Finite Mixture Models with Applications

Partha Deb

Hunter College and the Graduate Center, CUNY
NBER

July 2009

Introduction

- The finite mixture model provides a natural representation of heterogeneity in a finite number of latent classes
- It concerns modeling a statistical distribution by a mixture (or weighted sum) of other distributions

Introduction

- The finite mixture model provides a natural representation of heterogeneity in a finite number of latent classes

- It concerns modeling a statistical distribution by a mixture (or weighted sum) of other distributions

- Finite mixture models are also known as
  - latent class models
  - unsupervised learning models

- Finite mixture models are closely related to
  - intrinsic classification models
  - clustering
  - numerical taxonomy

- Heterogeneity of effects for different "classes" of observations
  - wine from different types of grapes
  - healthy and sick individuals
  - normal and complicated pregnancies
  - low and high responses to stress

- Estimating parameters of the distribution of lengths of halibut

- It is known that female halibut is longer, on average, than male fish and that the distribution of lengths is normal
- Gender cannot be determined at measurement
- Then distribution is a 2-component finite mixture of normals

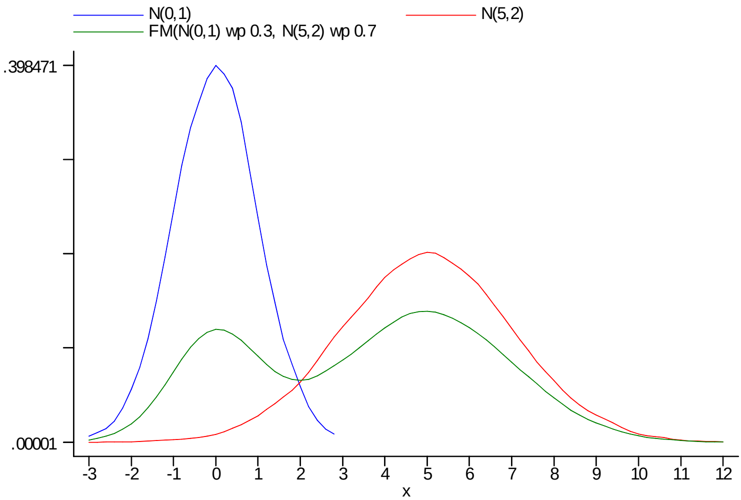
- Estimating parameters of the distribution of lengths of halibut

- It is known that female halibut is longer, on average, than male fish and that the distribution of lengths is normal
- Gender cannot be determined at measurement
- Then distribution is a 2-component finite mixture of normals

- A finite mixture model allows one to estimate:
  - mean lengths of male and female halibut
  - mixing probability

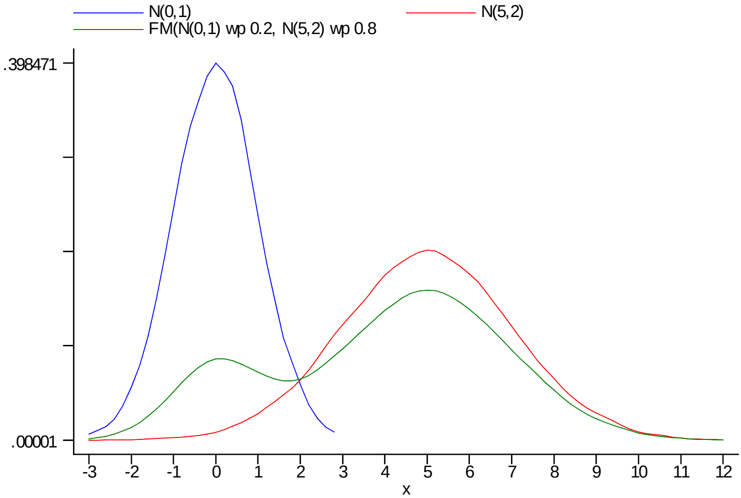
Introduction

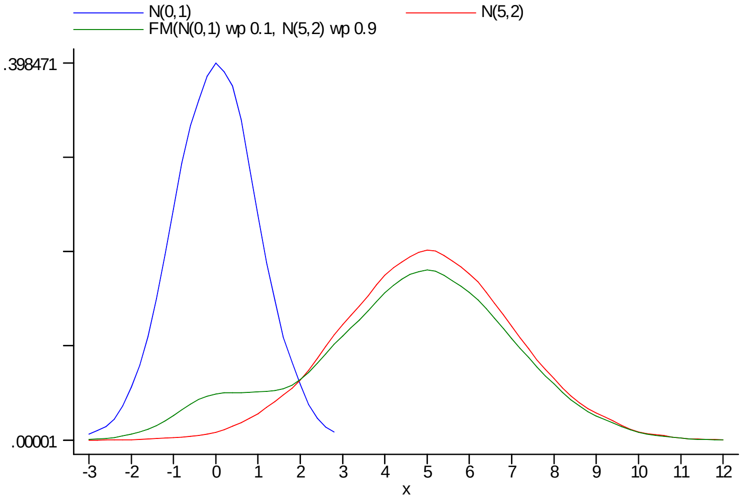
A graphical view



Introduction

A graphical view




N(0,1)
N(5,2)
FM(N(0,1) wp 0.1, N(5,2) wp 0.9
.398471
.00001
-3
-2
-1
0
1
2
3
4
5
6
7
8
9
10
11
12
x

- Characteristics of wine by cultivar

- Characteristics of wine by cultivar
- Infant Birthweight - two types of pregnancies “normal” and “complicated”

