?
  - Are disabilities compressed toward the end of life?
  - ... or is it a slow process over several years?

Motivation - Disability and Long-Term Care in the U.S.

- The number and proportion of older Americans is increasing rapidly. Seniors are living longer than in previous generations.
- Older people often need long-term care, specially in the presence of disabilities.
- How are older Americans aging?
  - Are disabilities compressed toward the end of life?
  - ... or is it a slow process over several years?
  - Is the process different for younger generations than for older ones?

Motivation - Disability and Long-Term Care in the U.S.

- The number and proportion of older Americans is increasing rapidly. Seniors are living longer than in previous generations.
- Older people often need long-term care, specially in the presence of disabilities.
- How are older Americans aging?
  - Are disabilities compressed toward the end of life?
  - ... or is it a slow process over several years?
  - Is the process different for younger generations than for older ones?
  - How is the population changing?

Motivation - Disability and Long-Term Care in the U.S.

- The number and proportion of older Americans is increasing rapidly. Seniors are living longer than in previous generations.
- Older people often need long-term care, specially in the presence of disabilities.
- How are older Americans aging?
  - Are disabilities compressed toward the end of life?
  - ... or is it a slow process over several years?
  - Is the process different for younger generations than for older ones?
  - How is the population changing?

Answers to these questions require a longitudinal view that also takes into account the heterogeneity of the population.

Data - The National Long Term Care Survey

- Longitudinal survey of people aged 65+

Data - The National Long Term Care Survey

- Longitudinal survey of people aged 65+
- Assess chronic disability

Data - The National Long Term Care Survey

- Longitudinal survey of people aged 65+
- Assess chronic disability
- 6 waves: 1982, 1984, 1989, 1994, 1999, 2004

Data - The National Long Term Care Survey

- Longitudinal survey of people aged 65+
- Assess chronic disability
- 6 waves: 1982, 1984, 1989, 1994, 1999, 2004
- Measures ADLs and IADLs:

Data - The National Long Term Care Survey

- Longitudinal survey of people aged 65+
- Assess chronic disability
- 6 waves: 1982, 1984, 1989, 1994, 1999, 2004
- Measures ADLs and IADLs:
  - Activities of daily living (ADL): Basic self-care (eating, bathing, etc.)—6 binary measures.

Data - The National Long Term Care Survey

- Longitudinal survey of people aged 65+
- Assess chronic disability
- 6 waves: 1982, 1984, 1989, 1994, 1999, 2004
- Measures ADLs and IADLs:
  - Activities of daily living (ADL): Basic self-care (eating, bathing, etc.)—6 binary measures.
  - Instrumental Activities of Daily Living (IADL): Related to independent living within a community (preparing meals, maintaining finances, etc.)—10 binary measures.

Data - The National Long Term Care Survey

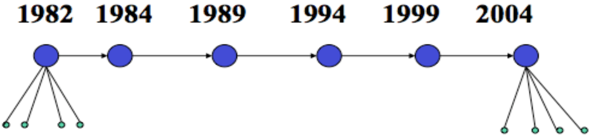
- Longitudinal survey of people aged 65+
- Assess chronic disability
- 6 waves: 1982, 1984, 1989, 1994, 1999, 2004
- Measures ADLs and IADLs:
  - Activities of daily living (ADL): Basic self-care (eating, bathing, etc.)—6 binary measures.
  - Instrumental Activities of Daily Living (IADL): Related to independent living within a community (preparing meals, maintaining finances, etc.)—10 binary measures.
- Each individual that enters the survey is reinterviewed in all subsequent waves until death.

Data - The National Long Term Care Survey

- Longitudinal survey of people aged 65+
- Assess chronic disability
- 6 waves: 1982, 1984, 1989, 1994, 1999, 2004
- Measures ADLs and IADLs:
  - Activities of daily living (ADL): Basic self-care (eating, bathing, etc.)—6 binary measures.
  - Instrumental Activities of Daily Living (IADL): Related to independent living within a community (preparing meals, maintaining finances, etc.)—10 binary measures.
- Each individual that enters the survey is reinterviewed in all subsequent waves until death.
- Approx. $20k$ individuals per wave. $45,009$ unique individuals sampled in all six waves together. Each wave incorporates $\approx 5k$ new subjects to replace those who have died.

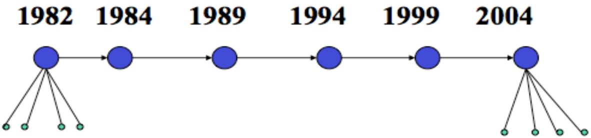
Longitudinal Modeling of NLTCS—Overview

- Sequential measurements on the same individuals allow to assess *individual* disability trajectories over time.



Longitudinal Modeling of NLTCS—Overview

- Sequential measurements on the same individuals allow to assess *individual* disability trajectories over time.

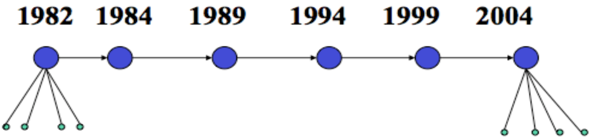


Specifically, we want to

- Understand evolution over time:

Longitudinal Modeling of NLTCS—Overview

- Sequential measurements on the same individuals allow to assess *individual* disability trajectories over time.

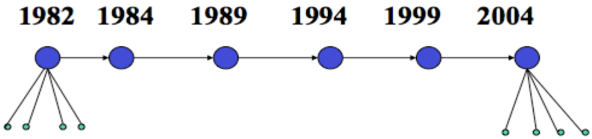


Specifically, we want to

- Understand evolution over time:
  - Individuals

Longitudinal Modeling of NLTCS—Overview

- Sequential measurements on the same individuals allow to assess *individual* disability trajectories over time.

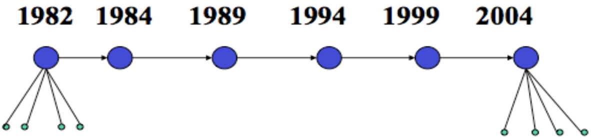


Specifically, we want to

- Understand evolution over time:
  - Individuals
  - Population

Longitudinal Modeling of NLTCS—Overview

- Sequential measurements on the same individuals allow to assess *individual* disability trajectories over time.

