Drifting semi-Markov models: Estimation and Simulation through the `dsmmR` package

V. Barbu, I. Mavrogiannis and N. Vergne

Laboratoire de Mathématiques Raphaël Salem, Université de Rouen Normandie

PMSMA, IMAG Montpellier, 5-7 juin 2023

Plan

Plan

Linear drifting Markov model

Definition (linear drifting Markov chain of order 1 and of length n)

A sequence $X_0, X_1, \ldots, X_n$ with state space $E = \{1, 2, \ldots, s\}$ is called a *linear drifting Markov chain (of order 1)* of length n between the Markov transition matrices $\mathbf{\Pi}_0$ and $\mathbf{\Pi}_1$, if the distribution of X_t, $t = 1, \ldots, n$, is defined by

$$\mathbb{P}(X_t = v \mid X_{t-1} = u) = \Pi_{\frac{t}{n}}(u, v), \ u, v \in E,$$

where $\Pi_{\frac{t}{n}}(u, v) = \left(1 - \frac{t}{n}\right) \Pi_0(u, v) + \frac{t}{n} \Pi_1(u, v), \ u, v \in E.$

Estimation method : Least squares estimators. We minimize

$$\sum_{t=1}^{n} \sum_{u \in E} \sum_{v \in E} \mathbb{1}_{\{X_{t-1}=u\}} \left(\Pi_{\frac{t}{n}}(u, v) - \mathbb{1}_{\{X_t = v\}} \right)^2.$$

Polynomial drifting Markov model

Definition (polynomial drifting Markov chain of order k and of length n)

A sequence $X_0, X_1, \ldots, X_n$ with state space $E = \{1, 2, \ldots, s\}$ is said to be a *polynomial drifting Markov chain of order k* and of length n if the distribution of X_t, $t = 1, \ldots, n$, is defined for $u_1, \ldots, u_k, v \in E$ by

$$\mathbb{P}(X_t = v \mid X_{t-1} = u_k, X_{t-2} = u_{k-1}, \ldots) = \mathbf{\Pi}_{\frac{t}{n}}(u_1, \ldots, u_k, v)$$
$$\text{where } \mathbf{\Pi}_{\frac{t}{n}}(u_1, \ldots, u_k, v) = \sum_{i=0}^{d} A_i(t)\mathbf{\Pi}_{\frac{i}{d}}(u_1, \ldots, u_k, v)$$

with A_i polynomials of degree d such as, for any $i, j \in \{0, 1, \ldots, d\}$, $A_i\left(\frac{nj}{d}\right) = \mathbb{1}_{\{i=j\}}$. Note : for $t = ni/d$, we have $\mathbf{\Pi}_{\frac{t}{n}} = \mathbf{\Pi}_{\frac{i}{d}}$.

Remark (A_i are Lagrange polynomials : chosen to have stochastic $\mathbf{\Pi}_{\frac{t}{n}}$.)

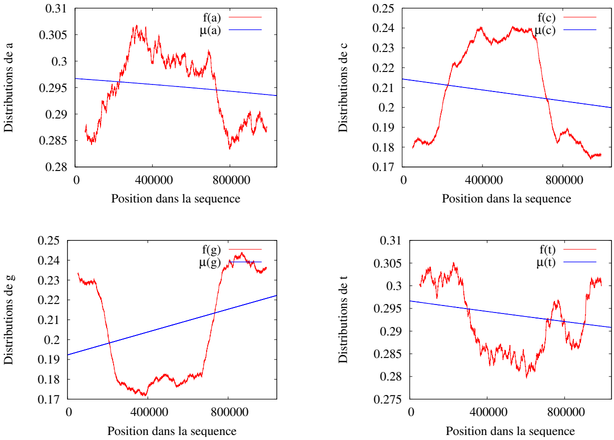
Indeed $\sum_{v \in E} \mathbf{\Pi}_{\frac{t}{n}}(u_1, \ldots, u_k, v) = \sum_{i=0}^{d} A_i(t) = A(t)$ *where* A *is a polynomial of degree* d *equal to one in* $(d+1)$ *points ; then* A *is a constant polynomial equal to one.*

Proposition

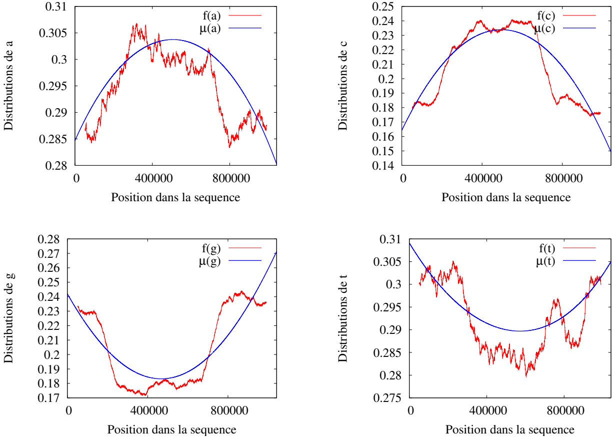
Let $\mathbb{1}_u := \mathbb{1}_{\{X_{t-k}\dots X_{t-1}=u_1\dots u_k\}}$ *and* $\mathbb{1}_{uv} := \mathbb{1}_{\{X_{t-k}\dots X_{t-1}=u_1\dots u_k, X_t=v\}}$. *For* $(X_0, X_1, \dots, X_n)$ *a sample path of a polynomial drifting Markov chain of order* k *and degree* d, *for any states* $u_1, \dots, u_k, v \in E$, *the estimators of* $\Pi_{\frac{i}{d}}(u_1, \dots u_k, v)$ *are given by solving the following linear system*

