A dynamic perspective to evaluate multiple treatments through a causal latent Markov model

Fulvia Pennoni

Department of Statistics and Quantitative Methods

University of Milano-Bicocca

http://www.statistica.unimib.it/utenti/pennoni/

Email: fulvia.pennoni@unimib.it

joint work with F. Bartolucci and G. Vittadini

Outline

- *Fulvia Pennoni is associate professor of statistics at Department of Statistics and Quantitative Methods, University of Milano-Bicocca. Her research interest includes latent variable models for longitudinal data, graphical and multilevel models;*

- *Francesco Bartolucci is full professor of statistics at the Department of Economics, University of Perugia. His research interests are about latent variable models and mixture models.*

- *Giorgio Vittadini is full professor of methodological statistics at Department of Statistics and Quantitative Methods, University of Milano-Bicocca. His research interests are on statistical models with latent variables and structural equation models.*

Outline

- We propose a novel approach for estimating Average Causal Effects (ATEs) in the presence of multiple treatments;

- We propose a new Potential outcome model with a longitudinal Latent Markov (LM) structure of dependence for multivariate responses in the context of observational studies.

Potential outcome framework

- The causal effects are defined according to the potential outcomes: the set of all outcomes related to all the feasible treatments for every sample unit;
- The interest is to consider the Average Treatment Effect: the difference (over the population) between the difference over two potential outcomes corresponding to two treatments (or between treatment and control);
- We follow the proposal of Lanza *et al.* (2013): they use a mixed method between matching and weighting to estimate causal effects within a latent class model (Goodman, 1974 and Lazarsfeld and Henry, 1968);
- Matching and weighting are two statistical approaches to balance groups according to pre treatment covariates.

The notation

We consider this notation in the following:

- n: number of individuals;
- T: number of time points (occasions) defining the longitudinal structure of the data;
- $\boldsymbol{Y}_{it}$: vector of categorical response variables Y_{ijt} (at occasion t, variable j, individual i);
- $\boldsymbol{X}_i$: corresponding column vector of pre-treatment specific covariates related to each individual i.

The illustrative example

In the illustrative example we aim at studying the development of the Human Capital (HC), and the related university-to-work transition phenomenon, due to the different types of degree, therefore we consider

- n: number of observed graduates in a certain year;
- T: each available time at which each response can be observed after graduation;
- $\boldsymbol{Y}_{it}$: vector of categorical response variables where $r = 3$ and $c_j = 3$ related to:
  - *i*) *contract type* with categories: none, temporary, and permanent;
  - *ii*) *skill* with categories: none, low/medium, and high;
  - *iii*) *gross income* in Euros (€) (quarterly) with categories: none, $\leq$3750 €, and $>$3750 €.
- $\boldsymbol{X}_i$: corresponding column vector of pre-treatment covariates: *gender*, *district of birth*, *final grade at high school diploma*, *type of high school*.

Latent variable

- The aim is to consider the effect of the acquired type of degree on the development of a latent concept like the Human Capital (HC);
- It is a latent construct related to the skills, competencies, and attributes embodied in individuals that are relevant to the economic activities, with particular reference of the labor market;
- From an economic point of view we take into account the definition of Harpan and Draghici (2014):"*the generic knowledge and skill accumulated during experiences and education*";
- We conceive the HC as a latent variable affected by the investment in education, own ability and type of higher education;
- In the applicative example, we propose to measure it by considering the the above three response variables related to the economic climate just before the economic recession.

The causal LM model

- HC is a latent variable denoted by its "potential versions" as $H_{it}^{(z)}$, with $i = 1, \ldots, n$, $t = 1, \ldots, T$, $z = 1, \ldots, l$, where l is the number of the possible treatments;
- Only one of the possible outcomes $H_{it}^{(z)}$ is indeed selected and this selection may depend on the value of the variables themselves;
- $H_{it}^{(z)}$ corresponds to the latent state of individual i at occasion t if he/she had taken treatment z;
- We consider a marginal model for $H_{it}^{(z)}$ given the pre-treatment covariates in the sense of Robins et al. (2000).

