

Estimation statistique dans une population branchante structurée par une diffusion

Aline Marguet

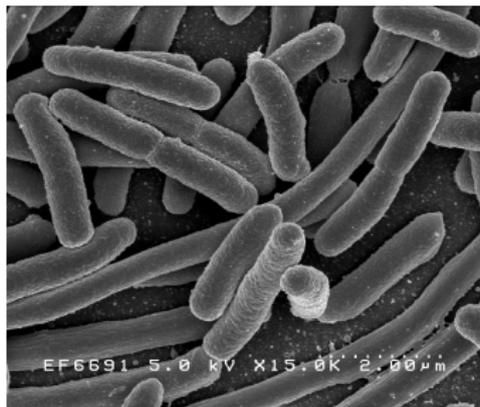
En collaboration avec Marc Hoffmann

29 septembre 2017



Un modèle pour décrire l'évolution d'une population de cellules

Population de cellules caractérisées par un trait (âge, nombre de parasites, taille, ...).



Étude du vieillissement cellulaire : [DLJB14].

1 Modèle probabiliste

2 Estimation non-paramétrique

3 Estimation paramétrique du taux de division

Description du modèle

- Le trait $(\phi_x(t))_{t \geq 0}$ de chaque individu évolue suivant une diffusion d'équation :

$$d\phi_x(t) = r(\phi_x(t))dt + \sigma(\phi_x(t))dW_t, \quad \phi_x(0) = x,$$

où $r, \sigma : \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions mesurables et $(W_t)_{t \geq 0}$ est un mouvement brownien standard.

Description du modèle

- Le trait $(\phi_x(t))_{t \geq 0}$ de chaque individu évolue suivant une diffusion d'équation :

$$d\phi_x(t) = r(\phi_x(t))dt + \sigma(\phi_x(t))dW_t, \quad \phi_x(0) = x,$$

où $r, \sigma : \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions mesurables et $(W_t)_{t \geq 0}$ est un mouvement brownien standard.

- Chaque individu se divise au temps t à un taux $B(\phi_x(t))$, i.e.

$$\mathbb{P}(\beta(u) > t | \alpha(u), (\phi_x(s))_{\alpha(u) \leq s \leq t}) = \exp\left(-\int_{\alpha(u)}^t B(\phi_x(s))ds\right),$$

avec $\alpha(u), \beta(u)$ les temps de division et de naissance de u .

Description du modèle

- Le trait $(\phi_x(t))_{t \geq 0}$ de chaque individu évolue suivant une diffusion d'équation :

$$d\phi_x(t) = r(\phi_x(t))dt + \sigma(\phi_x(t))dW_t, \quad \phi_x(0) = x,$$

où $r, \sigma : \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions mesurables et $(W_t)_{t \geq 0}$ est un mouvement brownien standard.

- Chaque individu se divise au temps t à un taux $B(\phi_x(t))$, i.e.

$$\mathbb{P}(\beta(u) > t | \alpha(u), (\phi_x(s))_{\alpha(u) \leq s \leq t}) = \exp\left(-\int_{\alpha(u)}^t B(\phi_x(s))ds\right),$$

avec $\alpha(u), \beta(u)$ les temps de division et de naissance de u .

- À sa mort, un individu de trait x est remplacé par 2 descendants de traits à la naissance θx et $(1 - \theta)x$, où $\theta \in [0, 1]$ est une v.a. de densité associée $\kappa(\theta)$ symétrique.

Description du modèle

- Le trait $(\phi_x(t))_{t \geq 0}$ de chaque individu évolue suivant une diffusion d'équation :

$$d\phi_x(t) = r(\phi_x(t))dt + \sigma(\phi_x(t))dW_t, \quad \phi_x(0) = x,$$

où $r, \sigma : \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions mesurables et $(W_t)_{t \geq 0}$ est un mouvement brownien standard.