$$\begin{pmatrix} \sum_{t=k}^{n} \mathbb{1}_u A_0(t) A_0(t) & \cdots & \sum_{t=k}^{n} \mathbb{1}_u A_0(t) A_d(t) \\ \vdots & & \vdots \\ \sum_{t=k}^{n} \mathbb{1}_u A_d(t) A_0(t) & \cdots & \sum_{t=k}^{n} \mathbb{1}_u A_d(t) A_d(t) \end{pmatrix} \begin{pmatrix} \widehat{\Pi_{0;n}} \\ \vdots \\ \widehat{\Pi_{1;n}} \end{pmatrix} = \begin{pmatrix} \sum_{t=k}^{n} A_0(t) \mathbb{1}_{uv} \\ \vdots \\ \sum_{t=k}^{n} A_d(t) \mathbb{1}_{uv} \end{pmatrix}.$$

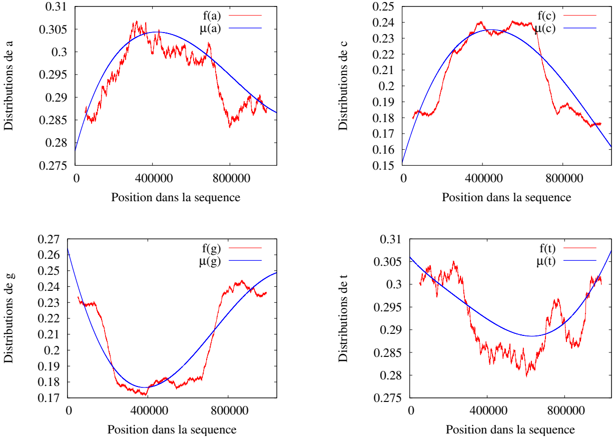
Fréquences / Lois stationnaires (degré 1) sur *Chlamydia trachomatis*



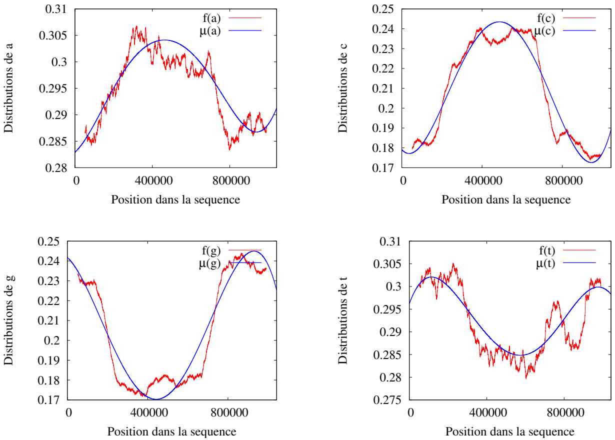
Fréquences / Lois stationnaires (degré 2)



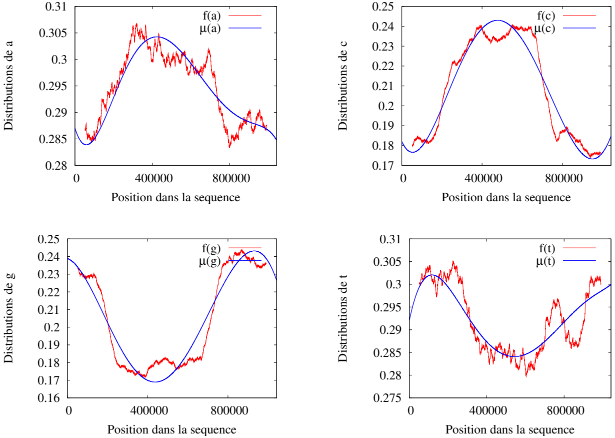
Fréquences / Lois stationnaires (degré 3)



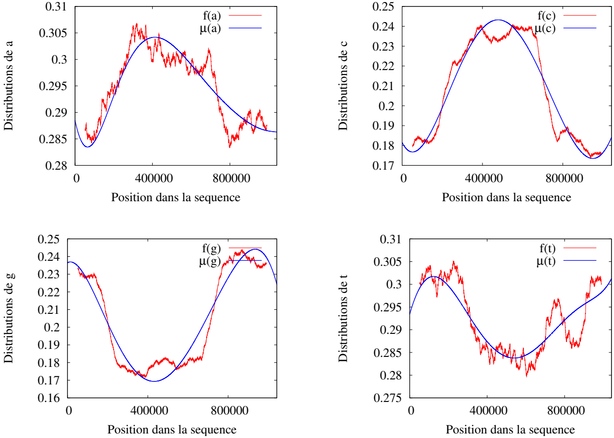
Fréquences / Lois stationnaires (degré 4)