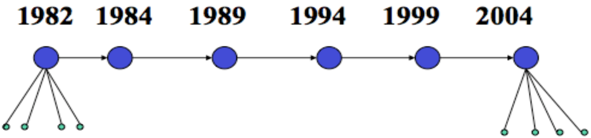


Specifically, we want to

- Understand evolution over time:
  - Individuals
  - Population
- Identify ‘typical’ evolutions over time

Longitudinal Modeling of NLTCS—Overview

- Sequential measurements on the same individuals allow to assess *individual* disability trajectories over time.



Specifically, we want to

- Understand evolution over time:
  - Individuals
  - Population
- Identify 'typical' evolutions over time
- Account for and understand individual variability

Data—NLTCS Longitudinal View and Notation

		$t=1$	$t=2$	$t=3$	$t=4$
$i=1$	$j=1$	0	0	1	1
	$j=2$	0	0	0	1
			$\vdots$		
	$j=J$	0	0	0	0
	Age	67	69	74	79
	Sex		F		

$\vdots$

		$t=1$	$t=2$	$t=3$	$t=4$
$i=N$	$j=1$	0	0	1	–
	$j=2$	0	1	1	–
			$\vdots$		
	$j=J$	1	1	1	–
	Age	80	82	87	–
	Sex		M		

- N individuals indexed by $i \in \{1, 2, ..., N\}$

Data—NLTCS Longitudinal View and Notation

		$t = 1$	$t = 2$	$t = 3$	$t = 4$
$i = 1$	$j = 1$	0	0	1	1
	$j = 2$	0	0	0	1
	$\vdots$				
	$j = J$	0	0	0	0
	Age	67	69	74	79
	Sex	F			

$\vdots$

		$t = 1$	$t = 2$	$t = 3$	$t = 4$
$i = N$	$j = 1$	0	0	1	–
	$j = 2$	0	1	1	–
	$\vdots$				
	$j = J$	1	1	1	–
	Age	80	82	87	–
	Sex	M			

- N individuals indexed by $i \in \{1, 2, ..., N\}$

Data—NLTCS Longitudinal View and Notation

		$t=1$	$t=2$	$t=3$	$t=4$
$i=1$	$j=1$	0	0	1	1
	$j=2$	0	0	0	1
	⋮				
	$j=J$	0	0	0	0
	Age	67	69	74	79
	Sex	F			

⋮

		$t=1$	$t=2$	$t=3$	$t=4$
$i=N$	$j=1$	0	0	1	–
	$j=2$	0	1	1	–
	⋮				
	$j=J$	1	1	1	–
	Age	80	82	87	–
	Sex	M			

- N individuals indexed by $i \in \{1, 2, ..., N\}$
- J binary response items indexed by $j \in \{1, 2, ..., J\}$

Data—NLTCS Longitudinal View and Notation

		$t=1$	$t=2$	$t=3$	$t=4$
$i=1$	$j=1$	0	0	1	1
	$j=2$	0	0	0	1
	$\vdots$				
	$j=J$	0	0	0	0
	Age	67	69	74	79
	Sex	F			

$\vdots$

		$t=1$	$t=2$	$t=3$	$t=4$
$i=N$	$j=1$	0	0	1	–
	$j=2$	0	1	1	–
	$\vdots$				
	$j=J$	1	1	1	–
	Age	80	82	87	–
	Sex	M			

- N individuals indexed by $i \in \{1, 2, ..., N\}$
- J binary response items indexed by $j \in \{1, 2, ..., J\}$
- Measurements indexed by $t \in \{1, 2, ..., N_i\}$

Data—NLTCS Longitudinal View and Notation

		$t=1$	$t=2$	$t=3$	$t=4$
$i=1$	$j=1$	0	0	1	1
	$j=2$	0	0	0	1
	$\vdots$				
	$j=J$	0	0	0	0
	Age	67	69	74	79
	Sex	F			

$\vdots$

		$t=1$	$t=2$	$t=3$	$t=4$
$i=N$	$j=1$	0	0	1	–
	$j=2$	0	1	1	–
	$\vdots$				
	$j=J$	1	1	1	–
	Age	80	82	87	–
	Sex	M			

- N individuals indexed by $i \in \{1, 2, ..., N\}$
- J binary response items indexed by $j \in \{1, 2, ..., J\}$
- Measurements indexed by $t \in \{1, 2, ..., N_i\}$
- Longitudinal/Cross-Sectional

Data—NLTCS Longitudinal View and Notation

		$t=1$	$t=2$	$t=3$	$t=4$
$i=1$	$j=1$	0	0	1	1
	$j=2$	0	0	0	1
	$\vdots$				
	$j=J$	0	0	0	0
	Age	67	69	74	79
	Sex	F			

$\vdots$

		$t=1$	$t=2$	$t=3$	$t=4$
$i=N$	$j=1$	0	0	1	–
	$j=2$	0	1	1	–
	$\vdots$				
	$j=J$	1	1	1	–
	Age	80	82	87	–
	Sex	M			

- N individuals indexed by $i \in \{1, 2, ..., N\}$
- J binary response items indexed by $j \in \{1, 2, ..., J\}$
- Measurements indexed by $t \in \{1, 2, ..., N_i\}$
- Longitudinal/Cross-Sectional

Data—NLTCS Longitudinal View and Notation

		$t=1$	$t=2$	$t=3$	$t=4$
$i=1$	$j=1$	0	0	1	1
	$j=2$	0	0	0	1
	$\vdots$				
	$j=J$	0	0	0	0
	Age	67	69	74	79
	Sex	F			

$\vdots$

		$t=1$	$t=2$	$t=3$	$t=4$
$i=N$	$j=1$	0	0	1	–
	$j=2$	0	1	1	–
	$\vdots$				
	$j=J$	1	1	1	–
	Age	80	82	87	–
	Sex	M			

- N individuals indexed by $i \in \{1, 2, ..., N\}$
- J binary response items indexed by $j \in \{1, 2, ..., J\}$
- Measurements indexed by $t \in \{1, 2, ..., N_i\}$
- Longitudinal/Cross-Sectional
- Other information