- Characteristics of wine by cultivar
- Infant Birthweight - two types of pregnancies “normal” and “complicated”
- Medical Services - two types of consumers “healthy” and “sick”

- Characteristics of wine by cultivar
- Infant Birthweight - two types of pregnancies “normal” and “complicated”
- Medical Services - two types of consumers “healthy” and “sick”
- Public goods experiments - selfish, reciprocal, and altruist

- Characteristics of wine by cultivar
- Infant Birthweight - two types of pregnancies "normal" and "complicated"
- Medical Services - two types of consumers "healthy" and "sick"
- Public goods experiments - selfish, reciprocal, and altruist
- Stock Returns in "typical" and "crisis" times

- Characteristics of wine by cultivar
- Infant Birthweight - two types of pregnancies "normal" and "complicated"
- Medical Services - two types of consumers "healthy" and "sick"
- Public goods experiments - selfish, reciprocal, and altruist
- Stock Returns in "typical" and "crisis" times
- Using somatic cell counts to classify records from healthy or infected goats

- Characteristics of wine by cultivar
- Infant Birthweight - two types of pregnancies “normal” and “complicated”
- Medical Services - two types of consumers “healthy” and “sick”
- Public goods experiments - selfish, reciprocal, and altruist
- Stock Returns in “typical” and “crisis” times
- Using somatic cell counts to classify records from healthy or infected goats
- Models of internet traffic

- FMM is a semiparametric / nonparametric estimator of the density (Lindsay)
- Experience suggests that usually only few latent classes are needed to approximate density well (Heckman)
- In practice FMM are flexible extensions to basic parametric models
  - can generate skewed distributions from symmetric components
  - can generate leptokurtic distributions from mesokurtic ones

Outline of talk

- Introduction

Outline of talk

- Introduction
- Example
    - color of wine

Outline of talk

- Introduction
- Example
  - color of wine
- Model
  - Formulation
  - Estimation
  - Popular densities
  - Properties
  - Model selection
  - Implementation

Outline of talk

- Applications
  - Birthweight and prenatal care – Conway and Deb, JHE 2005
  - Price elasticities of medical care use – Deb and Trivedi, JAE 1997; Deb and Trivedi, JHE 2002
  - Effects of job loss on BMI and alcohol consumption – Deb, Gallo, Ayyagari, Fletcher and Sindelar, working paper 2009

Outline of talk

- Applications
  - Birthweight and prenatal care – Conway and Deb, JHE 2005
  - Price elasticities of medical care use – Deb and Trivedi, JAE 1997; Deb and Trivedi, JHE 2002
  - Effects of job loss on BMI and alcohol consumption – Deb, Gallo, Ayyagari, Fletcher and Sindelar, working paper 2009
- Conclusions

Example
Color of Wine

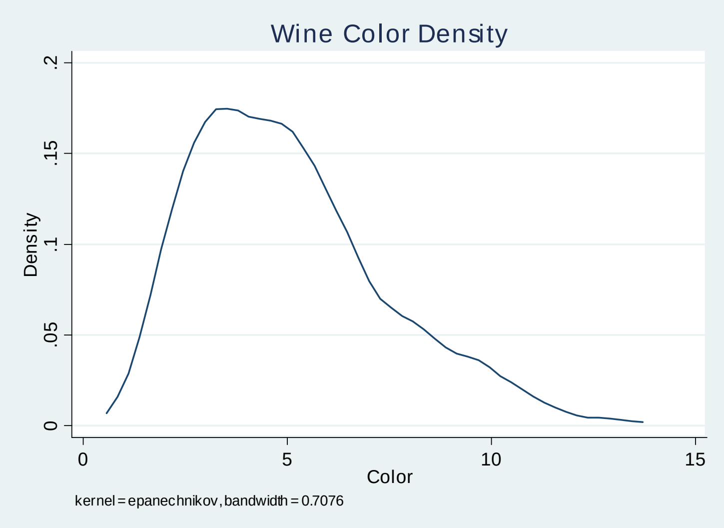
Results of a chemical analysis of wines grown in the same region in Italy but derived from three different cultivars (grape variety)

Data characteristics

Cultivar	Freq.	% of total	Color intensity (mean)
1	59	33.15	5.528
2	71	39.89	3.086
3	48	26.97	7.396
Total	178	100	5.058

Example

Color of Wine



Example

Color of Wine

- Finite mixture of Normals with 3 components

$$
\begin{aligned}
& f(y_i|\theta_1, \theta_2, ..., \theta_C; \pi_1, \pi_2, ..., \pi_C) \\
= & \sum_{j=1}^{C} \pi_j \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{1}{2\sigma_j^2}(y_i - x_i\beta_j)^2\right)
\end{aligned}
$$

Estimates from finite mixture of normals with 3 components

Parameter	component 1	component 2	component 3
Constant	4.929	7.548	2.803
	(0.334)	(0.936)	(0.244)
π	0.365	0.312	0.323
	(0.176)	(0.117)	(0.107)

Data characteristics

Cultivar	Freq.	% of total	Color (mean)
1	59	33.15	5.528
2	71	39.89	3.086
3	48	26.97	7.396
Total	178	100	5.058

Example

Color of Wine

Posterior probability (median)

Cultivar	component 1	component 2	component 3
1	0.737	0.195	9.00e-5
2	0.048	0.023	0.923
3	0.030	0.970	7.54e-14