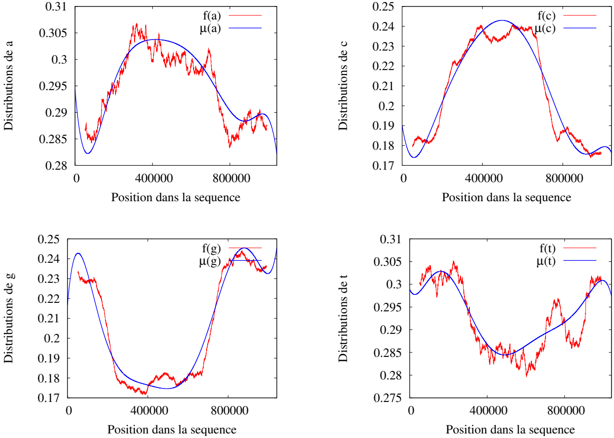
Fréquences / Lois stationnaires (degré 5)



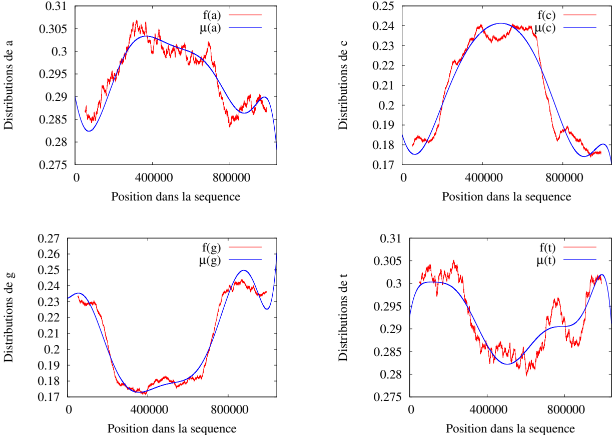
Fréquences / Lois stationnaires (degré 6)



Fréquences / Lois stationnaires (degré 7)



Fréquences / Lois stationnaires (degré 8)



Software, R Package, Web Interface and Applications

[1] **N. Vergne**, 2008. Drifting Markov models : Polynomial Drift. Applications to DNA sequences. *Statistical Applications in Genetics and Molecular Biology*, Vol 7, Iss 1, Article 6. Available at : https://doi.org/10.2202/1544-6115.1326.

[2] **V. S. Barbu, N. Vergne**, 2019. Reliability and Survival Analysis for Drifting Markov Models : Modeling and Estimation. *Methodology and Computing in Applied Probability*, 1-23. Available at https://link.springer.com/article/10.1007/s11009-018-9682-8

[3] **V. S. Barbu, G. Brelurut, A. Gilles, A. Lefebvre, C. Lothodé, V. Mataigne, A. Seiller** [1] **and N. Vergne**, 2021. drimmR : an R package for drifting Markov model estimation and reliability. Available at https://cran.r-project.org/web/packages/drimmR/index.html.

[4] **J. Komar, A. Lefebvre, H. Mayeur and N. Vergne**, 2021. WebDRIMM : a web interface for drifting Markov model estimation and reliability. Available at http://bioinfo.univ-rouen.fr/WebDRIMM/download.php.

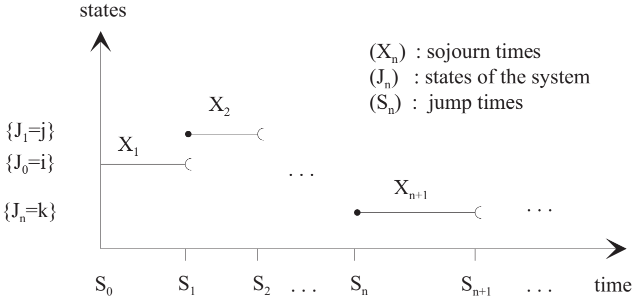
[5] **J. Komar, L. Seifert, N. Vergne and K.M. Newell**, 2023. Narrowing the coordination solution space during motor learning standardizes individual patterns of search strategy but diversifies learning rates. *Scientific Reports*, 13. Available at : https://www.nature.com/articles/s41598-023-29238-z