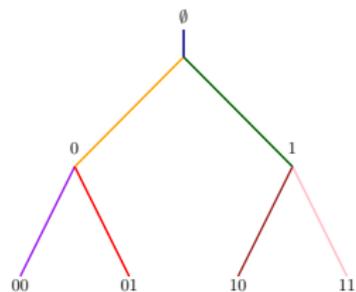
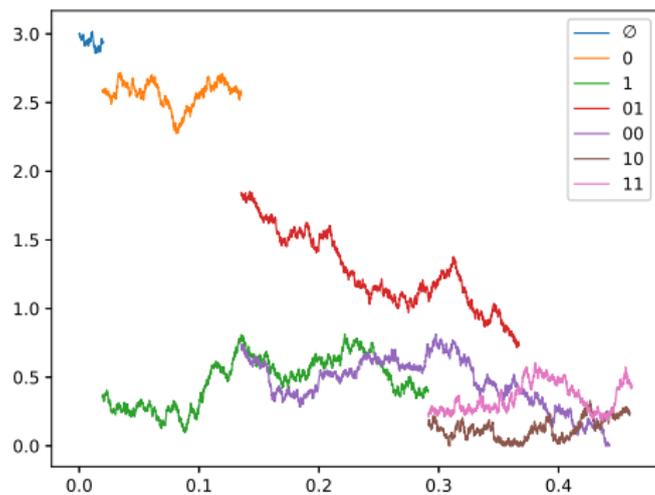
- Chaque individu se divise au temps t à un taux $B(\phi_x(t))$, i.e.

$$\mathbb{P}(\beta(u) > t | \alpha(u), (\phi_x(s))_{\alpha(u) \leq s \leq t}) = \exp\left(-\int_{\alpha(u)}^t B(\phi_x(s))ds\right),$$

avec $\alpha(u), \beta(u)$ les temps de division et de naissance de u .

- À sa mort, un individu de trait x est remplacé par 2 descendants de traits à la naissance θx et $(1 - \theta)x$, où $\theta \in [0, 1]$ est une v.a. de densité associée $\kappa(\theta)$ symétrique.
- Conditionnellement au trait de leur ancêtre, les descendants évoluent indépendamment les uns des autres.

Exemple de trajectoire



Chaîne de Markov bifurcante

On considère les notations de Ulam-Harris-Neveu, pour $n, m \geq 0$, on note :

$$\mathbb{T} = \bigcup_{m \in \mathbb{N}} \{0, 1\}^m, \quad \mathbb{T}_n = \bigcup_{m=0}^n \{0, 1\}^m, \quad \mathbb{T}_n^* = \mathbb{T}_n \setminus \{\emptyset\}.$$

On note X_u le trait à la naissance d'un individu $u \in \mathbb{T}$. On suppose que l'on dispose des observations

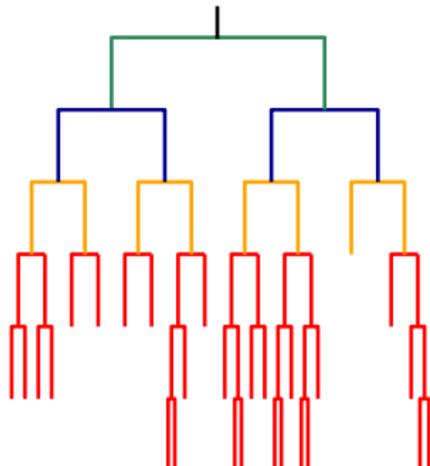
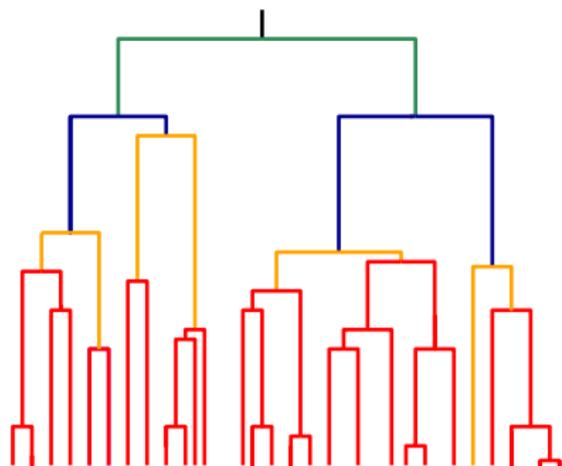
$$\mathbb{X}^n = (X_u)_{u \in \mathbb{T}_n}.$$