Data characteristics

Cultivar	Freq.	% of total	Color (mean)
1	59	33.15	5.528
2	71	39.89	3.086
3	48	26.97	7.396
Total	178	100	5.058

- The density function for a C-component finite mixture is

$$f(y|\mathbf{x};\theta_1,\theta_2,...,\theta_C;\pi_1,\pi_2,...,\pi_C) = \sum_{j=1}^{C} \pi_j f_j(y|\mathbf{x};\theta_j)$$

where $0 < \pi_j < 1$, and $\sum_{j=1}^{C} \pi_j = 1$

- The density function for a C-component finite mixture is

$$f(y|\mathbf{x};\theta_1,\theta_2,...,\theta_C;\pi_1,\pi_2,...,\pi_C)=\sum_{j=1}^{C}\pi_j f_j(y|\mathbf{x};\theta_j)$$

where $0<\pi_j<1$, and $\sum_{j=1}^{C}\pi_j=1$

- More generally

$$f(y|\mathbf{x};\mathbf{z};\theta_1,\theta_2,...,\theta_C;\pi_1,\pi_2,...,\pi_C)=\sum_{j=1}^{C}\pi_j(\mathbf{z})f_j(y|\mathbf{x};\theta_j)$$

Model
Estimation

- Maximum likelihood

$$\max_{\pi,\theta} \ln L = \sum_{i=1}^{N} \left(\log\left(\sum_{j=1}^{C} \pi_j f_j(y|\theta_j)\right) \right)$$

- Maximum likelihood

$$\max_{\pi,\theta} \ln L = \sum_{i=1}^{N} \left(\log\left(\sum_{j=1}^{C} \pi_j f_j(y|\theta_j)\right) \right)$$

- Trick to ensure $0 < \pi_j < 1$, and $\sum_{j=1}^{C} \pi_j = 1$

$$\pi_j = \frac{\exp(\gamma_j)}{\exp(\gamma_1) + \exp(\gamma_2) + \ldots + \exp(\gamma_{C-1}) + 1}$$

- Maximum likelihood

$$\max_{\boldsymbol{\pi},\boldsymbol{\theta}} \ln L = \sum_{i=1}^{N} \left(\log(\sum_{j=1}^{C} \pi_j f_j(y|\theta_j) \right)$$

- Trick to ensure $0 < \pi_j < 1$, and $\sum_{j=1}^{C} \pi_j = 1$

$$\pi_j = \frac{\exp(\gamma_j)}{\exp(\gamma_1) + \exp(\gamma_2) + \ldots + \exp(\gamma_{C-1}) + 1}$$

- EM
- Bayesian MCMC

Model

Popular mixture component densities

- Normal (Gaussian)
- Poisson
- Gamma
- Negative Binomial
- Student-t
- Weibull

- Conditional mean:

$$E(y_i|\mathbf{x}_i) = \sum_{j=1}^{C} \pi_j \lambda_j \text{ where } \lambda_j = E_j(y_i|\mathbf{x}_i)$$

Model

Some basic properties

- Conditional mean:

$$\mathrm{E}(y_i|\mathbf{x}_i) = \sum_{j=1}^{C} \pi_j \lambda_j \text{ where } \lambda_j = \mathrm{E}_j(y_i|\mathbf{x}_i)$$

- Marginal effects:

$$\frac{\partial \mathrm{E}_j(y_i|\mathbf{x}_i)}{\partial \mathbf{x}_i} = \frac{\partial \lambda_j}{\partial \mathbf{x}_i} \longrightarrow \text{within component}$$

$$\frac{\partial \mathrm{E}(y_i|\mathbf{x}_i)}{\partial \mathbf{x}_i} = \sum_{j=1}^{C} \pi_j \frac{\partial \lambda_j}{\partial \mathbf{x}_i} \longrightarrow \text{overall}$$

Model

Some basic properties

- Prior probability that observation y_i belongs to component c:

$$\Pr[y_i \in \text{population } c | \mathbf{x}_i, \boldsymbol{\theta}] = \pi_c$$

$c = 1, 2, ..C$

- Prior probability that observation y_i belongs to component c:

$$\Pr[y_i \in \text{population } c | \mathbf{x}_i, \boldsymbol{\theta}] = \pi_c$$

$c = 1, 2, ..C$

- Posterior probability that observation y_i belongs to component c:

$$\Pr[y_i \in \text{population } c | \mathbf{x}_i, y_{i;}\boldsymbol{\theta}] = \frac{\pi_c f_c(y_i | \mathbf{x}_i, \boldsymbol{\theta}_c)}{\sum_{j=1}^{C} \pi_j f_j(y_i | \mathbf{x}_i, \boldsymbol{\theta}_j)}$$

$c = 1, 2, ..C$

- The number of components has to be specified - we usually have little theoretical guidance
- Even if prior theory suggests a particular number of components we may not be able to reliably distinguish between some of the components
- In some cases additional components may simply reflect the presence of outliers in the data
- Likelihood function may have multiple local maxima

- Parameterize $\gamma_j = Z\alpha_j$ in

$$\pi_j = \frac{\exp(\gamma_j)}{\exp(\gamma_1) + \exp(\gamma_2) + \ldots + \exp(\gamma_{C-1}) + 1}$$

- Parameterizing mixing probabilities
  - may lead to finite sample identification issues
  - may lead to computational difficulties

Model

Selecting number of components

- Estimate models with 2 and then more components
- At each step calculate

$$AIC = -2\log(L) + 2K$$
$$BIC = -2\log(L) + K\log(N)$$

- Pick the model with the smallest AIC, BIC

- Stata package `fmm`

```
fmm depvar [indepvars] [if] [in] [weight],
components(#) mixtureof(density)
```

- where density is one of

```
gamma
negbin1
negbin2
normal
poisson
studentt
```

- `predict` and `mfx` give predictions and marginal effects of means, component means, prior and posterior probabilities

- Expanding prenatal care should improve infant health!
- Clinical interventions show positive effects
- Research using observational data typically finds weak, if any, effects

- Expanding prenatal care should improve infant health!
- Clinical interventions show positive effects
- Research using observational data typically finds weak, if any, effects

- Why?