The causal LM model

- The variables in the vector $\boldsymbol{H}_i^{(z)} = (H_{i1}^{(z)}, \ldots, H_{iT}^{(z)})$ are assumed as stochastic Markov processes of first-order;
- They have a discrete distribution with support $\{1, \ldots, k\}$ so that it is possible to identify (classify) individuals, with individuals in the same latent state sharing the same latent characteristics;
- We expect that there are sub-populations which cannot be directly observed having different expected behaviors with reference to the response variables.

The causal LM model

- We propose an innovative use of the propensity score method (Imbsen, 2000) which is extended to account for multiple treatments;

- First, we estimate a multinomial logit model for the probability of taking a certain type of treatment given the individual pre-treatment covariates $\boldsymbol{x}$ (see also, McCaffreys et al. 2013, among others):

$$\log \frac{p(Z_i = z|\boldsymbol{x}_i)}{p(Z_i = 1|\boldsymbol{x}_i)} = \eta_z + \boldsymbol{x}_i'\boldsymbol{\lambda}_z, \quad z = 2, \dots, l,$$

  where η_z and $\boldsymbol{\lambda}_z$ are regression parameters and so to obtain the individual weights

$$\hat{w}_i = n \frac{1/\hat{p}(z_i|\boldsymbol{x}_i)}{\sum_{m=1}^{n} 1/\hat{p}(z_m|\boldsymbol{x}_i)}, \quad i = 1, \dots, n.$$

- Second, we use the estimated values to weight the log-likelihood function which is used to estimate the model parameters within the Expectation-Maximization algorithm.

The causal LM model

- The probabilities at the first time period are modelled according to a baseline-category logit model:

$$\log \frac{p(H_{i1}^{(z)} = h)}{p(H_{i1}^{(z)} = 1)} = \alpha_h + \boldsymbol{d}(z)'\boldsymbol{\beta}_h, \quad h = 2, \dots, k,$$

  where the intercept is α_h, and $\boldsymbol{\beta}_h = (\beta_{h2}, \dots, \beta_{hl})'$ is a column vector of $l-1$ parameters, β_{hz} of $\boldsymbol{\beta}_h$ corresponds to the effect of the z-th treatment with respect to the first treatment;

- $\boldsymbol{\beta}_h$ corresponds to the ATE effect of the whole population of interest for the first time period.

The causal LM model

- The probabilities at the time periods different from the first one (*transition probabilities of the hidden chain*) are modelled according to a multinomial logit model:

$$\log \frac{p(H_{it}^{(z)} = h | H_{i,t-1}^{(z)} = \bar{h})}{p(H_{it}^{(z)} = 1 | H_{i,t-1}^{(z)} = \bar{h})} = \gamma_{\bar{h}h} + \boldsymbol{d}(z)'\boldsymbol{\delta}_h,$$

where $\gamma_{\bar{h}h}$ are the intercepts and $\boldsymbol{\delta}_h = (\delta_{h2}, \dots, \delta_{hl})'$, $\bar{h} = 1, \dots, k,\ h = 2, \dots, k,\ t = 2, \dots, T$ is the vectors of regression coefficients;

- Each coefficient δ_{hz} is again an ATE which is referred to the transition from level 1 to level h of the latent variable;

- The parameters of the measurement model are left unconstrained

$$\phi_{jy|h} = p(Y_{ijt} = y | H_{it}^{(z)} = h),$$

where $h = 1, \dots, k,\ j = 1, \dots, r,\ y = 0, \dots, c_j - 1$.

Causal LM model estimation procedure

- The weighted log-likelihood is considered

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \hat{w}_i \ell_i(\boldsymbol{\theta}), \quad \ell_i(\boldsymbol{\theta}) = \log p(\boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}_{iT} | z_i),$$

  where $\boldsymbol{\theta}$ is the vector of all causal LM model parameters and the $p(\boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}_{iT} | z_i)$ the *manifest probability*;

- The *complete data log-likelihood* $\ell^*(\boldsymbol{\theta})$ is maximized according to the following two steps of the EM algorithm:
  - **E-step**: compute the expected value of the frequencies and indicator variables the complete log-likelihood equation, given the observed data and the current value of the parameters, so as to obtain the expected value of $\ell^*(\boldsymbol{\theta})$;
  - **M-step**: update $\boldsymbol{\theta}$ by maximizing the expected value of $\ell^*(\boldsymbol{\theta})$ obtained at the E-step.