Le processus $(X_u, u \in \mathbb{T})$ est une chaîne de Markov bifurcante [Guy07] de noyau de transition \mathcal{P} de \mathbb{R} dans $\mathbb{R} \times \mathbb{R}$ tel que :

$$\mathbb{E} \left[\prod_{u \in \mathbb{G}_m} \psi_u(X_u, X_{u0}, X_{u1}) \mid \mathcal{F}_m \right] = \prod_{u \in \mathbb{G}_m} \mathcal{P} \psi_u(X_u),$$

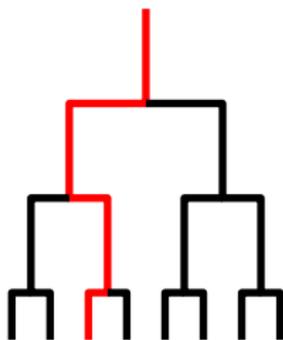
pour tout $m \geq 0$.

Temps généalogique



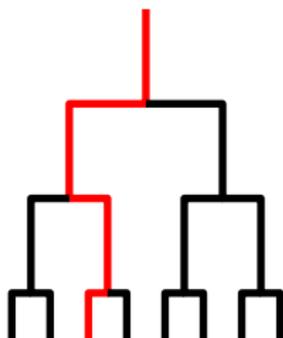
Processus de la cellule étiqueté

On considère la chaîne de Markov Y correspondant au trait d'un individu le long d'une lignée : à chaque branchement, on choisit l'un des deux descendants avec probabilité $1/2$.



Processus de la cellule étiquetée

On considère la chaîne de Markov Y correspondant au trait d'un individu le long d'une lignée : à chaque branchement, on choisit l'un des deux descendants avec probabilité $1/2$.



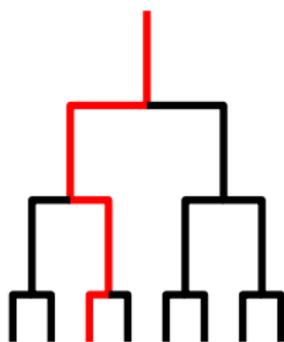
La transition de Y est donnée par $Q = (\mathcal{P}_1 + \mathcal{P}_2)/2$, avec

$$\mathcal{P}_1(x, dy) = \int_{\mathbb{R}} \mathcal{P}(x, dy_1 dy_2) = \mathcal{P}_2(x, dy) = \int_{\mathbb{R}} \mathcal{P}(x, dy_1 dy),$$

car κ est symétrique.

Processus de la cellule étiqueté

On considère la chaîne de Markov Y correspondant au trait d'un individu le long d'une lignée : à chaque branchement, on choisit l'un des deux descendants avec probabilité $1/2$.



La transition de Y est donnée par $Q = \mathcal{P}_1$, avec

$$\mathcal{P}_1(x, dy) = \int_{\mathbb{R}} \mathcal{P}(x, dy_1 dy_2) = \mathcal{P}_2(x, dy) = \int_{\mathbb{R}} \mathcal{P}(x, dy_1 dy),$$

car κ est symétrique.

Convergence de la chaîne étiquetée

On note

$$|\varphi|_V = \sup_{x \in \mathbb{R}} \frac{|\varphi(x)|}{1 + V(x)}.$$

Théorème

Sous des hypothèses appropriées sur la dynamique du flot, \mathcal{Q} admet une mesure invariante ν . De plus, il existe $C > 0$ et $\rho \in (0, 1)$ tels que pour tout $x \in \mathbb{R}$ et $m \in \mathbb{N}$:

$$|\mathcal{Q}^m \varphi - \nu(\varphi)|_V \leq C \rho^m |\varphi - \nu(\varphi)|_V$$

pour toute fonction mesurable $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ telle que $|\varphi|_V < \infty$, avec $V(x) = x^2$.

