Maximum likelihood estimation of hidden Markov models for continuous longitudinal data with missing responses and dropout

Fulvia Pennoni

Department of Statistics and Quantitative Methods - University of Milano-Bicocca (IT)

fulvia.pennoni@unimib.it

Francesco Bartolucci and **Silvia Pandolfi**

Department of Economics - University of Perugia (IT)

The 11th Conference of the The Asian Regional Section of the International Association for Statistical Computing (IASC-ARS)

UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA

Outline

- We propose a *hidden (or latent) Markov model* (HM; Bartolucci, *et al.*, 2013, Zucchini *et al.*, 2016) for *multivariate* longitudinal continuous responses to deal with three different types of missing pattern (Little and Rubin, 2002)

- We deal with the problem of *missing data* having the following patterns:

  1. *partially missing* outcomes at a given time occasion
  2. *completely missing outcomes* at a given time occasion (intermittent pattern)
  3. *dropout* before the end of the period of observation (monotone pattern)

- The HM model may be seen as *an extension of* finite mixture model of multivariate Gaussian distributions (McLachlan and Peel, 2000)

- For the first two types of missingness we develop an inferential approach based on the missing at random, MAR assumption
- We handle the third type of missingness (dropout) as *informative* or *non-ignorable*, including an *extra absorbing hidden state*
- To estimate the model parameters we extend the *Expectation-Maximization* (EM; Baum *et al.*, 1970; Dempster *et al.*, *1977*) algorithm usually employed to estimate the HM model without missing data
- The proposal is illustrated by a *Monte Carlo simulation study* and an application based on historical *data on primary biliary cholangitis*

Continuous Hidden Markov model

- Notation:
  - n: *sample size*
  - r: number of *continuous response variables*
  - $\boldsymbol{Y}_{it} = (Y_{i1t}, \ldots, Y_{irt})'$: vector of r continuous response variables for subject i $(i = 1, \ldots, n)$ at occasion t $(t = 1, \ldots, T_i)$
  - *Unbalanced panel data*: each subject i has a specific number of *time occasions* T_i, so that $t = 1, \ldots, T_i$ $(i = 1, \ldots, n)$
  - $\boldsymbol{Y}_i$: vector obtained by stacking $\boldsymbol{Y}_{it}$ for $t = 1, \ldots, T$
  - U_i: discrete latent variable affecting the distribution of the response variables
    - Its distribution is assumed to follow a first order Markov chain with state-space $\{1, \ldots, k\}$, where k is the number of hidden states

- *Model assumptions*:

  - A conditional Gaussian distribution is assumed for the responses

$$\boldsymbol{Y}_i | U_i = u \sim N(\boldsymbol{\mu}_u, \Sigma_u) \quad u = 1, \ldots, k$$

    - conditional *mean* vectors $\boldsymbol{\mu}_u$, $u = 1, \ldots, k$
    - *variance-covariance* matrix Σ_u, which may also be constant across states

  - We consider the case in which *covariates* denoted as $\boldsymbol{x}_i$ may affect

    - *initial probabilities* denoted as $\pi_{iu|\boldsymbol{x}} = p(U_1 = u|\boldsymbol{x}_i)$ for component u and individual i
    - *transition probabilities* $\pi_{i,u|\bar{u}\boldsymbol{x}} = p(U_{it} = u|U_{i,t-1} = \bar{u}, \boldsymbol{x}_{it})$ with $i = 1, \ldots, n$

Proposed model formulation

- *Informative dropout*: Indicator variable

$$D_{it} = \begin{cases} 0 & \text{if unit } i \text{ is still in the panel at occasion } t \\ 1 & \text{if unit } i \textit{ dropped out} \text{ at occasion } t \end{cases}$$

  - We assume the *conditional probabilities*:

$$P(D_{it} = d \mid U_{it} = u) = \begin{cases} 1 & \text{with } d = 0 \text{ and } u = 1, \ldots, k \\ & \text{or } d = 1 \text{ and } u = k+1 \\ 0 & \text{otherwise} \end{cases}$$