1. Thanks to FEDER DAISI

Plan

Different chains

- $Z = (Z_k)_{k \in \mathbb{N}}$, chain with state space $E = \{1, 2, \ldots, s\}$
- $S = (S_n)_{n \in \mathbb{N}}$, jump times
  $J = (J_n)_{n \in \mathbb{N}}$, visited states
  $X = (X_n)_{n \in \mathbb{N}}$, sojourn times of Z



semi-Markov kernel

the initial distribution $\alpha(i) := \mathbb{P}(J_0 = i)$
the homogeneous SM kernel $\mathbf{q} = \big(q(i,j;k)\big)_{i,j \in E, k \in \mathbb{N}}$

$$q(i,j;k) := \begin{cases} \mathbb{P}(J_{n+1} = j, X_{n+1} = k \mid J_n = i), & k \in \mathbb{N}^* \\ 0, & k = 0 \end{cases}$$

the conditional sojourn time distributions $\mathbf{f} = \big(f(i,j;k)\big)_{i,j \in E, k \in \mathbb{N}}$

$$f(i,j;k) := \mathbb{P}(X_{n+1} = k \mid J_n = i, J_{n+1} = j),\ f(i,j;0) := 0$$

the transition matrix of the MC $(J_n)_{n \in \mathbb{N}}$, $\mathbf{p} = (p(i,j))_{i,j \in E}$

$$p(i,j) := \mathbb{P}(J_{n+1} = j \mid J_n = i),\ p(i,i) := 0$$

Note that $q(i,j;k) = p(i,j)\ f(i,j;k)$

R Packages

[6] **V. S. Barbu, C. Bérard, D. Cellier, M. Sautreuil, N. Vergne**, 2017. `SMM` : a R package for Simulation and Estimation of Multi-State Discrete-Time Semi-Markov and Markov Models. Available at `https://cran.r-project.org/web/packages/SMM`.

[7] **V. S. Barbu, C. Bérard, D. Cellier, M. Sautreuil, N. Vergne**, 2018. SMM : An R package for estimation and simulation of discrete-time semi-Markov models. *The R journal*. Available at `https://rjournal.github.io/archive/2018/RJ-2018-050/index.html`.

[8] **V. S. Barbu, F. Lecocq** [2]**, C. Lothodé and N. Vergne**, 2021. `smmR` : a R package for Simulation, Estimation and Reliability of Semi-Markov Models. Available at `https://cran.r-project.org/web/packages/smmR/index.html`.

2. Thanks to RIN Asterics and FEDER DAISI

Plan

Drifting semi-Markov models – 1

Definition (linear drifting semi-Markov chain of length n – Model 1)

A sequence $Z_0, Z_1, \ldots, Z_{N(n)}$ with state space $E = \{1, 2, \ldots, s\}$ is called a *linear drifting Markov chain of length n of Model 1* between the semi-Markov kernels $\boldsymbol{q}_0$ and $\boldsymbol{q}_1$, if for $t = 0, \ldots, n$, we have

$$\mathbb{P}(J_{t+1} = v, S_{t+1} - S_t = \ell | J_t = u) = q_{\frac{t}{n}}(u, v, \ell),$$

where $q_{\frac{t}{n}}(u, v, \ell) = \left(1 - \frac{t}{n}\right) q_0(u, v, \ell) + \frac{t}{n} q_1(u, v, \ell), \; u, v \in E, \ell \in \mathbb{N}.$

Drifting semi-Markov models – 2

Definition (linear drifting semi-Markov chain of length n – Model 2)

A sequence $Z_0, Z_1, \ldots, Z_{N(n)}$ with state space $E = \{1, 2, \ldots, s\}$ is called a *linear drifting Markov chain of length n of Model 2* between the semi-Markov kernels $\boldsymbol{q}_0$ and $\boldsymbol{q}_1$, if for $t = 0, \ldots, n$, we have

$$\mathbb{P}(J_{t+1} = v, S_{t+1} - S_t = \ell | J_t = u) = q_{\frac{t}{n}}(u, v, \ell), \text{ where}$$

$$q_{\frac{t}{n}}(u, v; \ell) = \left(1 - \frac{t}{n}\right) f(u, v; \ell) p_0(u, v) + \frac{t}{n} f(u, v; \ell) p_1(u, v), \ u, v \in E, \ell \in \mathbb{N},$$

with $\boldsymbol{p}_0$ and $\boldsymbol{p}_1$ Markov kernels, $f(u, v; \ell)$ the conditional distribution of the sojourn in state u before jumping to v equal to ℓ.

Drifting semi-Markov models – 3

Definition (linear drifting semi-Markov chain of length n – Model 3)

A sequence $Z_0, Z_1, \ldots, Z_{N(n)}$ with state space $E = \{1, 2, \ldots, s\}$ is called a *linear drifting Markov chain of length n of Model 3* between the semi-Markov kernels $\boldsymbol{q}_0$ and $\boldsymbol{q}_1$, if for $t = 0, \ldots, n$, we have

$$\mathbb{P}(J_{t+1} = v, S_{t+1} - S_t = \ell | J_t = u) = q_{\frac{t}{n}}(u, v, \ell), \text{ where}$$

$$q_{\frac{t}{n}}(u, v; \ell) = \left(1 - \frac{t}{n}\right) f_0(u, v, \ell) p(u, v) + \frac{t}{n} f_1(u, v, \ell) p(u, v), \; u, v \in E, \ell \in \mathbb{N},$$

with $\boldsymbol{p}$ a Markov kernel, $f_0(u, v; \ell)$ and $f_1(u, v; \ell)$ conditional distributions of the sojourn in state u before jumping to v equal to ℓ.

Polynomial DSMM

Model 1 : p **and** f **are drifting**. The DSM kernel is given by :

$$q^{(1)}_{\frac{t}{n}} = \sum_{i=0}^{d} A_i(t) q^{(1)}_{\frac{i}{d}}(u, v, l) = \sum_{i=0}^{d} A_i(t) p_{\frac{i}{d}}(u, v) f_{\frac{i}{d}}(u, v, l)$$

Model 2 : **Only** p **is drifting** (f is not drifting). The DSM kernel is given by :

$$q^{(2)}_{\frac{t}{n}}(u, v, l) = \sum_{i=0}^{d} A_i(t) q^{(2)}_{\frac{i}{d}}(u, v, l) = \sum_{i=0}^{d} A_i(t) p_{\frac{i}{d}}(u, v) f(u, v, l)$$

Model 3 : **Only** f **is drifting** (p is not drifting). The DSM kernel is given by :

$$q^{(3)}_{\frac{i}{d}}(u, v, l) = \sum_{i=0}^{d} A_i(t) q^{(3)}_{\frac{i}{d}}(u, v, l) = \sum_{i=0}^{d} A_i(t) p(u, v) f_{\frac{i}{d}}(u, v, l)$$

Plan

Non-parametric Estimation for model 1

Model 1

We can estimate the DSM kernel by Least Square Estimation (LSE),

$$\widehat{q}^{\,(1)}_{\frac{t}{n}}(u,v,l) = \sum_{i=0}^{d} A_i(t)\widehat{q}^{\,(1)}_{\frac{i}{d}}(u,v,l)$$

$\forall u, v \in E, l \in \{1, \ldots, k_{max}\}$, where k_{max} is the maximum realized sojourn time in the sequence, we obtain the SM kernels $\widehat{q}^{\,(1)}_{\frac{i}{d}}(u,v,l), i = 0, \ldots, d$.

Estimators of $\widehat{p}_{\frac{i}{d}}(u,v)$ and $\widehat{f}_{\frac{i}{d}}(u,v,l)$ are obtained as in SMM:

$$\widehat{p}_{\frac{i}{d}}(u,v) = \sum_{l=1}^{k_{max}} \widehat{q}^{\,(1)}_{\frac{i}{d}}(u,v,l),$$

$$\widehat{f}_{\frac{i}{d}}(u,v,l) = \frac{\widehat{q}^{\,(1)}_{\frac{i}{d}}(u,v,l)}{\sum_{l=1}^{k_{max}} \widehat{q}^{\,(1)}_{\frac{i}{d}}(u,v,l)} = \frac{\widehat{q}^{\,(1)}_{\frac{i}{d}}(u,v,l)}{\widehat{p}_{\frac{i}{d}}(u,v)}.$$

Non-parametric Estimation for model 1

Model 1 : we solve $MJ = P$ to obtain $\widehat{q}_{\frac{i}{d}}^{(1)}(u,v,l)$

$$M = \begin{pmatrix} \sum_{t=1}^{n} \mathbb{1}_u(t) A_0(t) A_0(t) & \cdots & \sum_{t=1}^{n} \mathbb{1}_u(t) A_0(t) A_d(t) \\ \vdots & \ddots & \vdots \\ \sum_{t=1}^{n} \mathbb{1}_u(t) A_d(t) A_0(t) & \cdots & \sum_{t=1}^{n} \mathbb{1}_u(t) A_d(t) A_d(t) \end{pmatrix}$$

$$J = \begin{pmatrix} \widehat{q}_0^{(1)}(u,v,l) \\ \vdots \\ \widehat{q}_{\frac{i}{d}}^{(1)}(u,v,l) \\ \vdots \\ \widehat{q}_1^{(1)}(u,v,l) \end{pmatrix} \text{ and } P = \begin{pmatrix} \sum_{t=1}^{n} \mathbb{1}_{uvl}(t) A_0(t) \\ \vdots \\ \sum_{t=1}^{n} \mathbb{1}_{uvl}(t) A_i(t) \\ \vdots \\ \sum_{t=1}^{n} \mathbb{1}_{uvl}(t) A_1(t) \end{pmatrix}$$

Where $\mathbb{1}_u(t) = \mathbb{1}_{\{J_{t-1}=u\}}(t)$, and $\mathbb{1}_{uvl}(t) = \mathbb{1}_{\{J_t=v, J_{t-1}=u, X_t=l\}}(t)$.

Non-parametric Estimation for model 2

Model 2

$$\widehat{p}_{\frac{i}{d}}(u,v) = \sum_{l=1}^{k_{max}} \widehat{q}_{\frac{i}{d}}^{(1)}(u,v,l) \textbf{ (as in Model 1)}$$

$$\widehat{f}(u,v,l) = \frac{\sum_{i=0}^{d} \widehat{q}_{\frac{i}{d}}^{(1)}(u,v,l)}{\sum_{i=0}^{d} \sum_{l=1}^{k_{max}} \widehat{q}_{\frac{i}{d}}^{(1)}(u,v,l)}, \quad \text{with} \quad \sum_{l=1}^{k_{max}} \widehat{f}(u,v,l) = 1$$

This leads to the estimated SM kernel for **Model 2** $\widehat{q}_{\frac{i}{d}}^{(2)}(u,v,l)$ being described through model 1 :

$$\widehat{q}_{\frac{i}{d}}^{(2)}(u,v,l) = \frac{\left(\sum_{l=1}^{k_{max}} \widehat{q}_{\frac{i}{d}}^{(1)}(u,v,l)\right)\left(\sum_{i=0}^{d} \widehat{q}_{\frac{i}{d}}^{(1)}(u,v,l)\right)}{\sum_{i=0}^{d} \sum_{l=1}^{k_{max}} \widehat{q}_{\frac{i}{d}}^{(1)}(u,v,l)}$$

Non-parametric Estimation for model 3

Model 3

$$\widehat{p}(u,v) = \frac{\sum_{i=0}^{d}\sum_{l=1}^{k_{max}} \widehat{q}_{\frac{i}{d}}^{(1)}(u,v,l)}{d+1}, \quad \text{with} \sum_{v \in E} \widehat{p}(u,v) = 1$$

$$\widehat{f}_{\frac{i}{d}}(u,v,l) = \frac{\widehat{q}_{\frac{i}{d}}^{(1)}(u,v,l)}{\sum_{l=1}^{k_{max}} \widehat{q}_{\frac{i}{d}}^{(1)}(u,v,l)} \quad \textbf{(as in model 1)}.$$

This leads to the estimated SM kernel for **Model 3** $\widehat{q}_{\frac{i}{d}}^{(3)}(u,v,l)$ being described through model 1 :

$$\widehat{q}_{\frac{i}{d}}^{(3)}(u,v,l) = \frac{\widehat{q}_{\frac{i}{d}}^{(1)}(u,v,l)\sum_{i=0}^{d}\sum_{l=1}^{k_{max}} \widehat{q}_{\frac{i}{d}}^{(1)}(u,v,l)}{(d+1)\sum_{l=1}^{k_{max}} \widehat{q}_{\frac{i}{d}}^{(1)}(u,v,l)}$$

In conclusion :

Model 1

$$\widehat{q}^{\,(1)}_{\frac{t}{n}}(u,v,l) = \sum_{i=0}^{d} A_i \widehat{q}^{\,(1)}_{\frac{i}{d}}(u,v,l)$$

Model 2

$$\widehat{q}^{\,(2)}_{\frac{t}{n}}(u,v,l) = \sum_{i=0}^{d} A_i(t) \left(\frac{\left(\sum_{l=1}^{k_{max}} \widehat{q}^{\,(1)}_{\frac{i}{d}}(u,v,l)\right)\left(\sum_{i=0}^{d} \widehat{q}^{\,(1)}_{\frac{i}{d}}(u,v,l)\right)}{\sum_{i=0}^{d}\sum_{l=1}^{k_{max}} \widehat{q}^{\,(1)}_{\frac{i}{d}}(u,v,l)} \right)$$

Model 3

$$\widehat{q}^{\,(3)}_{\frac{t}{n}}(u,v,l) = \sum_{i=0}^{d} A_i(t) \left(\frac{\widehat{q}^{\,(1)}_{\frac{i}{d}}(u,v,l) \sum_{i=0}^{d}\sum_{l=1}^{k_{max}} \widehat{q}^{\,(1)}_{\frac{i}{d}}(u,v,l)}{(d+1)\sum_{l=1}^{k_{max}} \widehat{q}^{\,(1)}_{\frac{i}{d}}(u,v,l)} \right)$$

Parametric Estimation

Estimators of parameters for the discrete distribution of sojourn time

We note $\widehat{m_1} = \sum_{x=1}^{k_{max}} x\widehat{f}(x)$ and $\widehat{m_2} = \sum_{x=1}^{k_{max}} x^2\widehat{f}(x) - \widehat{m_1}^2$.

- Geometric(p) : $f(x) = p(1-p)^{x-1}, x = 1, \dots, k_{max}$. Then $\widehat{p} = 1/\widehat{m_1}$.
- Poisson(λ) : $f(x) = \frac{\lambda^{x-1} \exp -\lambda}{(x-1)!}$, $x = 1, \dots, k_{max}$. Then $\widehat{\lambda} = \widehat{m_1}$.
- Negative Binomial (α, p) : $f(x) = \frac{\Gamma(x+\alpha-1)}{\Gamma(\alpha)(x-1)!} p^\alpha (1-p)^{x-1}$, $x = 1, \dots, k_{max}$. Therefore:

$$\widehat{p} = \frac{\widehat{m_1}}{\widehat{m_2}}, \qquad \widehat{\alpha} = \widehat{m_1}\frac{\widehat{p}}{1-\widehat{p}} = \frac{\widehat{m_1}^2}{\widehat{m_2} - \widehat{m_1}}$$

- Discrete Weibull (q, β) : $f(x) = q^{(x-1)^\beta} - q^{x^\beta}$, $x = 1, \dots, k_{max}$. Therefore:

$$\widehat{q} = 1 - \widehat{f}(1), \qquad \widehat{\beta} = \frac{\sum_{i=2}^{k_{max}} \log_i(\log_{\widehat{q}}(\sum_{j=1}^{i} \widehat{f}(j))}{k_{max} - 1}$$

Plan

dsmmR - Main Functions

It constitutes of the main functions :

1. `fit_dsmm()`:

   Fit a DSMM on a given sequence. Parametric or non-parametric estimation is possible (**Model 1**, **2** or **3**).

2. `parametric_dsmm()` & `nonparametric_dsmm()`:

   Define a parametric or non-parametric DSMM (**Model 1**, **2** or **3**)

3. `simulate.dsmm()`:

   Generate a sequence of states with a maximum number of simulations equal to $n+1$ (model size).

4. `get_kernel()`:
   Compute the DSM kernel $q_{\frac{t}{n}}$.

$d+1$ Sojourn Time Distributions : f_0, $f_{\frac{1}{2}}$ and f_1

$f_0(u,v,l=1) = \begin{pmatrix} 0 & 0.2 & 0.7 \\ 0.3 & 0 & 0.4 \\ 0.2 & 0.8 & 0 \end{pmatrix}$	$f_{\frac{1}{2}}(u,v,l=1) = \begin{pmatrix} 0 & 0.3333333 & 0.4 \\ 0.3 & 0 & 0.4 \\ 0.2 & 0.1 & 0 \end{pmatrix}$	$f_1(u,v,l=1) = \begin{pmatrix} 0 & 0.3 & 0.3 \\ 0.3 & 0 & 0.5 \\ 0.05 & 0.1 & 0 \end{pmatrix}$
$f_0(u,v,l=2) = \begin{pmatrix} 0 & 0.3 & 0.2 \\ 0.2 & 0 & 0.5 \\ 0.1 & 0.15 & 0 \end{pmatrix}$	$f_{\frac{1}{2}}(u,v,l=2) = \begin{pmatrix} 0 & 0.3333333 & 0.4 \\ 0.4 & 0 & 0.2 \\ 0.3 & 0.4 & 0 \end{pmatrix}$	$f_1(u,v,l=2) = \begin{pmatrix} 0 & 0.2 & 0.6 \\ 0.3 & 0 & 0.35 \\ 0.9 & 0.2 & 0 \end{pmatrix}$
$f_0(u,v,l=3) = \begin{pmatrix} 0 & 0.5 & 0.1 \\ 0.5 & 0 & 0.1 \\ 0.7 & 0.05 & 0 \end{pmatrix}$	$f_{\frac{1}{2}}(u,v,l=3) = \begin{pmatrix} 0 & 0.3333333 & 0.2 \\ 0.3 & 0 & 0.4 \\ 0.5 & 0.5 & 0 \end{pmatrix}$	$f_1(u,v,l=3) = \begin{pmatrix} 0 & 0.5 & 0.1 \\ 0.4 & 0 & 0.15 \\ 0.05 & 0.7 & 0 \end{pmatrix}$

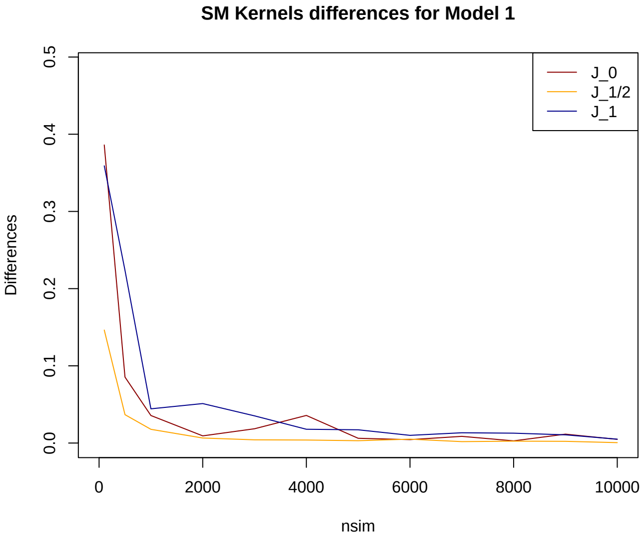
$d+1$ Transition Matrices - $p_0, p_{\frac{1}{2}}$ and p_1 and metric

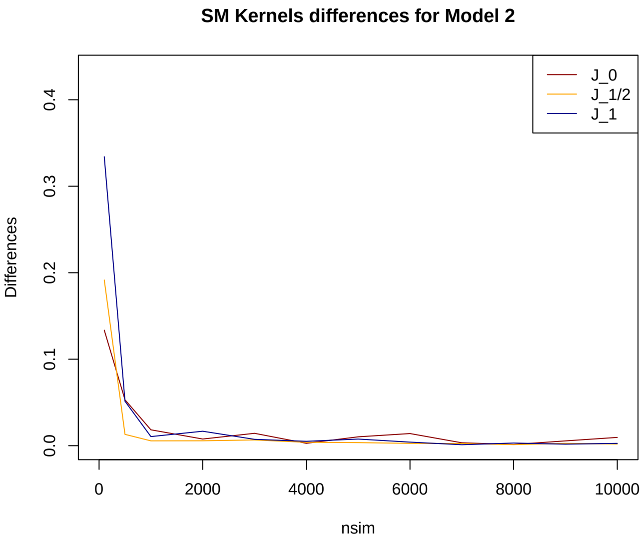
$p_0(u,v) = \begin{pmatrix} 0 & 0.1 & 0.9 \\ 0.5 & 0 & 0.5 \\ 0.3 & 0.7 & 0 \end{pmatrix}$	$p_{\frac{1}{2}}(u,v) = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.7 & 0 & 0.3 \\ 0.6 & 0.4 & 0 \end{pmatrix}$	$p_1(u,v) = \begin{pmatrix} 0 & 0.2 & 0.8 \\ 0.6 & 0 & 0.4 \\ 0.7 & 0.3 & 0 \end{pmatrix}$

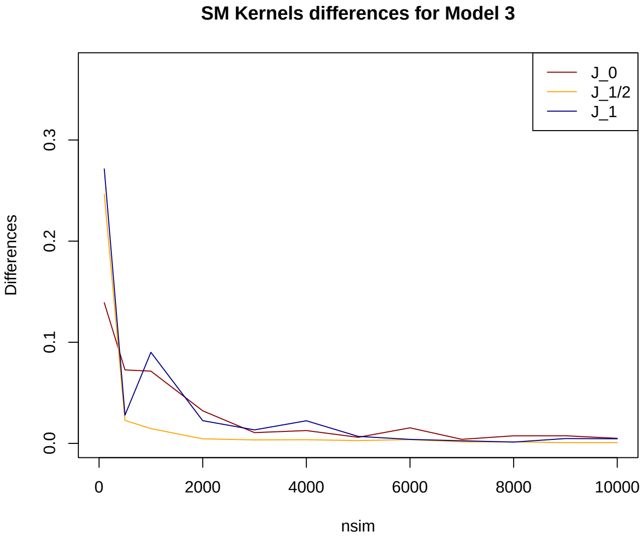
We are going to use the following metric, defining the distance between the $d+1$ theoretical kernels $q_{\frac{i}{d}}^{(M)}$ with the estimated ones $\widehat{q}_{\frac{i}{d}}^{(M)}$, for all 3 models $M = 1, 2, 3:$

$$d\left(q_{\frac{i}{d}}^{(M)}, \widehat{q}_{\frac{i}{d}}^{(M)}\right) = \sum_{uvl}\left(q_{\frac{i}{d}}^{(M)} - \widehat{q}_{\frac{i}{d}}^{(M)}\right)^2,$$

where $u, v \in E, l \in \{1, \ldots, k_{max}\}$.


SM Kernels differences for Model 1
J_0
J_1/2
J_1
Differences
0.0
0.1
0.2
0.3
0.4
0.5
0
2000
4000
6000
8000
10000
nsim


SM Kernels differences for Model 2
J_0
J_1/2
J_1
Differences
0.0
0.1
0.2
0.3
0.4
0
2000
4000
6000
8000
10000
nsim


SM Kernels differences for Model 3
J_0
J_1/2
J_1
Differences
0.0
0.1
0.2
0.3
0
2000
4000
6000
8000
10000
nsim

R Package

[9] **V. S. Barbu, I. Mavrogiannis**[3] **and N. Vergne**, 2022. dsmmR : a R package for Estimation and Simulation of Drifting Semi-Markov Models. Available at https://cran.r-project.org/web/packages/dsmmR/index.html.

3. Thanks to FEDER DATALAB

Plan

Concluding remarks

Conclusions

- `drimmR` : drifting Markov models, reliability, different types of datas.
- `smmR` : semi-Markov models, reliability, different types of sojourn time, censoring, non-parametric and parametric.
- `dsmmR` : drifting semi-Markov models, models 1, 2 and 3, non-parametric and parametric.

Future directions

- `hsmmR` : hidden semi-Markov models, reliability, different types of sojourn time, censoring, non-parametric and parametric (thanks to ANR HSMM-INCA).
- `gdrimmR` : generalized drifting Markov models, reliability (thanks to CNRS for E. Kalligeris contract).
- Eye-Tracking : work in collaboration with CETAPS to analyse climbing data. Clustering of models.