  1. difficulties in measuring prenatal care
  2. difficulties in modeling its endogeneity
  3. there are essentially two kinds of pregnancies, "complicated" and "normal" ones

- Study of the effect of the onset (timing) of prenatal care on birthweight

Applications

Infant Birthweight and Prenatal Care

- Study of the effect of the onset (timing) of prenatal care on birthweight

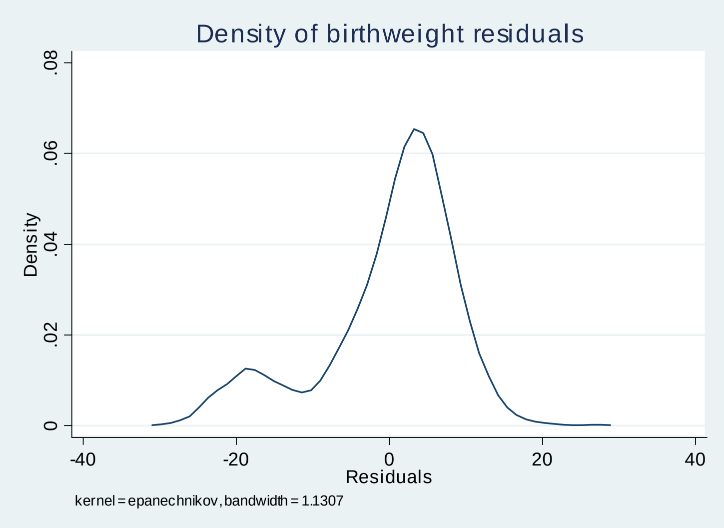
- Complicated pregnancies typically entail a large amount of prenatal care, but yield poorer outcomes
- Clinical evidence suggests that such births are quite difficult to prevent
- Observed factors such as prenatal care and other maternal behaviors may have different effects on each type of pregnancy
- Combining these pregnancies with normally progressing pregnancies could lead prenatal care to appear ineffective

- Study of the effect of the onset (timing) of prenatal care on birthweight
- Complicated pregnancies typically entail a large amount of prenatal care, but yield poorer outcomes
- Clinical evidence suggests that such births are quite difficult to prevent
- Observed factors such as prenatal care and other maternal behaviors may have different effects on each type of pregnancy
- Combining these pregnancies with normally progressing pregnancies could lead prenatal care to appear ineffective
- Econometric strategy: Use a finite mixture model

- Data from the National Maternal and Infant Health Survey
  - Contains information on 26,355 women who were pregnant in 1988
  - Oversamples fetal deaths and infant deaths
  - Oversamples blacks and low birthweight babies
  - Sample of singleton live births:
  - 3,350 nonHispanic blacks
  - 3,245 nonHispanic whites
- Number of observations: 5,219
- Number of covariates: 12

Applications

Infant Birthweight and Prenatal Care



Parameter estimates			
Variable	OLS	FMM	
		component 1	component 2
onsethat	-0.501**	-0.294*	0.006
	(0.183)	(0.127)	(0.234)

Parameter estimates

Variable	OLS	FMM	
		component 1	component 2
onsethat	-0.501**	-0.294*	0.006
	(0.183)	(0.127)	(0.234)
black	-1.213**	-1.231**	-0.775*
	(0.312)	(0.215)	(0.393)

Parameter estimates

Variable	OLS	FMM	
		component 1	component 2
onsethat	-0.501**	-0.294*	0.006
	(0.183)	(0.127)	(0.234)
black	-1.213**	-1.231**	-0.775*
	(0.312)	(0.215)	(0.393)
edu	0.353**	0.292**	0.040
	(0.074)	(0.050)	(0.102)

Parameter estimates

Variable	OLS	FMM	
		component 1	component 2
onsethat	-0.501**	-0.294*	0.006
	(0.183)	(0.127)	(0.234)
black	-1.213**	-1.231**	-0.775*
	(0.312)	(0.215)	(0.393)
edu	0.353**	0.292**	0.040
	(0.074)	(0.050)	(0.102)
numdead	-1.181**	-0.170	-0.585**
	(0.163)	(0.117)	(0.171)

Parameter estimates

Variable	OLS	FMM	
		component 1	component 2
onsethat	-0.501**	-0.294*	0.006
	(0.183)	(0.127)	(0.234)
black	-1.213**	-1.231**	-0.775*
	(0.312)	(0.215)	(0.393)
edu	0.353**	0.292**	0.040
	(0.074)	(0.050)	(0.102)
numdead	-1.181**	-0.170	-0.585**
	(0.163)	(0.117)	(0.171)
π		0.864	0.136
se(π)		(0.005)	(0.005)

- Remarkable robustness of the finite mixture model across samples, races and weighting schemes