Details on the estimation procedure

- The *complete data*, which correspond to the value of every latent variable, further to the observed covariates and response variables:

$$\ell^*(\boldsymbol{\theta}) = \sum_{h=1}^{k}\sum_{j=1}^{r}\sum_{t=1}^{T}\sum_{y=0}^{c_j-1} a_{hjty}\log\phi_{jy|h} + \sum_{h=1}^{k}\sum_{i=1}^{n}\hat{w}_i b_{hi1}\log p(H_{i1}=h|z_i)$$
$$+\sum_{\bar{h}=1}^{k}\sum_{h=1}^{k}\sum_{i=1}^{n}\sum_{t=2}^{T}\hat{w}_i b_{\bar{h}hit}\log p(H_{it}=h|H_{i,t-1}=\bar{h},z_i)$$

where: a_{hjty} is the frequency of subjects responding by y to the j-th response variable and belonging to latent state h at occasion t; b_{hit} is an indicator variable equal to 1 if subject i belongs to latent class h at occasion t, with $p(H_{i1}=h|z_i)$; and $b_{\bar{h}hit}=b_{\bar{h}i,t-1}b_{hit}$ is an indicator variable equal to 1 if the same subject moves from state $\bar{h}$ to state h at occasion t and w_i are the weights.

Asymptotic properties of the proposed estimator

- We make a proof of the consistency for the estimator:

- We consider a data generating model of the following type for $i = 1, \ldots, n$:
  1. the vector of covariates $\boldsymbol{x}_i$ is drawn from an unknown distribution $f(\boldsymbol{x})$;
  2. the potential latent outcomes $H_{it}^{(z)}$ are drawn given $\boldsymbol{x}_i$ from a Markov chain for $t = 1, \ldots, T$ and $z = 1, \ldots, l$;
  3. given $\boldsymbol{x}_i$, the treatment indicator z_i is generated from a multinomial logit model based on the probabilities $p_z(\boldsymbol{x}_i)$, $z = 1, \ldots, l$;
  4. the latent variables are generated as $H_{it}^{(z_i)}$ for $t = 1, \ldots, T$;
  5. given the generated value of $H_{it}^{(z_i)}$, the outcomes Y_{ijt} are generated, for $j = 1, \ldots, r$ and $t = 1, \ldots, T$, according to the LM model.

- Proposition: As $n \to \infty$, the estimator $\hat{\boldsymbol{\theta}}$ under the data generating model described above and the estimator $\tilde{\boldsymbol{\theta}}$ under the randomized sampling scheme converge in probability to the same point $\boldsymbol{\theta}_0$ of the parameter space.

Finite sample properties of the proposed estimator

- We assess some finite sample properties through simulations comparing the estimates obtained with the proposed estimator of the LM model with and without using weights;
- We show that bias of the proposed estimator is negligible; while the bias is large for the estimator not corrected for confounding and it does not decrease when the sample size and/or the time occasions are higher;
- The standard deviation of the proposed estimator decreases as n and T increase; in particular, it decreases with a rate close to $\sqrt{n}$;
- The table below shows the results of simulations according to the scenario of the following slide

				$k=2$		$k=3$			
				β_{22}	δ_{22}	β_{22}	β_{23}	δ_{22}	δ_{23}
$n=2,000$	$T=4$	Randomized	Mean	1.606	0.796	1.288	2.586	0.631	1.274
			SD	0.167	0.116	0.196	0.175	0.123	0.114
		Proposed	Mean	1.606	0.800	1.281	2.579	0.645	1.290
			Bias	0.000	0.004	−0.006	−0.007	0.014	0.016
			SD	0.219	0.151	0.240	0.230	0.178	0.160
		Naive	Mean	2.522	1.253	1.769	3.546	0.879	1.768
			Bias	0.917	0.457	0.481	0.959	0.248	0.493
			SD	0.211	0.135	0.196	0.205	0.146	0.134

Simulation design

- The estimates are compared with those obtained from the naive estimation method of the LM model without using PS weighting and those obtained in the case of randomized treatment;
- We assume the existence of two covariates affecting both the treatment and the value of the potential outcomes $H_{it}^{(z)}$ for all possible treatments z: one is continuous and is binary with two possible values;
- we drew 1,000 samples of size $n = 1000, 2000$ for a number of time occasions $T = 4, 8$, and we consider $k = 2, 3$ latent states; so that there are 16 scenarios overall;
- For $k = 2$ latent states and $l = 2$ treatments, and we fix $\alpha_2^* = -1$, $\beta_{22}^* = 2$, $\tau_2 = 1$, $\gamma_{12}^* = -1$, and $\gamma_{22}^* = 1$, with $\delta_{22}^* = \beta_{22}^*/2$ and $\psi_2 = \tau_2/2$.

Illustrative application: The administrative archives

- The data are from integrated *administrative archives*:[1]
  - the archive of the federal observatory of the labour market in Lombardy concerning the compulsory communications given by the employer from 2000 to nowadays, regarding activation, termination of the employment relationship;
  - the archive of the graduates of four universities in Milan, concerning the academic performance for all students gaining a university degree between 2003 and 2008;
  - the archive of the Italian office of revenues relative to the annual gross earned income of all residents declaring income in Lombardy (available years: 2007- 2008 for residents in Milan);
  - the archive of the Milan's City Hall recording annually the personal information about citizens.

[1] held by the Interuniversity Research Centre, visit `http://www.crisp-org.it/`

Illustrative application: data description

- We dispose of all the graduates in 2007 from four different universities with five years of university education: (pre-reform and post-reform)
- We exclude the faculties such as Law and Health often characterized by institutionalized stages for advancement in the associated professional careers;
- due to the fact that the archives covers different subpopulations and temporal periods we analyze a dataset of 1,624 individuals resident in the area surrounding Milan;
- The period of observation after graduation is one year
- The percentage of graduates for each type of treatment are the following

Table 1: *Percentage of graduates for each type of treatment*

Treatment		%
Degree type	Technical	22.78
	Architecture	8.93
	Economic	13.85
	Humanities	41.26
	Scientific	13.18

Illustrative application: response variables

- Unemployment rate is about 62% for the first quarter;
- The graduates who are employed in the first quarter, about 12% start working with a permanent contract and, within one year this value is about 19%;
- About 24% start working mainly with high skills and, within one year, graduates having high skill jobs are about 36%.

Table 3: *Percentage distribution of every response variable*

	Quarter (t)			
Contract type	1st	2nd	3rd	4th
None	61.58	53.51	50.68	47.23
Temporary	26.72	31.53	31.83	33.44
Permanent	11.70	14.96	17.49	19.33

	Quarter (t)			
Skill	1st	2nd	3rd	4th
None	61.58	53.51	50.68	47.23
Low/medium	14.72	15.21	16.07	17.00
High	23.72	31.28	33.25	35.78

	Quarter (t)			
Gross income	1st	2nd	3rd	4th
None	59.73	50.62	47.35	44.95
$\leqslant$ 3,750 (€)	31.28	29.74	27.34	25.31
$>$ 3,750 (€)	8.99	19.64	25.31	29.74

Illustrative application: Pre-treatment covariates

- A covariate need to be selected when it shows a strong dependence with the treatment when the hypothesis of equal means of this covariates for the different degrees is rejected;

- We used ANOVA (ANalysis Of VAriance) model the quantitative covariates and χ^2 test of independence for qualitative covariates.

- For example the proportions related to gender are showed in the following

Table 2: *Means (or proportions) for each available pretreatment covariate (†coded as: 1 for single, 2 for single with children, 3 for couple with no children, 4 for couple with children, 5 for other)*

		University degree				
Pretreatment covariate		Techn.	Arch.	Econ.	Human.	Scien.
Gender	Male	0.802	0.483	0.547	0.202	0.577
	Female	0.198	0.517	0.453	0.798	0.423

Illustrative application: selection of the pre-treatment covariates

- We assess the balance among Architecture, Economic, Humanities, Scientific degrees by considering the pre-treatment covariates through a multinomial logit model;
- Percentage of females among graduates in a technical subject is around 20% and it is around 80% among graduates in a humanities subject. By employing weights these percentages are 45% and 49%

Table 2: *Means (or proportions) for each available pretreatment covariate (†coded as: 1 for single, 2 for single with children, 3 for couple with no children, 4 for couple with children, 5 for other)*

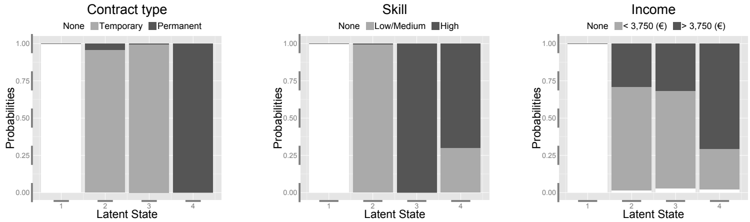
		University degree				
Pretreatment covariate		Techn.	Arch.	Econ.	Human.	Scien.
Gender	Male	0.802	0.483	0.547	0.202	0.577
	Female	0.198	0.517	0.453	0.798	0.423

TABLE 6.
Weighted Means (or Proportions) for Each Pretreatment Covariate Included in the Multinomial Logit Model Used to Compute Individual Weights

		University Degree				
Covariate		Technical	Architecture	Economic	Humanities	Scientific
Gender	Male	0.552	0.525	0.504	0.514	0.536
	Female	0.448	0.475	0.496	0.486	0.464
District of birth	Milan	0.784	0.726	0.762	0.750	0.791
	Lombardy	0.055	0.060	0.060	0.079	0.059
	Italy	0.146	0.185	0.145	0.147	0.127
	Others	0.015	0.029	0.032	0.024	0.023

Illustrative application: results

- Estimated conditional probabilities of labor condition ($\phi_{jy|h}$) when $k = 4$;



- *1st class*: lowest HC level; *2nd class*: intermediate HC level with high probability of a temporary job, requiring a low/medium skill level, and intermediate income level; *3rd class*: similar to the 2nd class with the exception of a higher skill level and a slightly higher income; *4th class*: high HC level with permanent contract, intermediate-high skills, and high income.

Illustrative application: estimates on the initial probabilities

- Every parameter β_{hz} is an ATE of degree (treatment) z, with respect to a technical degree ($z = 1$), in terms of initial probabilities expressed on the logit scale;
- Standard errors for the parameter estimates are obtained by a non-parametric bootstrap method;
- At the beginning of the period of observation, there is a statistically significant difference in terms of effect on HC of technical degrees with respect to architecture and humanities degrees.

TABLE 8.
Estimates of the Logit Regression Parameters for the Initial Probabilities of the Latent Process Under the Selected LM Causal Model with k = *4 Latent States*

	Latent State (h)		
Treatment	2	3	4
Technical ($\hat{\alpha}_h$)	−1.239**	−0.584**	−1.005**
Architecture versus technical ($\hat{\beta}_{h2}$)	−1.177**	−1.267**	−1.369**
Economic versus technical ($\hat{\beta}_{h3}$)	0.405	−0.232	−0.118
Humanities versus technical ($\hat{\beta}_{h4}$)	−0.522	−0.776**	−1.623**
Scientific versus technical ($\hat{\beta}_{h5}$)	−0.434	−0.981†	−0.903

Illustrative application: results

- According with the estimates of the above parameters it is possible to dispose of the estimated initial probabilities;
- At the beginning of the period of observation, there is a statistically significant difference in terms of effect on HC of technical degrees with respect to architecture and humanities degrees and in favor of the first ones;
- Recall that the first and the last latent classes, corresponding to the subpopulations of individuals with the lowest and the highest HC level, respectively

TABLE 9.
Estimated Initial Probabilities for Each Type of Treatment Under the Selected Causal LM Model with k = 4 *Latent States*

	Latent State (h)			
Treatment	1	2	3	4
Technical	0.452	0.131	0.252	0.165
Architecture	0.747	0.067	0.116	0.070
Economic	0.454	0.197	0.201	0.148
Humanities	0.666	0.115	0.171	0.048
Scientific	0.647	0.121	0.136	0.096

Illustrative application: estimates on the transition probabilities

- Every parameter $\gamma_{\bar{h}h}$ is an ATE of the degree z, with respect to a technical degree ($z = 1$), on transition probabilities;
- The causal effects of the degree type on the evolution of the HC level through transition probabilities are different;
- There are significant differences between technical degrees and all the other types of degree during the period of observation

TABLE 10.
Estimates of the Logit Regression Parameters for the Transition Probabilities of the Latent Process Under the Selected Causal LM Model with k = *4 Latent States*

Treatment	Latent state (h) 2	3	4
Technical $\bar{h} = 1(\hat{\gamma}_{1h})$	−3.020**	−1.620**	−1.894**
Technical $\bar{h} = 2(\hat{\gamma}_{2h})$	1.880**	−0.226	0.405
Technical $\bar{h} = 3(\hat{\gamma}_{3h})$	−2.497†	2.185**	0.171
Technical $\bar{h} = 4(\hat{\gamma}_{4h})$	−14.382**	−0.039	5.934**

Illustrative application: results

- Estimated transition probabilities according with the degree type:

TABLE 11.
Estimates of the Transition Probabilities Under the Selected Causal LM Model with k = 4 Latent States Referred to Each Treatment

		Latent state (h)			
Degree	$\bar{h}$	1	2	3	4
Technical	1	0.716	0.035	0.142	0.107
	2	0.102	0.665	0.081	0.152
	3	0.009	0.007	0.797	0.106
	4	0.003	0.000	0.002	0.995
Architecture	1	0.886	0.030	0.074	0.010
	2	0.167	0.758	0.057	0.018
	3	0.204	0.012	0.767	0.017
	4	0.036	0.000	0.014	0.950
Economic	1	0.835	0.044	0.086	0.035
	2	0.112	0.795	0.046	0.047
	3	0.165	0.015	0.765	0.055
	4	0.009	0.000	0.005	0.986
Humanities	1	0.846	0.035	0.099	0.020
	2	0.138	0.764	0.065	0.033
	3	0.152	0.011	0.808	0.029
	4	0.016	0.000	0.009	0.974
Scientific	1	0.858	0.034	0.079	0.029
	2	0.139	0.764	0.051	0.046
	3	0.183	0.013	0.756	0.048
	4	0.012	0.000	0.005	0.983

Illustrative application: results

- By considering the estimated transition probability matrices of Table 11 we conclude that
  - for technical degrees there is the lowest probability of remaining in the first latent class (0.72) and then the highest probability of moving away from this class;
  - The second lowest probability of remaining in the first class is for economic degrees (0.84);
  - The third lowest probability of remaining in the first class is for humanities degrees (0.85),
  - For scientific degrees we have the second lowest probability of persistence in the first latent class (0.86) and the third lowest probability of remaining in the last class (0.97),
  - For architecture we have the highest probability for the first class (0.89) and the lowest for the last class (0.95).

Conclusions

- We propose a new estimator under the potential outcome framework with longitudinal observational data;
- We provide a proposition of for the consistency of the novel estimator by showing that it converges to the same point of the parameter space at which the standard estimator converges under a perfectly randomized sampling scheme;
- This new statistical model has a potential use in a wide range of observational studies;
- The proposal is feasible also for more time occasions by which it can be possible to study the long-term effects of the treatment;
- The illustrative example dealing with the effect of the degree may be also considered to implement another way to rank university degrees so that they can be also evaluated in terms of their impact on HC levels.

Main References

- Bartolucci F, Pennoni F, Vittadini G (2016). Causal Latent Markov Model for the Comparison of Multiple Treatments in Observational Longitudinal Studies. *Journal of Educational and Behavioral Statistics*, **41**, 146–179. http://journals.sagepub.com/doi/abs/10.3102/1076998615622234

- Bartolucci, F., and Pennoni, F. (2011). Impact evaluation of job training programs by a latent variable model. In S. Ingrassia, R. Rocci, and M. Vichi (Eds.), *New Perspectives in Statistical Modeling and Data Analysis* (pp. 65–73). Berlin, Germany: Springer- Verlag.

- Bartolucci, F., Pennoni, F., and Vittadini, G. (2011). Assessment of school performance through a multilevel latent Markov Rasch model. *Journal of Educational and Behavioural Statistics*, **36**, 491–522.

- Bartolucci, F., Farcomeni, A., and Pennoni, F. (2014). Latent Markov models: A review of a general framework for the analysis of longitudinal data with covariates (with discussion). *Test*, **23**, 433–486.

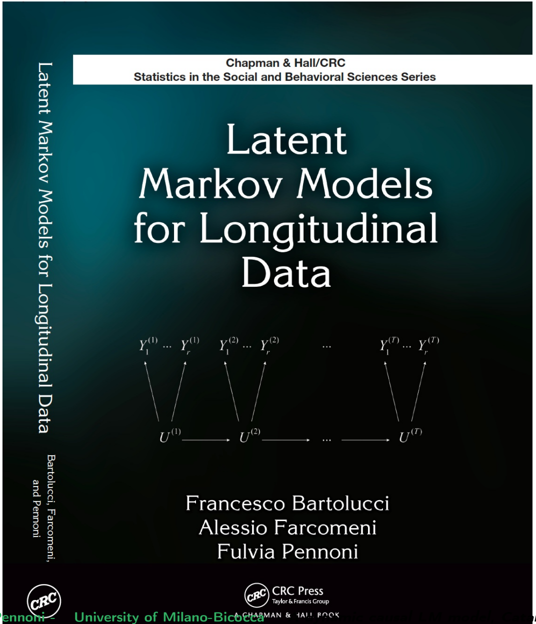
- Bartolucci, F., Farcomeni, A., and Pennoni, F. (2013). *Latent Markov models for longitudinal data*. Boca Raton, FL: Chapman & Hall/CRC.

Main References

- Guo, S., and Fraser, M. W. (2010). *Propensity score analysis: Statistical methods and applications*. Thousand Oaks, CA: Sage.
- Harpan, I., and Draghici, A. (2014). Debate on the multilevel model of the human capital measurement. *Procedia–Social and Behavioral Sciences*, **124**, 170–177.
- Hirano, K., Imbens, G. W., and Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, **71**, 1161–1189.
- Keribin, C. (2000). Consistent estimation of the order of mixture models. *Sankhya: The Indian Journal of Statistics, Series A*, **62**, 49–66.
- McCaffrey, D. F., Griffin, B. A., Almirall, D., Slaughter, M. E., Ramchand, R., and Burgette, L. F. (2013). A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Statistics in Medicine*, **32**, 338–3414.
- Robins, J., Hernan, M., and Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, **11**, 550–560.
- Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical* , **100**, 322–331.

- For more details on the main basic and advanced theory of the latent Markov models consider the book and the summer school of the following slides

Latent Markov models for longitudinal data



Italian Statistical Society and University of Milano-Bicocca Summer School

Course of the SIS' school

Latent Markov models with applications
Milano, 16 - 20 September 2017
Department of Statistics and Quantitative Methods
University of Milano-Bicocca

Local Organizing Committee

Fulvia Pennoni
(University of Milano-Bicocca)

Scientific Committee

Francesco Bartolucci
(University of Perugia)

Alessio Farcomeni
(Sapienza University of Rome)

Silvia Pandolfi
(University of Perugia)

Fulvia Pennoni
(University of Milano-Bicocca)

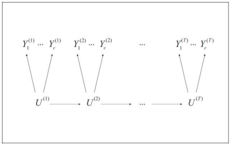
For additional details and student housing
http://www.summerschoolbicocca.com/

Administrative Secretary
SIS Secretariat
Salita de' Crescenzi 26 – 00186 Roma
tel. +39 06 6869845 – Fax +39 06 68806742
e-mail: sis@sis-statistica.it

Scuola
della
Società Italiana di Statistica

Latent Markov models with applications

s



Department of Statistics and Quantitative Methods
University of Milano-Bicocca
16 - 20 September 2017