  - *Transition probabilities* are considered with an additional $(k+1)$-th *absorbing hidden state*, under the constraints:

$$\pi_{k+1|k+1,\boldsymbol{x}} = 1, \quad \pi_{u|k+1,\boldsymbol{x}} = 0, \quad u = 1, \ldots, k,$$

  once an individual reaches the $(k+1)$-th latent state, he/she yields *missing values* until the end of the study

- *Intermittent missing responses*:
  - We *partition* the response vector: $\boldsymbol{Y}_{it} = (\boldsymbol{Y}_{it}^{o}, \boldsymbol{Y}_{it}^{m})'$, where $\boldsymbol{Y}_{it}^{o}$ is observed and $\boldsymbol{Y}_{it}^{m}$ is missing so that

$$\boldsymbol{\mu}_u = \begin{pmatrix} \boldsymbol{\mu}_u^m \\ \boldsymbol{\mu}_u^o \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma^{oo} & \Sigma^{om} \\ \Sigma^{mo} & \Sigma^{mm} \end{pmatrix}$$

  - For the *observed responses* we have

$$\boldsymbol{Y}_{it}^{o} \mid U_{it} = u, D_{it} = 0 \sim N(\boldsymbol{\mu}_u^o, \Sigma^{oo}), \quad u = 1, \ldots, k.$$

  - *The manifest distribution for the observed data*:

$$f(\boldsymbol{d}_i, \mathcal{Y}_i^o) = \sum_{\boldsymbol{u}_i} \left[\prod_{t=1}^{T_i} f(\boldsymbol{y}_{it}^o | d_{it}, u_{it})\, p(d_{it} | u_{it}) \right] \left(\pi_{u_{i\mathbf{1}}|\boldsymbol{x}} \prod_{t=2}^{T_i} \pi_{u_{it}|u_{i,t-\mathbf{1}},\boldsymbol{x}} \right)$$

    - $\mathcal{Y}_i^o = \{\boldsymbol{y}_{it}^o, t = 1, \ldots, T_i : d_{it} = 0\}$: set of vectors $\boldsymbol{y}_{it}^o$ observed when $d_{it} = 0$, for $i = 1, \ldots, n$
    - $\boldsymbol{d}_i$: observed vector of indicator variables D_{it} for individual i

- *Latent model*:
  - We adopt a *multinomial parameterization* for the initial and transition Markov chain probabilities
  - For the *initial probabilities* we assume

$$\log \frac{\pi_{iu|\boldsymbol{x}}}{\pi_{i1|\boldsymbol{x}}} = \beta_{0u} + \boldsymbol{x}'_{i1}\boldsymbol{\beta}_{1u}, \quad u = 2, \ldots, k,$$

  - For the *transition probabilities* we assume

$$\log \frac{\pi_{i,u|\bar{u}\boldsymbol{x}}}{\pi_{i,\bar{u}|\bar{u}\boldsymbol{x}}} = \gamma_{0\bar{u}u} + \boldsymbol{x}'_{it}\boldsymbol{\gamma}_{1\bar{u}u}, \quad \bar{u} = 1, \ldots, k, \; u = 1, \ldots, k+1, \; \bar{u} \neq u.$$

where there are parameter vectors to be estimated and those in $\boldsymbol{\gamma}_{\bar{u}u} = (\gamma_{0\bar{u}u}, \boldsymbol{\gamma}'_{1\bar{u}u})'$ are properly constrained *to avoid transitions* from the latent absorbing state

Model inference

- Assuming independence between sample units, the *log-likelihood* referred to the observed data can be written as

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log f(\boldsymbol{d}_i, \mathcal{Y}_i^o)$$

  - $\boldsymbol{\theta}$: vector of all *model parameters*
  - $f(\boldsymbol{d}_i, \mathcal{Y}_i^o)$: *manifest distribution* of the observed response data