- Remarkable robustness of the finite mixture model across samples, races and weighting schemes
- Clear message – getting prenatal care one week earlier significantly increases birth weights in 'normal' pregnancies by 30-50 grams per week

- Remarkable robustness of the finite mixture model across samples, races and weighting schemes
- Clear message – getting prenatal care one week earlier significantly increases birth weights in 'normal' pregnancies by 30-50 grams per week
- Estimated probability of having a 'normal' pregnancy is 0.856 to 0.873 across samples, races and weighting schemes

- Remarkable robustness of the finite mixture model across samples, races and weighting schemes
- Clear message – getting prenatal care one week earlier significantly increases birth weights in 'normal' pregnancies by 30-50 grams per week
- Estimated probability of having a 'normal' pregnancy is 0.856 to 0.873 across samples, races and weighting schemes
- Estimated magnitudes are on target with the medical literature

- Study of the price (insurance) elasticity of demand for healthcare services

- Study of the price (insurance) elasticity of demand for healthcare services
- The two-part model (TPM) is the methodological cornerstone of empirical analysis in the analysis of medical care use
    - "... the decision to receive some care is largely the consumer's, while the physician influences the decision about the level of care" (Manning et al. 1981, p. 109)
    - "... while at the first stage it is the patient who determines whether to visit the physician, it is essentially up to the physician to determine the intensity of the treatment" (Pohlmeier and Ulrich, 1995, p. 340)
    - "...where the first part relates to the patient who decides whether to contact the physician and the second to the decision about repeated visits and/or referrals, which is determined largely by the preferences of the physician" (Gerdtham, 1997, p. 308)

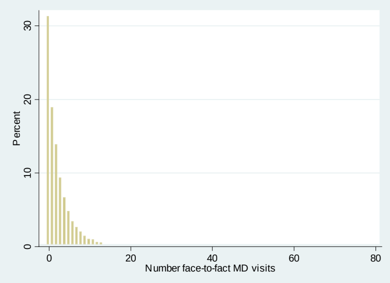
- The sharp dichotomy between users and nonusers may not be tenable in the case of typical cross-sectional data-sets

- The sharp dichotomy between users and nonusers may not be tenable in the case of typical cross-sectional data-sets
- A better distinction for such data may be between infrequent users and frequent users of medical care

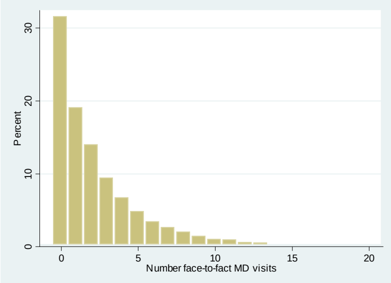
- The sharp dichotomy between users and nonusers may not be tenable in the case of typical cross-sectional data-sets
- A better distinction for such data may be between infrequent users and frequent users of medical care

- The finite mixture model provides a better framework for distinguishing between infrequent and frequent users

- Data from the Rand Health Insurance Experiment (RHIE)
  - conducted by the RAND Corporation from 1974 to 1982
  - individuals were randomized into insurance plans
  - widely regarded as the basis of the most reliable estimates of price elasticities
  - collected from about 8,000 enrollees in 2,823 families from six sites across the country
  - each family was enrolled in one of fourteen different insurance plans for either three or five years
  - the FFS plans ranged from free care to 95 % coinsurance
- Data from all 5 years of the experiment
- Number of observations: 20,186
- Number of covariates: 17



Percent
0
10
20
30
0
20
40
60
80
Number face-to-fact MD visits



Percent
0
10
20
30
0
5
10
15
20
Number face-to-fact MD visits

Explanatory variables	
LC	ln(coinsurance+1), 0≤coinsurance≤100
IDP	1 if individual deductible plan, 0 otherwise
LPI	f(annual participation incentive payment)
FMDE	f(maximum dollar expenditure)
LINC	ln(family income)
LFAM	ln(family size)
EDUCDEC	education of the household head in years
PHYSLIM	1 if the person has a physical limitation
NDISEASE	number of chronic diseases
HLTHG, F, P	self-rated health
AGE, FEMALE, CHILD, FEMALE * CHILD, BLACK	

- The density of the C-component finite mixture is specified as

$$f(y_i|\theta_1, \theta_2, ..., \theta_C; \pi_1, \pi_2, ..., \pi_C)$$
$$= \sum_{j=1}^{C} \pi_j \frac{\Gamma(y_i + \psi_{j,i})}{\Gamma(\psi_{j,i})\Gamma(y_i + 1)} \left(\frac{\psi_{j,i}}{\lambda_{j,i} + \psi_{j,i}}\right)^{\psi_{c,i}} \left(\frac{\lambda_{j,i}}{\lambda_{j,i} + \psi_{j,i}}\right)^{y_i}$$

where $\lambda = \exp(x\beta)$ and $\psi = (1/\alpha)\lambda^k$

- $k = 1$ NB-2
- $k = 0$ NB-1
- $k = 0$ fits best

Parameter Estimates

	nb1	fmm nb1	
		component 1	component 2
logc	-0.149*	-0.203*	-0.024
	(0.012)	(0.020)	(0.031)

Parameter Estimates

	nb1	fmm nb1	
		component 1	component 2
logc	-0.149*	-0.203*	-0.024
	(0.012)	(0.020)	(0.031)
educdec	0.023*	0.027*	0.015
	(0.003)	(0.005)	(0.010)