Preuve : [HM11]

On considère la mesure empirique donnée pour toute fonction test ψ par :

$$\mathcal{M}_n(\psi) = \frac{1}{|\mathbb{T}_n^*|} \sum_{u \in \mathbb{T}_n^*} \psi(X_{u^-}, X_u)$$

On considère la mesure empirique donnée pour toute fonction test ψ par :

$$\mathcal{M}_n(\psi) = \frac{1}{|\mathbb{T}_n^*|} \sum_{u \in \mathbb{T}_n^*} \psi(X_{u^-}, X_u)$$

Théorème

Soit $\mu \in \mathcal{M}_P(\mathbb{R})$ telle que $\mu(V^2) < \infty$ et $\psi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ une fonction à support compact. Sous des hypothèses assurant l'ergodicité de Y , si $\rho \leq \frac{1}{2}$,

$$\mathbb{E}_\mu [(\mathcal{M}_n(\psi) - \nu(\psi))^2] \lesssim |\mathbb{T}_n|^{-1} C(\psi), \quad \forall n \in \mathbb{N},$$

où \lesssim signifie à une constante près dépendante de \mathcal{Q} et $\text{supp}(\psi)$ et

$$C(\psi) = |\psi^2|_\mu + |\psi^* \psi|_\mu + \left(1 + \mu(V^2)^{1/2}\right) |\psi_\star|_1 |\psi|_1.$$

$$\psi^*(x) = \sup_{y \in \mathbb{R}} |\psi(x, y)|, \quad \psi_\star(y) = \sup_{x \in \mathbb{R}} |\psi(x, y)|$$

$$|\psi|_\mu = \int_{\mathbb{R} \times \mathbb{R}} |\psi(x, y)| \mu(dx) dy + |\psi_\star|_1 \wedge |\psi|_1.$$

Plan

- 1 Modèle probabiliste
- 2 Estimation non-paramétrique**
- 3 Estimation paramétrique du taux de division

Définition

Une fonction $G : \mathbb{R} \rightarrow \mathbb{R}$ est un noyau d'ordre k si elle est à support compact et si elle satisfait $\int_{\mathbb{R}} x^l G(x) dx = \mathbf{1}_{\{l=0\}}$ pour tout $l = 0, \dots, k$.

Exemples :

- $x \rightarrow \mathbf{1}_{\{x \leq 1\}}$ est un noyau d'ordre 0,
- $x \rightarrow 1/\sqrt{2\pi} e^{-x^2/2}$ est un noyau d'ordre 1,
- on peut construire des noyaux de tout ordre à l'aide des polynômes de Legendre.

Estimateur à noyau de la mesure invariante

Soit G un noyau d'ordre k et $h > 0$ un paramètre de lissage. On pose

$$G_h(y) = h^{-1}G(h^{-1}y), \quad \forall y \in \mathbb{R}.$$

Comme $\mathcal{M}_n(\varphi)$ converge vers $\nu(\varphi)$, on a

$$\mathcal{M}_n(G_h(\cdot - x_0)) \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}} G_h(x - x_0)\nu(x)dx.$$

Estimateur à noyau de la mesure invariante

Soit G un noyau d'ordre k et $h > 0$ un paramètre de lissage. On pose

$$G_h(y) = h^{-1}G(h^{-1}y), \quad \forall y \in \mathbb{R}.$$

Comme $\mathcal{M}_n(\varphi)$ converge vers $\nu(\varphi)$, on a

$$\mathcal{M}_n(G_h(\cdot - x_0)) \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}} G_h(x - x_0)\nu(x)dx.$$

Un estimateur de $\nu(x_0)$, pour $x_0 \in \mathbb{R}$ est donc donné par

$$\hat{\nu}_n(x_0) = \mathcal{M}_n(G_h(\cdot - x_0)).$$

Estimateur à noyau de la densité

De la même façon, on considère G un noyau d'ordre k et $h_1, h_2 > 0$ deux paramètres de lissage. On pose

$$G_{h_1, h_2}^{\otimes 2}(x, y) = h_1^{-1} h_2^{-1} G(h_1^{-1} x) G(h_2^{-1} y), \quad \forall x, y \in \mathbb{R}.$$