Data—NLTCS Longitudinal View and Notation

		$t=1$	$t=2$	$t=3$	$t=4$
$i=1$	$j=1$	0	0	1	1
	$j=2$	0	0	0	1
	⋮				
	$j=J$	0	0	0	0
	Age	67	69	74	79
	Sex	F			

⋮

		$t=1$	$t=2$	$t=3$	$t=4$
$i=N$	$j=1$	0	0	1	–
	$j=2$	0	1	1	–
	⋮				
	$j=J$	1	1	1	–
	Age	80	82	87	–
	Sex	M			

- N individuals indexed by $i \in \{1, 2, ..., N\}$
- J binary response items indexed by $j \in \{1, 2, ..., J\}$
- Measurements indexed by $t \in \{1, 2, ..., N_i\}$
- Longitudinal/Cross-Sectional
- Other information
  - Time dependent (e.g. Age)

Data—NLTCS Longitudinal View and Notation

		$t=1$	$t=2$	$t=3$	$t=4$
$i=1$	$j=1$	0	0	1	1
	$j=2$	0	0	0	1
			$\vdots$		
	$j=J$	0	0	0	0
	Age	67	69	74	79
	Sex	F			

$\vdots$

		$t=1$	$t=2$	$t=3$	$t=4$
$i=N$	$j=1$	0	0	1	–
	$j=2$	0	1	1	–
			$\vdots$		
	$j=J$	1	1	1	–
	Age	80	82	87	–
	Sex	M			

- N individuals indexed by $i \in \{1, 2, ..., N\}$
- J binary response items indexed by $j \in \{1, 2, ..., J\}$
- Measurements indexed by $t \in \{1, 2, ..., N_i\}$
- Longitudinal/Cross-Sectional
- Other information
  - Time dependent (e.g. Age)
  - Fixed (e.g. DOB, Sex)

Proposed Approach

- Combine the Grade of Membership Model (Woodbury et al. 1978, Erosheva et al. 2007) and the Multivariate Latent Trajectories Model (Connor, 2006).

Proposed Approach

- Combine the Grade of Membership Model (Woodbury et al. 1978, Erosheva et al. 2007) and the Multivariate Latent Trajectories Model (Connor, 2006).
- Latent Trajectories:

Proposed Approach

- Combine the Grade of Membership Model (Woodbury et al. 1978, Erosheva et al. 2007) and the Multivariate Latent Trajectories Model (Connor, 2006).
- Latent Trajectories:
  - Specify latent classes of individuals and model their time-dependent trajectories of probability of acquiring disabilities.

Proposed Approach

- Combine the Grade of Membership Model (Woodbury et al. 1978, Erosheva et al. 2007) and the Multivariate Latent Trajectories Model (Connor, 2006).
- Latent Trajectories:
  - Specify latent classes of individuals and model their time-dependent trajectories of probability of acquiring disabilities.
  - Summarize the typical ways of acquisition of disabilities.

Proposed Approach

- Combine the Grade of Membership Model (Woodbury et al. 1978, Erosheva et al. 2007) and the Multivariate Latent Trajectories Model (Connor, 2006).
- Latent Trajectories:
  - Specify latent classes of individuals and model their time-dependent trajectories of probability of acquiring disabilities.
  - Summarize the typical ways of acquisition of disabilities.
- Grade of Membership:

Proposed Approach

- Combine the Grade of Membership Model (Woodbury et al. 1978, Erosheva et al. 2007) and the Multivariate Latent Trajectories Model (Connor, 2006).
- Latent Trajectories:
  - Specify latent classes of individuals and model their time-dependent trajectories of probability of acquiring disabilities.
  - Summarize the typical ways of acquisition of disabilities.
- Grade of Membership:
  - ”Soft classification”.

Proposed Approach

- Combine the Grade of Membership Model (Woodbury et al. 1978, Erosheva et al. 2007) and the Multivariate Latent Trajectories Model (Connor, 2006).
- Latent Trajectories:
  - Specify latent classes of individuals and model their time-dependent trajectories of probability of acquiring disabilities.
  - Summarize the typical ways of acquisition of disabilities.
- Grade of Membership:
  - ”Soft classification”.
  - Individuals “belong” to more than one class simultaneously.

Proposed Approach

- Combine the Grade of Membership Model (Woodbury et al. 1978, Erosheva et al. 2007) and the Multivariate Latent Trajectories Model (Connor, 2006).
- Latent Trajectories:
  - Specify latent classes of individuals and model their time-dependent trajectories of probability of acquiring disabilities.
  - Summarize the typical ways of acquisition of disabilities.
- Grade of Membership:
  - ”Soft classification”.
  - Individuals “belong” to more than one class simultaneously.
  - *Individual* degrees of membership.

Proposed Approach

- Combine the Grade of Membership Model (Woodbury et al. 1978, Erosheva et al. 2007) and the Multivariate Latent Trajectories Model (Connor, 2006).
- Latent Trajectories:
  - Specify latent classes of individuals and model their time-dependent trajectories of probability of acquiring disabilities.
  - Summarize the typical ways of acquisition of disabilities.
- Grade of Membership:
  - ”Soft classification”.
  - Individuals “belong” to more than one class simultaneously.
  - *Individual* degrees of membership.
  - Acknowledges the fact that real individuals have unique trajectories.

Mixed-Membership Models

- Grade of Membership is a special example of mixed-membership models that have been applied in other contexts.

Mixed-Membership Models

- Grade of Membership is a special example of mixed-membership models that have been applied in other contexts.
  - Population genetics—people come from “mixed” origins.

Mixed-Membership Models

- Grade of Membership is a special example of mixed-membership models that have been applied in other contexts.
  - Population genetics—people come from “mixed” origins.
  - Document analysis—using words and references.

Mixed-Membership Models

- Grade of Membership is a special example of mixed-membership models that have been applied in other contexts.
  - Population genetics—people come from “mixed” origins.
  - Document analysis—using words and references.
  - Network stochastic blockmodels with mixed-membership applied to protein-protein interactions and friendship among monks in a monastery.

Mixed-Membership Models

- Grade of Membership is a special example of mixed-membership models that have been applied in other contexts.
  - Population genetics—people come from “mixed” origins.
  - Document analysis—using words and references.
  - Network stochastic blockmodels with mixed-membership applied to protein-protein interactions and friendship among monks in a monastery.
  - Aspects of the spread of malaria in the Amazon.