- In order to estimate these parameters, we maximize $\ell(\boldsymbol{\theta})$ by the EM algorithm based on the *complete-data log-likelihood* expressed as the sum of *three components* that are maximized separately:

$$\ell^*(\boldsymbol{\theta}) = \ell_1^*(\boldsymbol{\theta}) + \ell_2^*(\boldsymbol{\theta}) + \ell_3^*(\boldsymbol{\theta})$$

Expectation-Maximization algorithm

- The EM algorithm is based on the complete-data log-likelihood

$$\begin{aligned}\ell^*(\boldsymbol{\theta}) &= \sum_{i=1}^{n} \sum_{\substack{t=1 \\ (d_{it}=0)}}^{T_i} \sum_{u=1}^{k} z_{itu} \log f(\boldsymbol{y}_{it} \mid D_{it} = 0, u) \\ &+ \sum_{i=1}^{n} \sum_{u=1}^{k} z_{i1u} \log \pi_{u|\boldsymbol{x}} \\ &+ \sum_{i=1}^{n} \sum_{t=2}^{T_i} \sum_{\bar{u}=1}^{k+1} \sum_{u=1}^{k+1} z_{it\bar{u}u} \log \pi_{u|\bar{u}\boldsymbol{x}}\end{aligned}$$

- $z_{itu} = I(u_{it} = u)$: indicator variable equal to 1 if individual i is in latent state u at time t
- $z_{it\bar{u}u} = z_{i,t-1,\bar{u}}\, z_{itu}$: indicator variable for the transition from state $\bar{u}$ to state u of individual i at time occasion t

Expectation-Maximization algorithm

E-step: compute the posterior expected value of the dummy variables given the observed data and the current value of the parameters by means of suitable *recursions* so as to obtain the posterior probabilities

$$\begin{aligned} \hat{z}_{itu} &= P(U_{it} = u | \boldsymbol{y}_{it}, \boldsymbol{d}_i, \boldsymbol{x}_{it}), \quad t = 1, \ldots, T_i, \ u = 1, \ldots, k+1 \\ \hat{z}_{it\bar{u}u} &= P(U_{it} = u, U_{i,t-1} = \bar{u} | \boldsymbol{y}_{it}, \boldsymbol{d}_i, \boldsymbol{x}_{it}), \quad t = 2, \ldots, T_i, \ \bar{u}, u = 1, \ldots, k+1 \end{aligned}$$

- When $d_{it} = 1 \implies$ $\hat{z}_{isu} = 0, \quad u = 1, \ldots, k, \quad s = t, \ldots, T_i$
  $\hat{z}_{is,k+1} = 1, \quad s = t, \ldots, T_i$

Expectation-Maximization algorithm

- When $d_{it} = 0$, given the presence of *missing observations assumed as MAR*, the **E-step** also includes the computation of the following expected values:

$$\mathrm{E}(\boldsymbol{Y}_{it} \mid \boldsymbol{y}_{it}^{o}, u) = \begin{pmatrix} \boldsymbol{y}_{it}^{o} \\ \boldsymbol{\mu}_u^{m} + \Sigma^{mo}(\Sigma^{oo})^{-1}(\boldsymbol{y}_{it}^{o} - \boldsymbol{\mu}_u^{o}) \end{pmatrix}$$

$$\begin{aligned} &\mathrm{E}\left[(\boldsymbol{Y}_{it} - \boldsymbol{\mu}_u)(\boldsymbol{Y}_{it} - \boldsymbol{\mu}_u)' \mid \boldsymbol{y}_{it}^{o}, u\right] = \\ &= \mathrm{Var}(\boldsymbol{Y}_{it} \mid \boldsymbol{y}_{it}^{o}) + [E(\boldsymbol{Y}_{it} \mid \boldsymbol{y}_{it}^{o}, u) - \boldsymbol{\mu}_u][E(\boldsymbol{Y}_{it} \mid \boldsymbol{y}_{it}^{o}, u) - \boldsymbol{\mu}_u]' \end{aligned}$$

where

$$\mathrm{Var}(\boldsymbol{Y}_{it} | \boldsymbol{y}_{it}^{o}) = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \Sigma^{mm} - \Sigma^{mo}(\Sigma^{oo})^{-1}\Sigma^{om} \end{pmatrix}$$