On a alors la convergence suivante :

$$\begin{aligned} & \mathcal{M}_n(G_{h_1, h_2}^{\otimes 2}(\cdot - x_0, \cdot - y_0)) \\ & \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R} \times \mathbb{R}} G_{h_1, h_2}^{\otimes 2}(x - x_0, y - y_0) q(x, y) \nu(x) dy dx. \end{aligned}$$

Estimateur à noyau de la densité

De la même façon, on considère G un noyau d'ordre k et $h_1, h_2 > 0$ deux paramètres de lissage. On pose

$$G_{h_1, h_2}^{\otimes 2}(x, y) = h_1^{-1} h_2^{-1} G(h_1^{-1} x) G(h_2^{-1} y), \quad \forall x, y \in \mathbb{R}.$$

On a alors la convergence suivante :

$$\begin{aligned} & \mathcal{M}_n(G_{h_1, h_2}^{\otimes 2}(\cdot - x_0, \cdot - y_0)) \\ & \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R} \times \mathbb{R}} G_{h_1, h_2}^{\otimes 2}(x - x_0, y - y_0) q(x, y) \nu(x) dy dx. \end{aligned}$$

Un estimateur de la densité $q(x_0, y_0)$, pour $x_0, y_0 \in \mathbb{R}$ est alors donné par

$$\hat{q}_n(x_0, y_0) = \frac{\mathcal{M}_n(G_{h_1, h_2}^{\otimes 2}(\cdot - x_0, \cdot - y_0))}{\mathcal{M}_n(G_h(\cdot - x_0)) \vee \varpi}.$$

Convergence des estimateurs

Soit $\alpha, \beta > 0$ et G un noyau d'ordre $k > \max(\alpha, \beta)$. On pose

$$h = |\mathbb{T}_n|^{\frac{-1}{2\beta+1}}, \quad h_1 = |\mathbb{T}_n|^{\frac{-s(\alpha,\beta)}{(\alpha \wedge \beta)(2s(\alpha,\beta)+1)}}, \quad h_2 = |\mathbb{T}_n|^{\frac{-s(\alpha,\beta)}{\beta(2s(\alpha,\beta)+1)}}, \quad \varpi_n \rightarrow 0.$$

Théorème

Sous des hypothèses assurant l'ergodicité de la chaîne Y , on a

$$\left(\mathbb{E} \left[\left(\hat{\nu}_n(x_0) - \nu(x_0) \right)^2 \right] \right)^{1/2} \lesssim |\mathbb{T}_n|^{-\beta/(2\beta+1)},$$

$$\left(\mathbb{E} \left[\left(\hat{q}_n(x_0, y_0) - q(x_0, y_0) \right)^2 \right] \right)^{1/2} \lesssim \varpi_n^{-1} |\mathbb{T}_n|^{-s(\alpha,\beta)/(2s(\alpha,\beta)+1)},$$

avec $s(\alpha, \beta)^{-1} = (\alpha \wedge \beta)^{-1} + \beta^{-1}$, uniformément en \mathcal{Q} pour \mathcal{Q} dans une certaine classe de régularité de Hölder dépendant de α et β .

Plan

- 1 Modèle probabiliste
- 2 Estimation non-paramétrique
- 3 Estimation paramétrique du taux de division

Estimation du taux de division

La dépendance en B de la transition est complexe mais explicite :

$$q(x, y) = \int_0^1 \frac{\kappa(z)}{z} B(y/z) \sigma(y/z)^{-2} \mathbb{E} \left[\int_0^\infty e^{-\int_0^t B(\phi_x(s)) ds} dL_t^{y/z}(\phi_x) \right] dz,$$

avec $L_t^y(\phi_x)$ le temps local au temps t au point y de $(\phi_x(t), t \geq 0)$.

Pour l'estimation par maximum de vraisemblance de B , on se restreint au cas d'un trait évoluant dans un compact $[0, L] \subset \mathbb{R}$:

$$d\phi_x(t) = r(\phi_x(t))dt + \sigma(\phi_x(t))dW_t + d\ell_t,$$

où $(\ell_t)_{t \geq 0}$, vérifiant $\ell_t = \int_0^t (\mathbf{1}_{\{X_s=0\}} + \mathbf{1}_{\{X_s=L\}}) d\ell_s$, traduit les réflexions au bord du domaine et $(W_t)_{t \geq 0}$ est un mouvement brownien standard.

Modèle paramétrique

On suppose que le taux de division B appartient à une classe

$$\mathcal{B} = \{B : [0, L] \rightarrow \mathbb{R}, B(x) = B_0(\vartheta, x), x \in [0, L], \vartheta \in \Theta\},$$

où $x \mapsto B_0(x, \vartheta)$ est connu à un paramètre $\vartheta \in \Theta$ près, et $\Theta \subset \mathbb{R}^d$ est compact.

But

Estimer ϑ à partir de $(X_u, u \in \mathbb{T}_n)$.

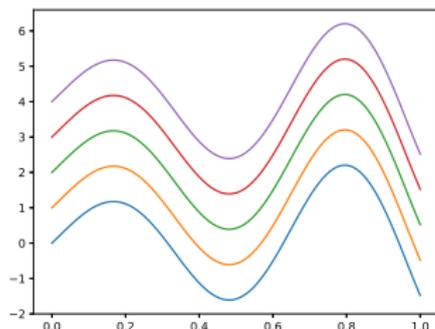
Identifiabilité du modèle

Définition

Une classe \mathcal{B} de fonctions de $[0, L]$ dans $[0, \infty)$ est ordonnée si $\varphi_1, \varphi_2 \in \mathcal{B}$ entraîne soit $\varphi_1(x) \leq \varphi_2(x)$ pour tout $x \in [0, L]$ ou $\varphi_2(x) \leq \varphi_1(x)$ pour tout $x \in [0, L]$.

Exemples :

- $\{x \rightarrow \vartheta, \vartheta \in \Theta\}$,
- $\{x \rightarrow 1 + \vartheta x, \vartheta \in \Theta\}$.



Vraisemblance

La vraisemblance est donnée par :

$$\mathcal{L}_n(\vartheta, (X_u, u \in \mathbb{T}_n)) = \prod_{u \in \mathbb{T}_n^*} q_{\vartheta}(X_{u^-}, X_u),$$

où X_{u^-} correspond au trait de l'ancêtre de u . On considère alors l'estimateur du maximum de vraisemblance (EMV) donné par :

$$\hat{\vartheta}_n \in \operatorname{argmax}_{\vartheta \in \Theta} \left\{ \frac{1}{|\mathbb{T}_n^*|} \sum_{u \in \mathbb{T}_n^*} \log(q_{\vartheta}(X_{u^-}, X_u)) \right\}.$$

Théorème

Sous des hypothèses de régularité du taux de division, $\hat{\vartheta}_n$ converge en probabilité vers ϑ lorsque n tend vers l'infini.

Normalité asymptotique

On définit $\Psi(\vartheta)$ la matrice d'information de Fisher dont les coefficients sont donnés pour tout $1 \leq i, j \leq d$ par :

$$\Psi(\vartheta)_{i,j} = \nu_{\vartheta} \mathcal{Q}(\vartheta) \left(\frac{\partial_{\vartheta_i} \mathbf{q}_{\vartheta} \partial_{\vartheta_j} \mathbf{q}_{\vartheta}}{q_{\vartheta}^2} \right).$$

Théorème

Sous des hypothèses de régularité et si $\Psi(\vartheta)$ est inversible, pour tout ϑ dans l'intérieur de Θ , on a :

$$\mathbb{T}_n^{1/2} \left(\widehat{\vartheta}_n - \vartheta \right) \rightarrow \mathcal{N} \left(0, \Psi(\vartheta)^{-1} \right),$$

en loi lorsque n tend vers l'infini, où $\mathcal{N} \left(0, \Psi(\vartheta)^{-1} \right)$ désigne la loi normale d -dimensionnelle de moyenne 0 et de matrice de covariance l'inverse de la matrice de Fisher $\psi(\vartheta)$.

Simulations

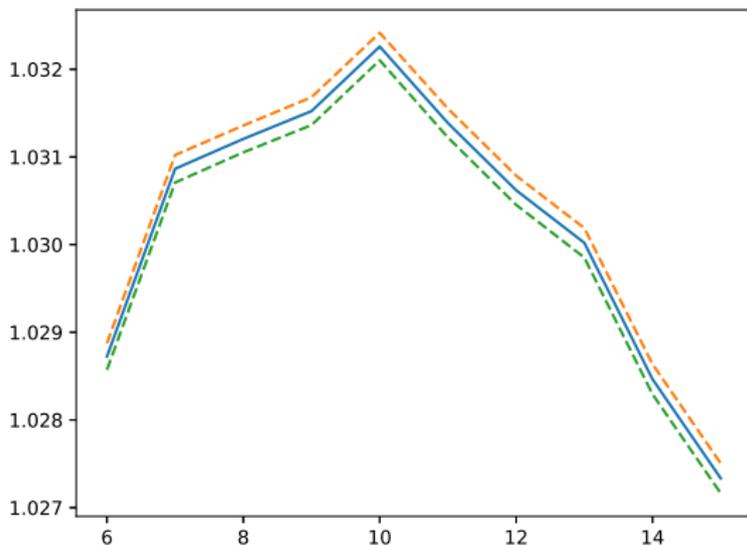


Figure: Log-vraisemblance calculée à partir de données simulées avec $B(x, \vartheta) = 1 + \vartheta x$, $\vartheta = 10$, $d\phi_x(t) = -10dt + dW_t$.

 A. Denoth Lippuner, T. Julou, and Y. Barral.

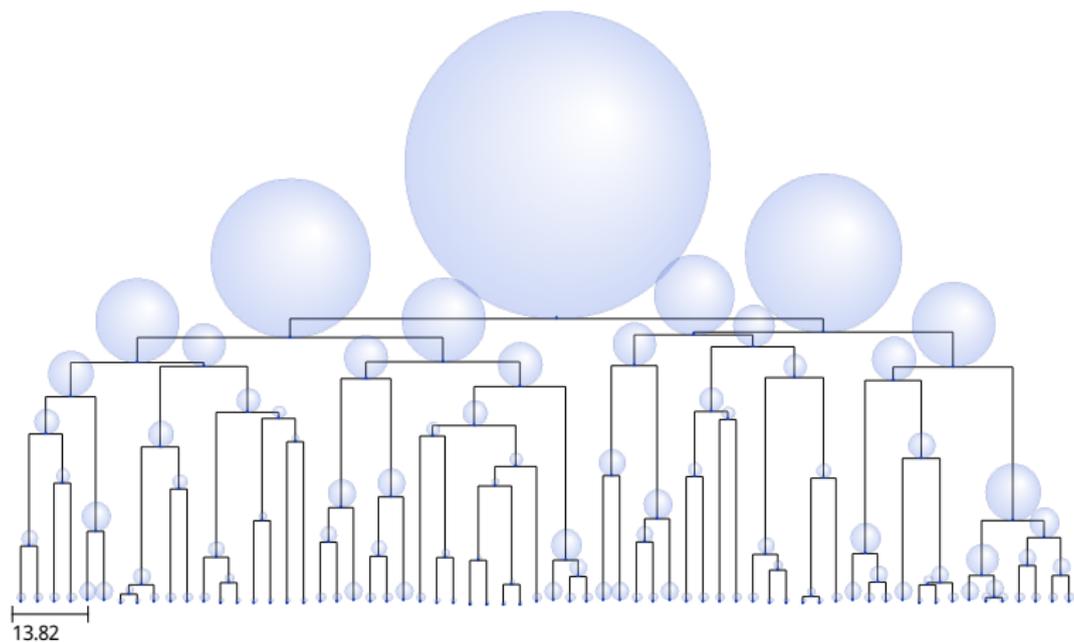
Budding yeast as a model organism to study the effects of age.
FEMS Microbiology Reviews, 38(2):300–325, 2014.

 J. Guyon.

Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging.
The Annals of Applied Probability, 17(5/6):1538–1569, 2007.

 M. Hairer and J. C Mattingly.

Yet another look at Harris' ergodic theorem for Markov chains.
Seminar on Stochastic Analysis, Random Fields and Applications VI,
pages 109–117, 2011.



Merci pour votre attention !