Mixed-Membership Models

- Grade of Membership is a special example of mixed-membership models that have been applied in other contexts.
  - Population genetics—people come from “mixed” origins.
  - Document analysis—using words and references.
  - Network stochastic blockmodels with mixed-membership applied to protein-protein interactions and friendship among monks in a monastery.
  - Aspects of the spread of malaria in the Amazon.
- The specific longitudinal models we describe have potential application to the study of other phenomena measured at discrete points in time.

Modeling—Construction of a trajectory GoM model (1)

- Assume the existence of K “ideal classes” or “extreme profiles”
- Assign each individual a *Membership Vector*:

$$g_i = (g_{i1}, g_{i2}, ..., g_{iK})$$

with $g_{ik} > 0$ and $\sum_{k=1}^{K} g_{ik} = 1$ ($g_i \in \Delta_{K-1}$).
- For the “ideal” individuals, specify the marginal distribution of response j, at measurement time t, as a function of some time-dependent covariates.

$$\Pr\left(Y_{ijt} = y_{ijt} \mid g_{ik} = 1, X_i, \theta\right) = f_{\theta_{j|k}}\left(y_{ijt} \mid X_{it}\right)$$

Modeling—Construction of a trajectory GoM model (1)

- Assume the existence of K "ideal classes" or "extreme profiles"
- Assign each individual a *Membership Vector*:

$$g_i = (g_{i1}, g_{i2}, ..., g_{iK})$$

with $g_{ik} > 0$ and $\sum_{k=1}^{K} g_{ik} = 1$ ($g_i \in \Delta_{K-1}$).

- For the "ideal" individuals, specify the marginal distribution of response j, at measurement time t, as a function of some time-dependent covariates.

$$\Pr\left(Y_{ijt} = y_{ijt} \mid g_{ik} = 1, X_i, \theta\right) = f_{\theta_{j|k}}\left(y_{ijt} \mid X_{it}\right)$$

Modeling—Construction of a trajectory GoM model (1)

- Assume the existence of K "ideal classes" or "extreme profiles"
- Assign each individual a *Membership Vector*:

$$g_i = (g_{i1}, g_{i2}, ..., g_{iK})$$

with $g_{ik} > 0$ and $\sum_{k=1}^{K} g_{ik} = 1$ ($g_i \in \Delta_{K-1}$).

- For the "ideal" individuals, specify the marginal distribution of response j, at measurement time t, as a function of some time-dependent covariates.

$$\Pr\left(Y_{ijt} = y_{ijt} \mid g_{ik} = 1, X_i, \theta\right) = f_{\theta_{j|k}}\left(y_{ijt} \mid X_{it}\right)$$

Modeling—Construction of a trajectory GoM model (2)

- Mixed Membership: For a generic individual *i*, we model

$$\Pr\left(Y_{ijt} = y_{ijt}\,|g_i, X_i, \theta\right) = \sum_{k=1}^{K} g_{ik} f_{\theta_{j|k}}(y_{ijt}|X_{it})$$

- Assuming conditional independence,

$$\Pr\left(Y_i = y_i\,|g_i, X_i, \theta\right) = \prod_{j=1}^{J}\prod_{t=1}^{N_i}\sum_{k=1}^{K} g_{ik} f_{\theta_{j|k}}(y_{ijt}|X_{it})$$

- Assume that the membership vectors are an iid sample from a common distribution with support on the $K-1$ dimensional unit simplex (Δ_{K-1}):

$$g_i|\alpha \overset{iid}{\sim} G_\alpha$$

Modeling—Construction of a trajectory GoM model (2)

- Mixed Membership: For a generic individual i, we model

$$\Pr\left(Y_{ijt} = y_{ijt} \,|\, g_i, X_i, \theta\right) = \sum_{k=1}^{K} g_{ik} f_{\theta_{j|k}}(y_{ijt}|X_{it})$$

- Assuming conditional independence,

$$\Pr(Y_i = y_i \,|\, g_i, X_i, \theta) = \prod_{j=1}^{J} \prod_{t=1}^{N_i} \sum_{k=1}^{K} g_{ik} f_{\theta_{j|k}}(y_{ijt}|X_{it})$$

- Assume that the membership vectors are an iid sample from a common distribution with support on the $K-1$ dimensional unit simplex (Δ_{K-1}):

$$g_i|\alpha \overset{iid}{\sim} G_\alpha$$

Modeling—Construction of a trajectory GoM model (2)

- Mixed Membership: For a generic individual i, we model

$$\Pr\left(Y_{ijt} = y_{ijt} \,|g_i, X_i, \theta\right) \;\; = \;\; \sum_{k=1}^{K} g_{ik} f_{\theta_{j|k}}(y_{ijt}|X_{it})$$

- Assuming conditional independence,

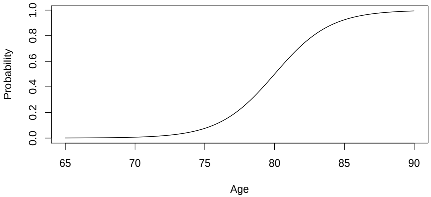
$$\Pr\left(Y_i = y_i \,|g_i, X_i, \theta\right) = \prod_{j=1}^{J} \prod_{t=1}^{N_i} \sum_{k=1}^{K} g_{ik} f_{\theta_{j|k}}(y_{ijt}|X_{it})$$

- Assume that the membership vectors are an iid sample from a common distribution with support on the $K-1$ dimensional unit simplex (Δ_{K-1}):

$$g_i|\alpha \overset{iid}{\sim} G_{\alpha}$$

Basic Model—Extreme profile Trajectories

- For each extreme profile ($g_k = 1$) specify trajectories of probability of disability in ADLs as a monotone function of Age:



$$
\begin{aligned}
y_{ijt} &\sim \text{Bernoulli}\left[\lambda_{j|k}(Age_{it})\right] \\
\lambda_{j|k}(X_{it}) &= \text{logit}^{-1}\left[\beta_{0j|k} + \beta_{1j|k} \times Age_{it}\right]
\end{aligned}
$$

(Connor, 2006)

Basic Model—Distribution for g_i ($\sim G_\alpha$)

- Membership vectors from a Dirichlet distribution

$$g_i \overset{iid}{\sim} \text{Dirichlet}(\alpha_0 \times \xi)$$

with $\alpha_0 > 0$ and $\xi = (\xi_1, \xi_2, ..., \xi_K) \in \Delta_{K-1}$.

Basic Model—Distribution for g_i ($\sim G_\alpha$)

- Membership vectors from a Dirichlet distribution

$$g_i \overset{iid}{\sim} \text{Dirichlet}(\alpha_0 \times \xi)$$

with $\alpha_0 > 0$ and $\xi = (\xi_1, \xi_2, ..., \xi_K) \in \Delta_{K-1}$.

- Parametrization:

Basic Model—Distribution for g_i ($\sim G_\alpha$)

- Membership vectors from a Dirichlet distribution

$$g_i \overset{iid}{\sim} \text{Dirichlet}(\alpha_0 \times \xi)$$

with $\alpha_0 > 0$ and $\xi = (\xi_1, \xi_2, ..., \xi_K) \in \Delta_{K-1}$.

- Parametrization:
  - $\xi = (\xi_1, \xi_2, ..., \xi_K)$: Represents the relative weight of the k-th component. (Think of “percentage”).

Basic Model—Distribution for g_i ($\sim G_\alpha$)

- Membership vectors from a Dirichlet distribution

$$g_i \overset{iid}{\sim} \text{Dirichlet}(\alpha_0 \times \xi)$$

with $\alpha_0 > 0$ and $\xi = (\xi_1, \xi_2, ..., \xi_K) \in \Delta_{K-1}$.

- Parametrization:
  - $\xi = (\xi_1, \xi_2, ..., \xi_K)$: Represents the relative weight of the k-th component. (Think of "percentage").
  - α_0: The concentration of probability toward the center ($\alpha_0 > K$) or the vertices ($\alpha_0 < K$) of the simplex.

Basic Model—Distribution for g_i ($\sim G_\alpha$)

- Membership vectors from a Dirichlet distribution

$$g_i \overset{iid}{\sim} \text{Dirichlet}(\alpha_0 \times \xi)$$

  with $\alpha_0 > 0$ and $\xi = (\xi_1, \xi_2, ..., \xi_K) \in \Delta_{K-1}$.
- Parametrization:
  - $\xi = (\xi_1, \xi_2, ..., \xi_K)$: Represents the relative weight of the k-th component. (Think of "percentage").
  - α_0: The concentration of probability toward the center ($\alpha_0 > K$) or the vertices ($\alpha_0 < K$) of the simplex.
- Interesting cases:

Basic Model—Distribution for g_i ($\sim G_\alpha$)

- Membership vectors from a Dirichlet distribution

$$g_i \overset{iid}{\sim} \text{Dirichlet}(\alpha_0 \times \xi)$$

with $\alpha_0 > 0$ and $\xi = (\xi_1, \xi_2, ..., \xi_K) \in \Delta_{K-1}$.

- Parametrization:
  - $\xi = (\xi_1, \xi_2, ..., \xi_K)$: Represents the relative weight of the k-th component. (Think of "percentage").
  - α_0: The concentration of probability toward the center ($\alpha_0 > K$) or the vertices ($\alpha_0 < K$) of the simplex.
- Interesting cases:
  - Dirichlet($\alpha_0 \times \xi$) with $\alpha_0 = K$ and $\xi = (\frac{1}{K}, \frac{1}{K}, ..., \frac{1}{K})$ defines a uniform distribution over Δ_{K-1}

Basic Model—Distribution for g_i ($\sim G_\alpha$)

- Membership vectors from a Dirichlet distribution

$$g_i \overset{iid}{\sim} \text{Dirichlet}(\alpha_0 \times \xi)$$

with $\alpha_0 > 0$ and $\xi = (\xi_1, \xi_2, ..., \xi_K) \in \Delta_{K-1}$.

- Parametrization:
  - $\xi = (\xi_1, \xi_2, ..., \xi_K)$: Represents the relative weight of the k-th component. (Think of "percentage").
  - α_0: The concentration of probability toward the center ($\alpha_0 > K$) or the vertices ($\alpha_0 < K$) of the simplex.
- Interesting cases:
  - Dirichlet($\alpha_0 \times \xi$) with $\alpha_0 = K$ and $\xi = (\frac{1}{K}, \frac{1}{K}, ..., \frac{1}{K})$ defines a uniform distribution over Δ_{K-1}

Basic Model—Priors

- For the membership distribution

$$g_i \overset{iid}{\sim} \text{Dirichlet}(\alpha)$$

We use the same priors as Erosheva (2002):

$$\begin{aligned} \alpha &= \alpha_0 \times \xi \\ \alpha_0 &\sim \text{Gamma}(1, 5) \\ \xi &\sim \text{Uniform}[\Delta_{K-1}] \end{aligned}$$

Basic Model—Priors

- For the membership distribution

$$g_i \overset{iid}{\sim} \text{Dirichlet}(\alpha)$$

We use the same priors as Erosheva (2002):

$$\begin{aligned} \alpha &= \alpha_0 \times \xi \\ \alpha_0 &\sim \text{Gamma}(1,5) \\ \xi &\sim \text{Uniform}[\Delta_{K-1}] \end{aligned}$$

- And complete the specification with

$$\begin{aligned} \beta_{0j|k} &\overset{iid}{\sim} N(0,100) \\ \beta_{1j|k} &\overset{iid}{\sim} N(0,100) \end{aligned}$$

Estimation—MCMC sampling

- MCMC algorithm based on a method from Erosheva (2002) for fitting GoM model to cross sectional data. Using an equivalent latent class representation for the GoM model.
- Difficult to run
  - Huge latent space.
  - Nonstandard distributions.
  - Numerical problems.
- 40,000 long chains (using the “improved” algorithm). $5 \sim 7$h runs

Test Computations—Data

- Tested for six ADLs:

ADL (j-index)	Abbrv	Description
1	EAT	Eating
2	BED	Getting in and out of bed
3	MOB	Inside mobility
4	DRS	Dressing
5	BTH	Bathing
6	TLT	Toileting

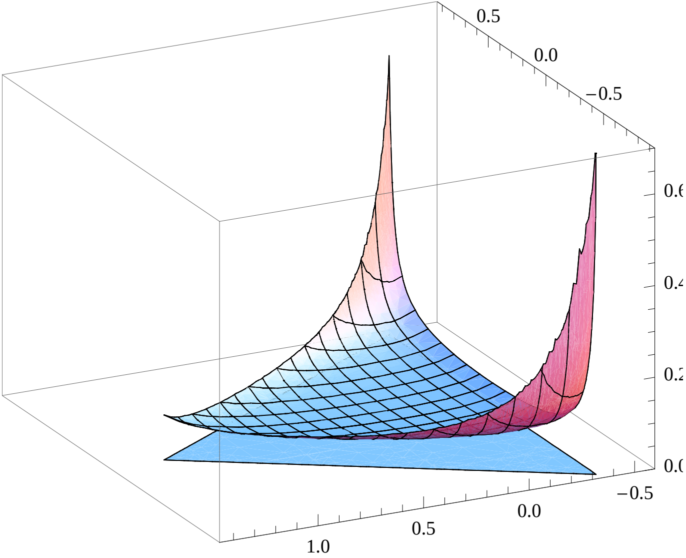
- Data from 6 waves (1982 - 2004).
- Individuals from 2004 are only those that were already in the 1999 sample.
- $N \approx 40K$

Test Computations - Posterior Summaries ($K = 3$)

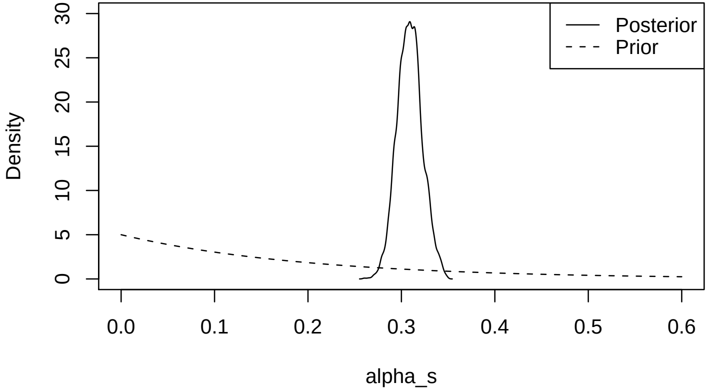
Parameter	Estimate (sd)
α_0	0.264 (0.00489)

Parameter	ADL(j)	Estimate Extreme Profile-k (sd)					
		$k = 1$		$k = 2$		$k = 3$	
ξ	—	0.645	(0.004)	0.252	(0.003)	0.104	(0.002)
β_0	1 (EAT)	-8.845	(0.313)	-3.103	(0.057)	-0.066	(0.044)
	2 (BED)	-7.02	(0.144)	-1.739	(0.053)	3.581	(0.142)
	3 (MOB)	-5.339	(0.093)	-0.759	(0.044)	5.803	(0.277)
	4 (DRS)	-7.912	(0.216)	-2.256	(0.051)	2.042	(0.082)
	5 (BTH)	-4.458	(0.075)	-0.23	(0.035)	6.257	(0.28)
	6 (LTL)	-6.59	(0.148)	-1.768	(0.047)	2.506	(0.098)
β_1	1 (EAT)	0.357	(0.017)	0.347	(0.008)	0.105	(0.006)
	2 (BED)	0.394	(0.01)	0.551	(0.013)	0.29	(0.012)
	3 (MOB)	0.348	(0.007)	0.52	(0.012)	0.426	(0.022)
	4 (DRS)	0.392	(0.013)	0.463	(0.011)	0.203	(0.008)
	5 (BTH)	0.295	(0.006)	0.426	(0.009)	0.445	(0.022)
	6 (LTL)	0.337	(0.009)	0.475	(0.011)	0.234	(0.009)
$Age_{1/2}$	1 (EAT)	104.82	(0.46)	88.945	(0.163)	80.641	(0.444)
	2 (BED)	97.838	(0.167)	83.154	(0.089)	67.657	(0.173)
	3 (MOB)	95.338	(0.137)	81.458	(0.083)	66.389	(0.151)
	4 (DRS)	100.212	(0.231)	84.869	(0.104)	69.934	(0.192)
	5 (BTH)	95.118	(0.151)	80.54	(0.082)	65.927	(0.16)
	6 (LTL)	99.553	(0.222)	83.725	(0.092)	69.3	(0.175)

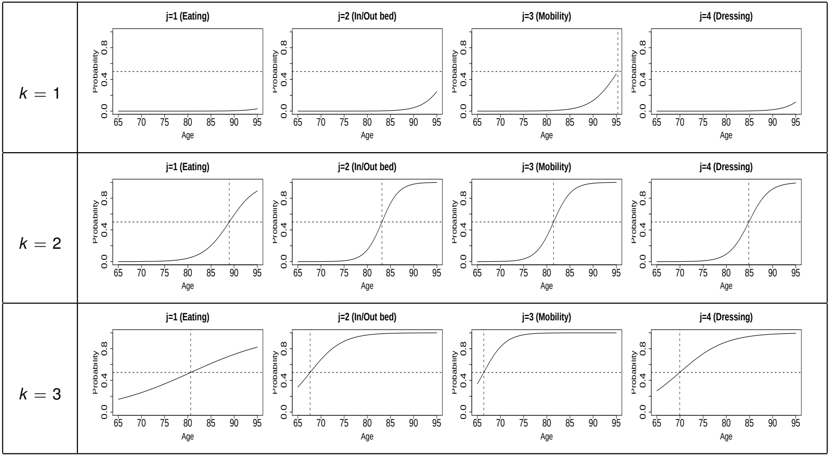
Computations - "Posterior density" for g_i (K=3)



Computations - prior/posterior for α_0 ($K = 3$)

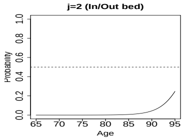


Test Computations—Profiles for $K = 3$

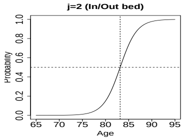


Test Computations—From profiles to Individuals

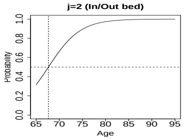
$k = 1$



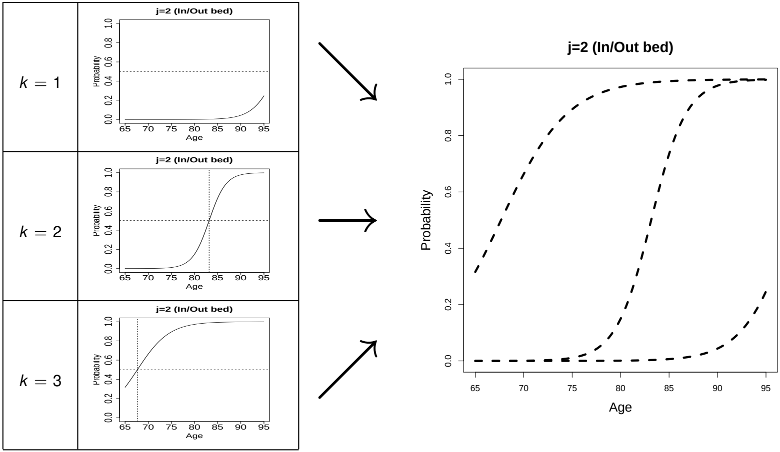
$k = 2$



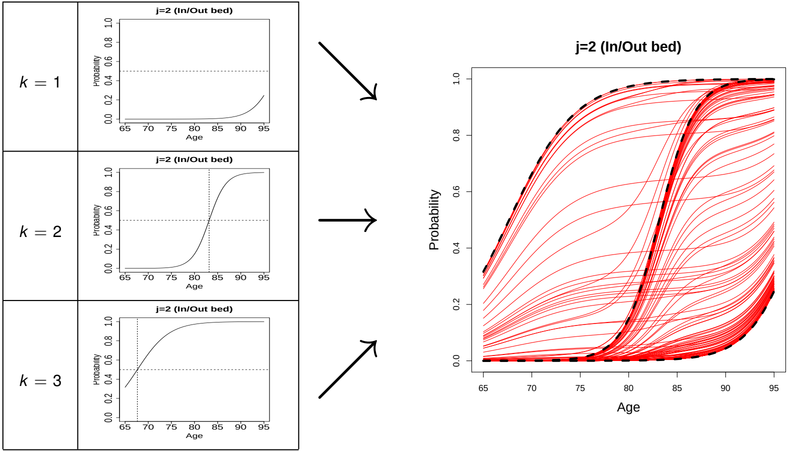
$k = 3$



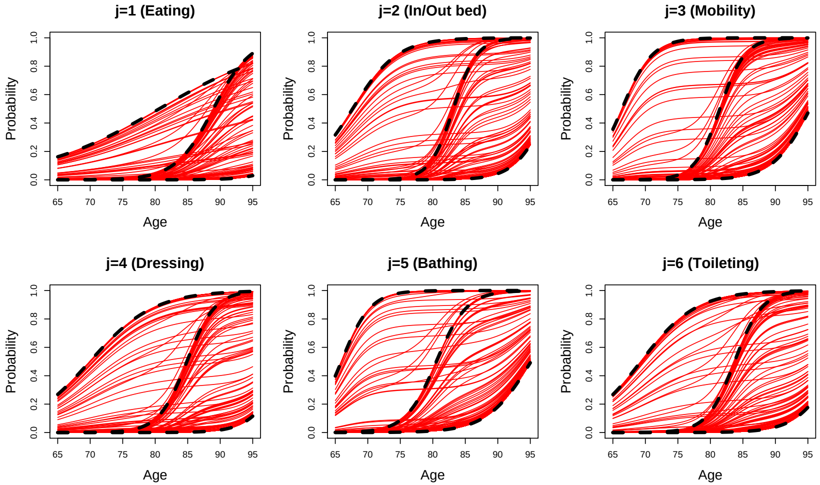
Test Computations—From profiles to Individuals



Test Computations—From profiles to Individuals



Test Computations—Individual Trajectories



Extensions

1. Modeling of generational differences through generation-dependent group membership distributions.

Extensions

1. Modeling of generational differences through generation-dependent group membership distributions.
   - Assess changes over time as a function of individual's "generational group".

Extensions

1. Modeling of generational differences through generation-dependent group membership distributions.
   - Assess changes over time as a function of individual's "generational group".
   - Allow to answer the question: "Are younger generations acquiring disabilities in a different way than older ones?"

Extensions

1. Modeling of generational differences through generation-dependent group membership distributions.
   - Assess changes over time as a function of individual's "generational group".
   - Allow to answer the question: "Are younger generations acquiring disabilities in a different way than older ones?"
2. Joint modeling of survival times and disability acquisition.

Extensions

1. Modeling of generational differences through generation-dependent group membership distributions.
   - Assess changes over time as a function of individual's "generational group".
   - Allow to answer the question: "Are younger generations acquiring disabilities in a different way than older ones?"
2. Joint modeling of survival times and disability acquisition.
   - Use survival information to achieve better classification.

Extensions

1. Modeling of generational differences through generation-dependent group membership distributions.
   - Assess changes over time as a function of individual's "generational group".
   - Allow to answer the question: "Are younger generations acquiring disabilities in a different way than older ones?"
2. Joint modeling of survival times and disability acquisition.
   - Use survival information to achieve better classification.
3. Other trajectory functions (e.g. step functions).

Extensions (1) - Modeling Generational differences

Have population disability patterns changed?

Extensions (1) - Modeling Generational differences

Have population disability patterns changed?

- Study generations (i.e. people that were born in the same period) to understand changes in how people acquire disabilities.

Extensions (1) - Modeling Generational differences

Have population disability patterns changed?

- Study generations (i.e. people that were born in the same period) to understand changes in how people acquire disabilities.
- Longitudinal data allows us to compare whole aging life trajectories for different individuals from different generations.

Extensions (2) - Modeling Generational differences

- Approach: make group membership dependent on the generation to which the individual belongs, keeping the extreme trajectories fixed:

$$\begin{aligned} \Pr\left(Y_{ijt} = y_{ijt} \,|g_i, X_i, \theta\right) &= \sum_{k=1}^{K} g_{ik} f_{\theta_{j|k}}(y_{ijt}|Age_{it}) \\ g_i|Gen_i &\sim G_{\alpha(Gen_i)} \end{aligned}$$

Extensions (2) - Modeling Generational differences

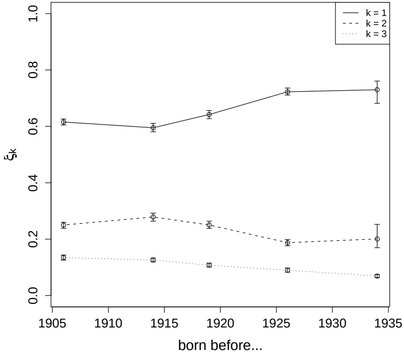
- Approach: make group membership dependent on the generation to which the individual belongs, keeping the extreme trajectories fixed:

$$\Pr\left(Y_{ijt} = y_{ijt}\,|g_i, X_i, \theta\right) = \sum_{k=1}^{K} g_{ik} f_{\theta_{j|k}}(y_{ijt}|Age_{it})$$
$$g_i|Gen_i \sim G_{\alpha(Gen_i)}$$

- This way we can assess the distribution of membership scores for different generational groups.

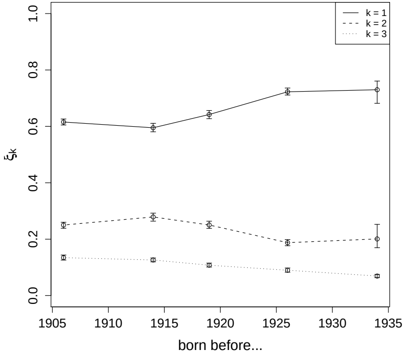
Extensions (2) - Modeling Generational Differences

- 5 (rather arbitrary) “Generational Groups”. Born...
  1. ...before 1906 (up to 1873)
  2. between 1906 and 1914,
  3. between 1914 and 1919,
  4. between 1919 and 1926,
  5. after 1926 (up to 1934).



Extensions (2) - Modeling Generational Differences

- 5 (rather arbitrary) "Generational Groups". Born...
  1. ...before 1906 (up to 1873)
  2. between 1906 and 1914,
  3. between 1914 and 1919,
  4. between 1919 and 1926,
  5. after 1926 (up to 1934).



Individuals form younger generations tend to be closer to "healthy" trajectory profiles.

Other Extensions

1. Joint modeling of survival times and disability acquisition.
   - Use survival information to achieve better classification.
   - Understand the relationship between disability and mortality.
2. Other trajectory functions (e.g. step functions)
   - Test the constraints imposed by the selection of disability trajectory curves.
3. Using full database,
4. Model choice—picking the value of K.
5. Incorporating fuller set of covariates.
6. Adapting all of these models and methods for other surveys—e.g., HRS and NHAT.

Summary

- We've developed a method that allows understanding discrete longitudinal data from heterogeneous populations.

Summary

- We've developed a method that allows understanding discrete longitudinal data from heterogeneous populations.
- Modeling a small number of extreme profiles uncovers and summarize typical progressions over time.

Summary

- We've developed a method that allows understanding discrete longitudinal data from heterogeneous populations.
- Modeling a small number of extreme profiles uncovers and summarize typical progressions over time.
- Mixed membership accounts for individual variability.

Summary

- We've developed a method that allows understanding discrete longitudinal data from heterogeneous populations.
- Modeling a small number of extreme profiles uncovers and summarize typical progressions over time.
- Mixed membership accounts for individual variability.
- In this particular application,

Summary

- We've developed a method that allows understanding discrete longitudinal data from heterogeneous populations.
- Modeling a small number of extreme profiles uncovers and summarize typical progressions over time.
- Mixed membership accounts for individual variability.
- In this particular application,
  - Extreme profile trajectories illustrate typical ways of aging.

Summary

- We've developed a method that allows understanding discrete longitudinal data from heterogeneous populations.
- Modeling a small number of extreme profiles uncovers and summarize typical progressions over time.
- Mixed membership accounts for individual variability.
- In this particular application,
  - Extreme profile trajectories illustrate typical ways of aging.
  - Mixed membership acknowledges the fact that not everybody ages the same way!

Summary

- We've developed a method that allows understanding discrete longitudinal data from heterogeneous populations.
- Modeling a small number of extreme profiles uncovers and summarize typical progressions over time.
- Mixed membership accounts for individual variability.
- In this particular application,
  - Extreme profile trajectories illustrate typical ways of aging.
  - Mixed membership acknowledges the fact that not everybody ages the same way!
  - Making individual membership scores dependent on the individual's generation allows to assess changes on the ways of aging.

Summary

- We've developed a method that allows understanding discrete longitudinal data from heterogeneous populations.
- Modeling a small number of extreme profiles uncovers and summarize typical progressions over time.
- Mixed membership accounts for individual variability.
- In this particular application,
  - Extreme profile trajectories illustrate typical ways of aging.
  - Mixed membership acknowledges the fact that not everybody ages the same way!
  - Making individual membership scores dependent on the individual's generation allows to assess changes on the ways of aging.
- General methodology. Can be applied in other settings!

Bibliography

Airoldi, E. M., Blei, D. M., Fienberg, S. E., and Xing, E. P. (2008) Mixed membership stochastic blockmodels. *JMLR*, **9**, 1981–2014.

Airoldi, E. M., Fienberg, S. E., C. Joutard, and Love, T. (2010) Hierarchical Bayesian mixed-membership models and latent pattern discovery. in *Frontier of Statistical Decision Making and Bayesian Analysis*, Chen, Dey, P. Mueller, Sun, and Ye, eds., Springer.

Connor, J. T. (2006) *Multivariate Mixture Models to Describe Longitudinal Patterns of Frailty in American Seniors,* Ph.D. thesis, Department of Statistics & Heinz School. Carnegie Mellon University.

Erosheva, E. A. (2002) *Grade of membership and latent structures with application to disability survey data*, Ph.D. thesis, Department of Statistics. Carnegie Mellon University.

Erosheva, E. A., Fienberg, S. E. and Lafferty, J. (2004) Mixed membership models of scientific publications. *PNAS*, **101**, Suppl. 1, (2004), 5220–5227.

Erosheva, E. A., Fienberg, S. E. and Joutard, C. (2007) Describing disability through individual-level mixture models for multivariate binary data. *Ann. Appl. Statist.*, 1 502–537.

The End