M-step: update the estimates of $\boldsymbol{\theta}$ by maximizing the expected value of $\ell^*(\boldsymbol{\theta})$ obtained at the E-step

We have a closed form solution for the means

$$\boldsymbol{\mu}_u = \frac{\sum_{i=1}^{n} \sum_{\substack{t=1 \\ (d_{it}=0)}}^{T_i} \hat{z}_{itu} \mathrm{E}(\boldsymbol{Y}_{it} | \boldsymbol{y}_{it}^o, u)}{\sum_{i=1}^{n} \sum_{\substack{t=1 \\ (d_{it}=0)}}^{T_i} \hat{z}_{itu}}, \quad u = 1, \ldots, k$$

$$\boldsymbol{\Sigma} = \frac{1}{\sum_{i=1}^{n} (T_i - d_{i+})} \sum_{i=1}^{n} \sum_{\substack{t=1 \\ (d_{it}=0)}}^{T_i} \sum_{u=1}^{k} \hat{z}_{itu} \Big[\mathrm{Var}(\boldsymbol{Y}_{it} | \boldsymbol{y}_{it}^o) + [\mathrm{E}(\boldsymbol{Y}_{it} | \boldsymbol{y}_{it}^o, u) - \boldsymbol{\mu}_u][\mathrm{E}(\boldsymbol{Y}_{it} | \boldsymbol{y}_{it}^o, u) - \boldsymbol{\mu}_u]' \Big]$$

- With individual covariates in order to update the latent model parameters we maximize the referred components of $\ell^*(\boldsymbol{\theta})$ with respect to $\boldsymbol{B}$ and Γ, by a *Newton-Raphson* algorithm

Other features of the estimation

- **Initialization** of the model parameters: a *multi-start strategy* is employed combining deterministic and random initializations of the EM algorithm
- **Selection of** k: the best model is selected by using the *Bayesian Information Criterion* (BIC, Schwarz, 1978)
- **Standard errors**: a *non-parametric bootstrap* procedure is employed (Davinson and Hinkley, 1997)
- **Dynamic clustering**: *local decoding* is performed to predict the sequence of latent states for each subject according to the estimated posterior probabilities of U_{it} provided by the EM algorithm
- **Code**: The functions written in `R` implemented for the proposal are *publicly available* and they extend the functions the package ***LMest*** (Bartolucci et. al., 2017)

Simulation study

To assess the performance of the inferential methods discussed so far, we conducted a *simulation study* based on:

- $B = 250$ *samples*
- varying number of *hidden states* $k = 2, 3$
- *sample size* $n = 500, 1000$
- number of *time occasions* $T = 5$
- number of *response variables* $r = 3$
- varying *proportions of missing responses* $p_{miss} = 0.01, 0.05, 0.10, 0.25$
- varying *dropout rates* $p_{drop} = 0.01, 0.05, 0.10, 0.25$
- uniform *initial probabilities* $\pi = 1/k$

- *Conditional means*: $k = 2 : \boldsymbol{\mu}_1 = \begin{pmatrix} -2 \\ -2 \\ 0 \end{pmatrix}, \ \boldsymbol{\mu}_2 = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}$

$$k = 3 : \boldsymbol{\mu}_1 = \begin{pmatrix} -2 \\ -2 \\ 0 \end{pmatrix}, \ \boldsymbol{\mu}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \ \boldsymbol{\mu}_3 = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}$$

- *Variance-covariance matrix*: $\Sigma = \begin{pmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{pmatrix}$