Parameter Estimates

	nb1	fmm nb1	
		component 1	component 2
logc	-0.149*	-0.203*	-0.024
	(0.012)	(0.020)	(0.031)
educdec	0.023*	0.027*	0.015
	(0.003)	(0.005)	(0.010)
disea	0.021*	0.019*	0.033*
	(0.001)	(0.002)	(0.004)

Parameter Estimates

	nb1	fmm nb1	
		component 1	component 2
logc	-0.149*	-0.203*	-0.024
	(0.012)	(0.020)	(0.031)
educdec	0.023*	0.027*	0.015
	(0.003)	(0.005)	(0.010)
disea	0.021*	0.019*	0.033*
	(0.001)	(0.002)	(0.004)
π		0.802	0.198
		(0.037)	(0.037)
log L	–42405	–42037	
BIC	84999	84461	

Marginal Effects

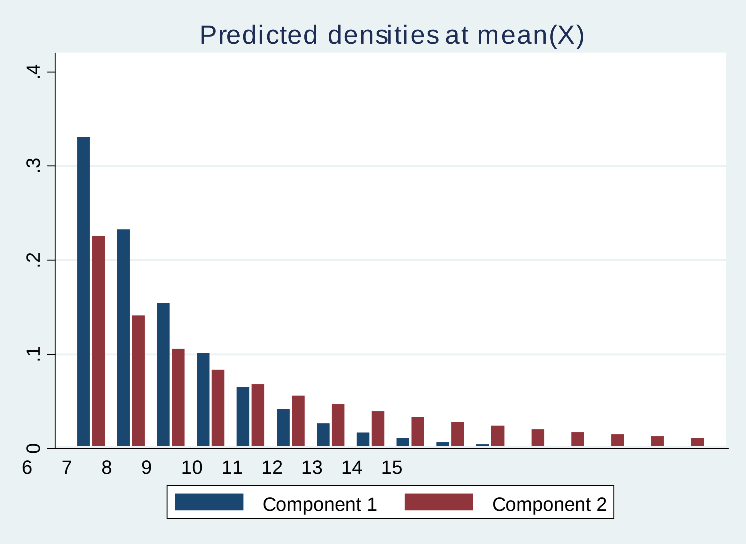
	nb1	fmm nb1		
	overall	overall	component 1	component 2
$E(y\|\bar{\mathbf{x}})$	2.561	2.511	1.887	5.038
logc	-0.382*	-0.331*	-0.382*	-0.121
	(0.030)	(0.032)	(0.032)	(0.158)

Marginal Effects

	nb1	fmm nb1		
	overall	overall	component 1	component 2
$E(y\|\bar{\mathbf{x}})$	2.561	2.511	1.887	5.038
logc	-0.382*	-0.331*	-0.382*	-0.121
	(0.030)	(0.032)	(0.032)	(0.158)
educdec	0.058*	0.056*	0.052*	0.073
	(0.007)	(0.009)	(0.008)	(0.053)

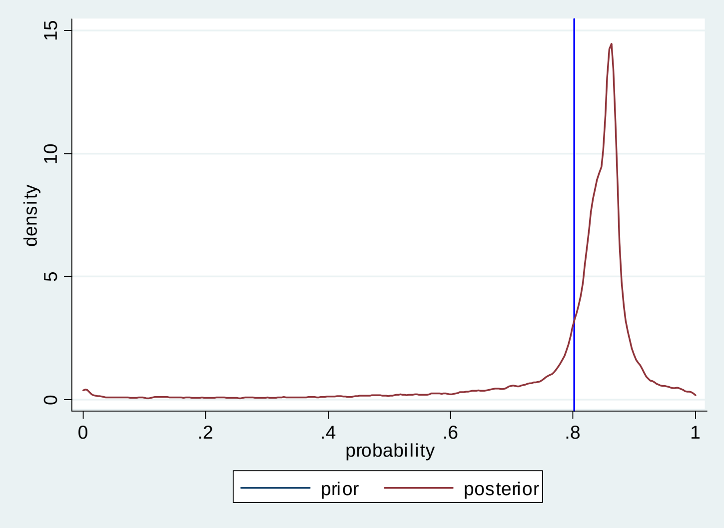
Marginal Effects

	nb1	fmm nb1		
	overall	overall	component 1	component 2
$E(y\|\bar{\mathbf{x}})$	2.561	2.511	1.887	5.038
logc	-0.382*	-0.331*	-0.382*	-0.121
	(0.030)	(0.032)	(0.032)	(0.158)
educdec	0.058*	0.056*	0.052*	0.073
	(0.007)	(0.009)	(0.008)	(0.053)
disea	0.054*	0.062*	0.036*	0.167*
	(0.003)	(0.004)	(0.004)	(0.024)



Predicted densities at mean(X)
0
.1
.2
.3
.4
6
7
8
9
10
11
12
13
14
15
Component 1
Component 2

Prior and posterior probabilities



- Study of effects of business closures on alcohol use and body mass index (BMI)

- Study of effects of business closures on alcohol use and body mass index (BMI)
- Job loss frequently involves a succession of stress-laden experiences
- Evidence on the health effects of job loss is mixed
- Evidence on changes in weight associated with unemployment is similarly ambiguous
- Evidence on the effects of job loss on mental health status is more definitive

- There is wide variation in the individual behavioral response to the stress of job loss
- Greater alcohol or food consumption could conceivably counterbalance neuro- or emotion-regulatory disturbances
- But unemployment introduces discretionary time which may be used to pursue health-promoting behaviors
- Income losses could constrain job losers to certain food choices or limit alcohol purchases

- There is wide variation in the individual behavioral response to the stress of job loss
- Greater alcohol or food consumption could conceivably counterbalance neuro- or emotion-regulatory disturbances
- But unemployment introduces discretionary time which may be used to pursue health-promoting behaviors
- Income losses could constrain job losers to certain food choices or limit alcohol purchases

- Econometric strategy: Use a finite mixture model

- Job loss has frequently been represented by layoff or some combination of involuntary termination (e.g., layoff, plant closing, and firing)
- Layoffs and firing are likely to be endogenous
- We use business closing

- Data from the Health and Retirement Study (HRS)
  - 12,652 individuals from 7,702 households at baseline
- Data from the first six HRS waves (1992-2002)
- Analysis sample restricted to participants who met the following criteria at the 1992 baseline

  1. were between ages 51 and 61
  2. were working for pay, but not self employed
  3. reported a minimum of two years of continuous employment with the 1992 employer
  4. provided at least one follow-up response

- Limit the sample to study subjects who reported continuous employment in the previous person-spell
- Number of observations: 6,726.

Parameter estimates

Variable	OLS	FMM Normal	
		component1	component2
Business closure	0.081	-0.192	1.083**
	(0.149)	(0.119)	(0.541)

Parameter estimates

Variable	OLS	FMM Normal	
		component1	component2
Business closure	0.081	-0.192	1.083**
	(0.149)	(0.119)	(0.541)
Non-housing net worth	-0.025	0.023	-0.162
	(0.035)	(0.022)	(0.134)

Parameter estimates

Variable	OLS	FMM Normal	
		component1	component2
Business closure	0.081	-0.192	1.083**
	(0.149)	(0.119)	(0.541)
Non-housing net worth	-0.025	0.023	-0.162
	(0.035)	(0.022)	(0.134)
Depressive symptoms	-0.042*	-0.023	-0.084
	(0.023)	(0.019)	(0.098)

Parameter estimates			
Variable	OLS	FMM Normal	
		component1	component2
Business closure	0.081	-0.192	1.083**
	(0.149)	(0.119)	(0.541)
Non-housing net worth	-0.025	0.023	-0.162
	(0.035)	(0.022)	(0.134)
Depressive symptoms	-0.042*	-0.023	-0.084
	(0.023)	(0.019)	(0.098)
Lagged BMI	0.956***	0.989***	0.850***
	(0.007)	(0.005)	(0.035)

Parameter estimates

Variable	OLS	FMM Normal	
		component1	component2
Business closure	0.081	-0.192	1.083**
	(0.149)	(0.119)	(0.541)
Non-housing net worth	-0.025	0.023	-0.162
	(0.035)	(0.022)	(0.134)
Depressive symptoms	-0.042*	-0.023	-0.084
	(0.023)	(0.019)	(0.098)
Lagged BMI	0.956***	0.989***	0.850***
	(0.007)	(0.005)	(0.035)
π		0.806***	
		(0.026)	

Parameter estimates

Variable	Poisson	FMM Poisson	
		component1	component2
Business closure	0.228	0.131	0.844***
	(0.173)	(0.109)	(0.242)

Parameter estimates

Variable	Poisson	FMM Poisson	
		component1	component2
Business closure	0.228	0.131	0.844***
	(0.173)	(0.109)	(0.242)
Non-housing net worth	0.082***	0.119***	-0.211
	(0.028)	(0.034)	(0.172)

Parameter estimates

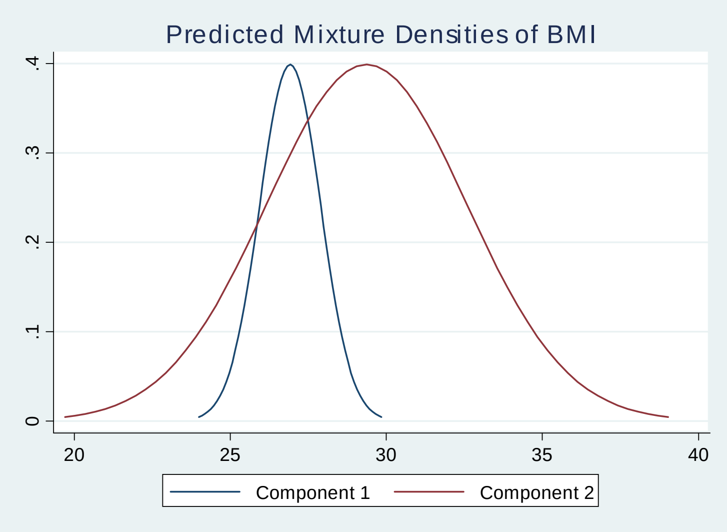
Variable	Poisson	FMM Poisson	
		component1	component2
Business closure	0.228	0.131	0.844***
	(0.173)	(0.109)	(0.242)
Non-housing net worth	0.082***	0.119***	-0.211
	(0.028)	(0.034)	(0.172)
Depressive symptoms	-0.086***	-0.146***	0.064
	(0.029)	(0.050)	(0.130)

Parameter estimates

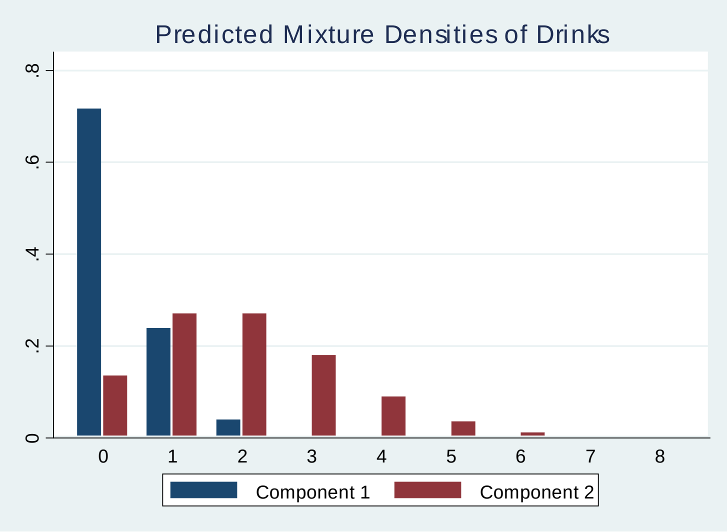
Variable	Poisson	FMM Poisson	
		component1	component2
Business closure	0.228	0.131	0.844***
	(0.173)	(0.109)	(0.242)
Non-housing net worth	0.082***	0.119***	-0.211
	(0.028)	(0.034)	(0.172)
Depressive symptoms	-0.086***	-0.146***	0.064
	(0.029)	(0.050)	(0.130)
Lagged number of drinks	1.587***	1.986***	0.191
	(0.040)	(0.067)	(0.189)

Parameter estimates

Variable	Poisson	FMM Poisson	
		component1	component2
Business closure	0.228	0.131	0.844***
	(0.173)	(0.109)	(0.242)
Non-housing net worth	0.082***	0.119***	-0.211
	(0.028)	(0.034)	(0.172)
Depressive symptoms	-0.086***	-0.146***	0.064
	(0.029)	(0.050)	(0.130)
Lagged number of drinks	1.587***	1.986***	0.191
	(0.040)	(0.067)	(0.189)
π		0.939***	
		(0.009)	



Predicted Mixture Densities of BMI
0
.1
.2
.3
.4
20
25
30
35
40
Component 1
Component 2



Predicted Mixture Densities of Drinks
0
.2
.4
.6
.8
0
1
2
3
4
5
6
7
8
Component 1
Component 2

Determinants of posterior probability of being in component 2: BMI

Variable	(1)	(2)	(3)
Married	-0.007	-0.000	0.000
	(0.008)	(0.008)	(0.008)
Non-housing net worth	-0.019***	-0.017***	-0.017***
	(0.004)	(0.005)	(0.005)
Depressive symptoms	0.011***	0.010***	0.010***
	(0.003)	(0.003)	(0.004)
Non-white		0.017*	0.017*
		(0.010)	(0.010)
Male		-0.025***	-0.025***
		(0.007)	(0.007)
Risk averse	0.004	0.003	0.003
	(0.004)	(0.004)	(0.004)

Determinants of posterior probability of being in component 2: Drinks

Variable	(1)	(2)	(3)
Married	-0.004	-0.012**	-0.012**
	(0.006)	(0.006)	(0.006)
Non-housing net worth	0.006*	0.006	0.006*
	(0.003)	(0.004)	(0.004)
Depressive symptoms score	-0.001	-0.001	-0.001
	(0.003)	(0.003)	(0.003)
Non-white		-0.007	-0.007
		(0.007)	(0.007)
Male		0.032***	0.033***
		(0.006)	(0.006)
Risk averse	-0.007**	-0.006**	-0.006**
	(0.003)	(0.003)	(0.003)

Conclusions

- Finite mixture models are a useful way to model unobserved heterogeneity

Conclusions

- Finite mixture models are a useful way to model unobserved heterogeneity
- FMM can uncover otherwise hidden relationships

Conclusions

- Finite mixture models are a useful way to model unobserved heterogeneity
- FMM can uncover otherwise hidden relationships

Conclusions

- Finite mixture models are a useful way to model unobserved heterogeneity
- FMM can uncover otherwise hidden relationships
- FMM can be applied when outcomes are continuous or discrete

Conclusions

- Finite mixture models are a useful way to model unobserved heterogeneity
- FMM can uncover otherwise hidden relationships
- FMM can be applied when outcomes are continuous or discrete
    - but not for binary or “severely” limited outcomes

Conclusions

- Finite mixture models are a useful way to model unobserved heterogeneity
- FMM can uncover otherwise hidden relationships
- FMM can be applied when outcomes are continuous or discrete
  - but not for binary or "severely" limited outcomes

Conclusions

- Finite mixture models are a useful way to model unobserved heterogeneity
- FMM can uncover otherwise hidden relationships
- FMM can be applied when outcomes are continuous or discrete
  - but not for binary or "severely" limited outcomes
- Extensions of FMM to panels are in the works

Conclusions

- Finite mixture models are a useful way to model unobserved heterogeneity
- FMM can uncover otherwise hidden relationships
- FMM can be applied when outcomes are continuous or discrete
  - but not for binary or "severely" limited outcomes
- Extensions of FMM to panels are in the works
  - Random effects models can be estimated with standard software using Mundlak-type specifications

Conclusions

- Finite mixture models are a useful way to model unobserved heterogeneity
- FMM can uncover otherwise hidden relationships
- FMM can be applied when outcomes are continuous or discrete
  - but not for binary or "severely" limited outcomes
- Extensions of FMM to panels are in the works
  - Random effects models can be estimated with standard software using Mundlak-type specifications
  - Fixed effects models are being worked on