- *Homogeneous transition probabilities*:

$$k = 2 : \Pi = \begin{pmatrix} 0.90 & 0.10 \\ 0.10 & 0.90 \end{pmatrix}, \quad k = 3 : \Pi = \begin{pmatrix} 0.90 & 0.09 & 0.01 \\ 0.08 & 0.84 & 0.08 \\ 0.01 & 0.09 & 0.90 \end{pmatrix}$$

Results

- The average *bias*, *sd*, and *rmse* *are small*, they tend to increase when considering a higher number of hidden states
- When the sample size increases, we observe *a reduction* of the *sd* and *rmse*
- Higher rates of *missing data and/or dropout* only slightly affect the results in terms of average *bias*, *sd*, and *rmse*
- In general, model parameters are estimated with *good accuracy*, even in the presence of missing data
- The *absorbing state* is properly estimated as the frequency of dropout increases

Application: Primary Biliary Cholangitis

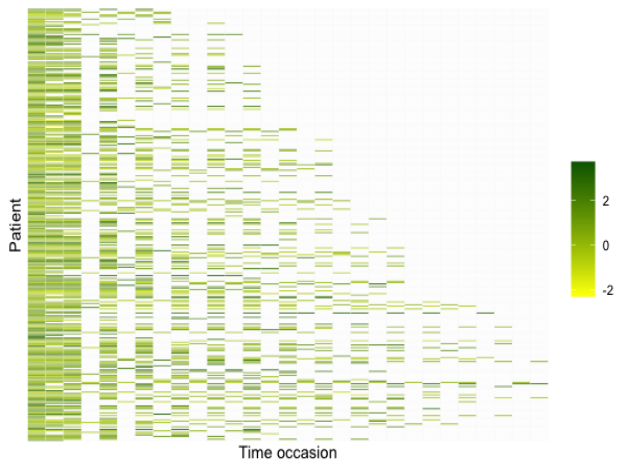
- *Historical data* are collected in a randomized control trial by the *Mayo Clinic* from January 1974 to May 1984 (Murtaugh et al., 1994)
- *p Primary biliary cholangitis* (PBC) is a liver disease causing inflammatory destruction of the bile ducts and eventually leads to cirrhosis of the liver (Dickson et al., 1989)
- Data refers to $n = 312$ *patients* of which 158 have been randomized to D-penicillamine and 154 with placebo
- Follow up times varied and we considered time occasions at 6 months from the baseline, responses as *seven biochemical variables* and three *covariates*: drug use, gender and age
- We analyzed the joint association of the levels of biomarkers with the *risk of death* along with the effects of covariates accounting for *missing observations*, *missing visits* and *dropout* (143 died) in a period of 29 times occasions

Primary Biliary Cholangitis

- The *biomarkers* are the following:
  - Serum bilirubin in mg/dl (values above 1.2 mg/dL are synonymous with liver failure)
  - Serum *cholesterol* in mg/dl
  - Serum albumin in gm/dl (low albumin values may result from liver malfunction)
  - *Platelets* per cubic ml/1000
  - *Prothrombin time* in seconds
  - Alkaline phosphatase in U/liter (high values of alkaline phosphatase can occur in the presence of liver disease)
  - Transaminase (SGOT in U/ml) (in the case of liver damage, an increase in blood SGOT concentration is observed)

Data description

- The following figure shows the *observed values* of each patient for serum bilirubin:



- Almost all the patients had a record at the first three visits: on average they *made 6.2 visits (sd 3.8)*

- *Females* were 88%, *treated* patients with D-penicillamine were 51% and patients with an *age* higher than 50 have an age ranging from 50 to 78.5 years (time-varying covariate)

Results: Means

- We estimate *a HM model* with states ranging from $k = 1$ to $k = 8$ *without covariates* for selecting the a number of hidden states
- To combine goodness of fit and parsimony *we select the HM model with $k = 5$ hidden states* according on the *BIC value*,
- The estimated *conditional means*:

	1	2	3	4	5
Bilirubin	0.136	0.865	*2.020*	-0.432	*2.411*
Cholesterol	5.783	5.496	6.146	5.508	5.415
Albumin	1.270	1.137	1.177	1.279	*0.940*
Platelets	5.556	4.776	5.525	5.477	5.010
Prothrombin	2.339	2.441	2.396	2.363	*2.578*
Alkaline	7.270	6.824	7.611	6.430	*7.033*
Transaminase	4.769	4.664	5.163	4.086	5.070

Results: Transitions

- Estimated *initial and transition probabilities*:

	1	2	3	4	5	*drop*
$\hat{\pi}_u$	0.438	0.121	0.208	0.191	0.042	0.000
$\hat{\pi}_{u\|1}$	0.868	0.056	0.020	0.040	0.012	0.004
$\hat{\pi}_{u\|2}$	0.000	0.922	0.004	0.000	0.066	0.008
$\hat{\pi}_{u\|3}$	0.002	0.008	0.864	0.010	0.094	*0.022*
$\hat{\pi}_{u\|4}$	0.021	0.003	0.000	0.939	0.036	0.001
$\hat{\pi}_{u\|5}$	0.000	0.014	0.014	0.053	0.654	*0.265*
$\hat{\pi}_{u\|drop}$	0.000	0.000	0.000	0.000	0.000	1.000

- The *5th state* includes 4.2% patients with worse health conditions and the highest probability towards dropout is the 5th
- *Higher values* of serum *bilirubin* and *prothrombin* and *lower albumin* values are strongly related to the risk for death

Results: Covariates effects

- Estimated *regression parameters* for the initial probabilities

Effect	$\hat{\beta}_{12}$	$\hat{\beta}_{13}$	$\hat{\beta}_{14}$	$\hat{\beta}_{15}$
Intercept	-5.028**	-0.356	-4.403*	-13.506**
Drug	0.544	-0.322	0.341	-0.297
Female	-0.124	-0.412	1.204	6.634**
Age	0.068**	0.002	0.046*	0.093*

- The estimated *gender log-odds* relative to the 5th state is positive and significant indicating that the probability of being in the 5th state at the beginning of the study is higher for females with respect to males
- The log-odds related to *age* are positive, indicating that, at the baseline, older patients generally tend to belong to the other states with respect to the 1st

Results: Covariates effects

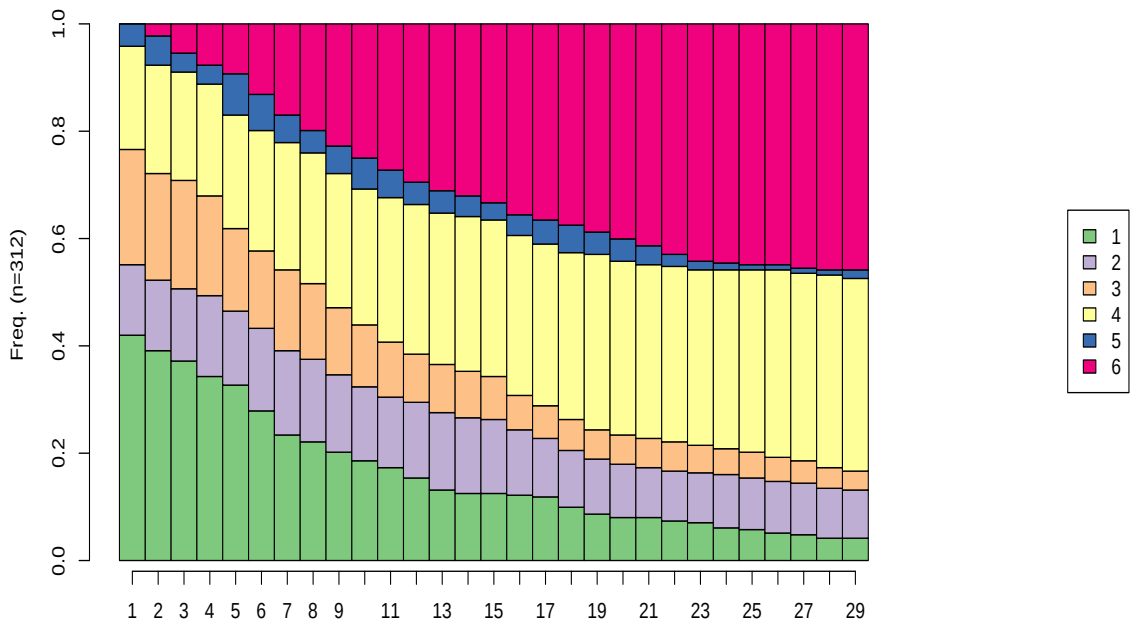
- The regression coefficients affecting the *transition probabilities* are also estimated, for example we show the average of the estimated transitions for *gender*: the matrix for females is more persistent

female	1	2	3	4	5	*drop*
$\hat{\pi}_{u\|1}$	*0.891*	0.046	0.020	0.039	0.000	0.004
$\hat{\pi}_{u\|2}$	0.000	*0.930*	0.000	0.000	0.065	0.005
$\hat{\pi}_{u\|3}$	0.002	0.009	*0.871*	0.012	0.083	0.023
$\hat{\pi}_{u\|4}$	0.000	0.003	0.000	*0.994*	0.002	0.001
$\hat{\pi}_{u\|5}$	0.000	0.016	0.013	0.060	*0.639*	0.272
$\hat{\pi}_{u\|drop}$	0.000	0.000	0.000	0.000	0.000	1.000
male	1	2	3	4	5	*drop*
$\hat{\pi}_{u\|1}$	0.691	*0.132*	0.021	0.053	0.103	0.000
$\hat{\pi}_{u\|2}$	0.000	0.862	0.038	0.000	0.070	0.030
$\hat{\pi}_{u\|3}$	0.000	0.000	0.806	0.000	0.183	0.011
$\hat{\pi}_{u\|4}$	0.180	0.000	0.000	0.524	*0.296*	0.000
$\hat{\pi}_{u\|5}$	0.000	0.000	0.022	0.000	0.768	0.210
$\hat{\pi}_{u\|drop}$	0.000	0.000	0.000	0.000	0.000	1.000

UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA

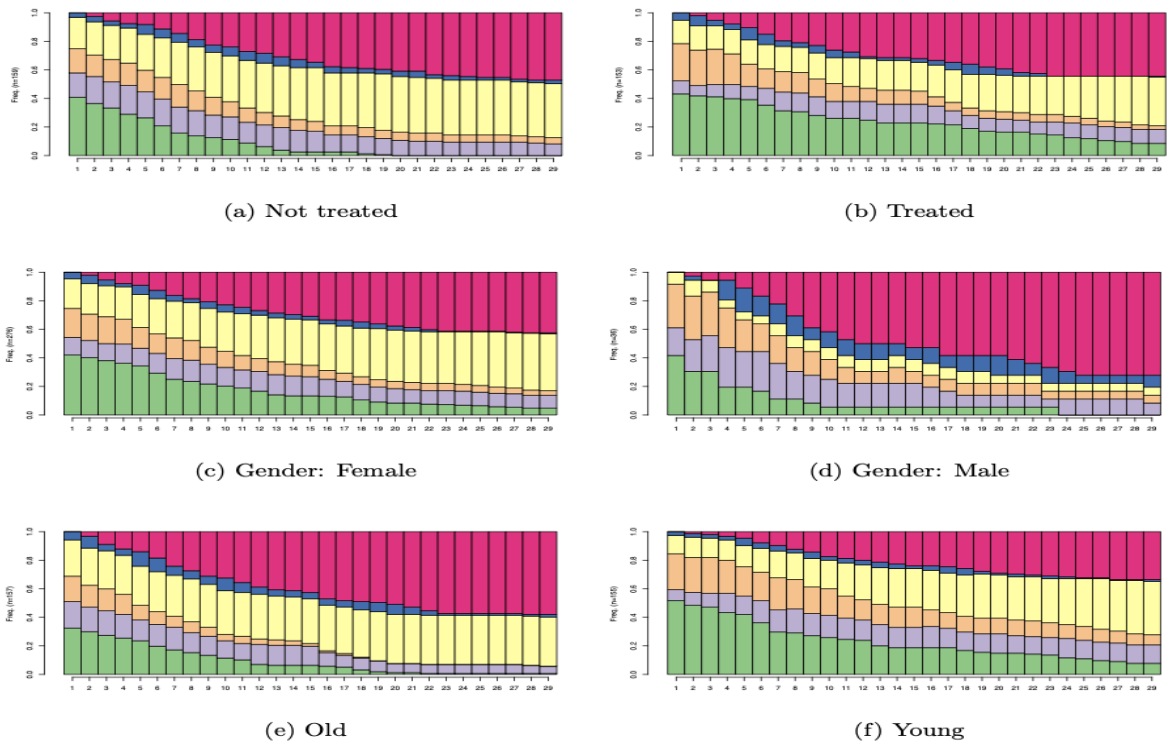
Results: Decoding

- *Dynamic clustering*: prediction the sequence of latent states to evaluate the time-varying patient risk of death
  - *Cluster 2* is more stable over time, *Cluster 1* rapidly decreases



Results: Decoding

- *Drug use* was not particularly effective in prolonging survival, *males* have a higher risk of dying compared to females, especially from the eighth visit, and *older patients* have a higher risk of dying compared to younger patients



(a) Not treated

(b) Treated

(c) Gender: Female

(d) Gender: Male

(e) Old

(f) Young

Conclusions

- An extended Hidden Markov model is proposed for *multivariate longitudinal data* to handle *(I)* completely, *(II)* partially missing responses and *(III)* dropout
- The model includes the effects of the *covariates* on the initial and transition probabilities to evaluate their influence towards the dropout state
- *Simulations* confirm that the model parameters are properly estimated; in the *application* using very sparse data we identified groups of patients differing for the severity of a rare disease and their transitions across groups and towards dropout
- The model may be also *extended* to include responses of a mixed nature and non only continuous

Main References

- Bartolucci, F. and Farcomeni, A. and Pennoni, F. (2013). Latent Markov Models for Longitudinal Data, Chapman and Hall/CRC press, Boca Raton
- Bartolucci, F., Pandolfi, S., and Pennoni, F. (2017). LMest: An R package for latent Markov models for longitudinal categorical data. Journal of Statistical Software, 81:1-38
- Delalleau, O., Courville, A., and Bengio, Y. (2018). Efficient EM training of Gaussian mixtures with missing data. *arXiv preprint arXiv:1209.0521*
- Dempster, A. P., Laird, N. M., Rubin, D.B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society B, 39: 138
- Dickson, E. R., Grambsch, P. M., Fleming, T. R., Fisher, L. D., Langworthy, A. (1989). Prognosis in primary biliary cirrhosis: model for decision making. Hepatology, 10:1-7
- Eirola, E., Lendasse, A., Vandewalle, V., and Biernacki, C. (2014). Mixture of Gaussians for distance estimation with missing data. Neurocomputing, 3:31-142
- Little R.J.A., Rubin D.B. (2002). Statistical Analysis with Missing Data, 2nd Edn. Wiley Series in Probability and Statistics, Wiley
- *Pandolfi, S., Bartolucci, F., Pennoni F. (2022). A hidden Markov model for continuous longitudinal data with missing responses and dropout.* Submitted and publicly available at the following repository https://arxiv.org/pdf/2106.15948.pdf

UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA