Atti Accademia Nazionale dei Lincei

Classe Scienze Fisiche Matematiche Naturali

Rendiconti

Colette Andrieu, Radu Theodorescu

Sur les probabilités absolues stationnaires pour des processus de Markov à temps continu

Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti, Serie 8, Vol. **54** (1973), n.6, p. 892–897.

Accademia Nazionale dei Lincei

<http://www.bdim.eu/item?id=RLINA_1973_8_54_6_892_0>

Calcolo delle probabilità. — *Sur les probabilités absolues stationnaires pour des processus de Markov à temps continu* (*). Nota di COLETTE ANDRIEU e RADU THEODORESCU, presentata (**) dal Socio B. SEGRE.

RIASSUNTO. — In questa Nota vengono indicati i legami tra la probabilità assoluta stazionaria (quando essa esiste, è unica e soddisfa a certe condizioni supplementari) relativa ad un processo di Markov omogeneo a tempo continuo, e quella relativa ad un processo di Markov omogeneo a tempo discreto che ne deriva. Per facilitare i calcoli, si esamina innanzitutto, ed in dettaglio, il caso di un numero finito di stati; si indica poi l'estensione ai casi in cui il numero degli stati risulta infinito.

Dans cette Note, nous indiquons les liens entre la probabilité absolue stationnaire (quand elle existe, est unique et sous certaines autres conditions) d'un processus de Markov homogène, à temps continu, et celle du processus de Markov homogène, à temps discret, qui en est extrait. Pour faciliter les calculs, nous examinons d'abord en détail le cas d'un nombre fini d'états. Nous indiquons ensuite l'extension aux cas d'un nombre non fini d'états.

1. Notons $N^* = \{1, 2, \cdots\}$, $R_+ = [0, \infty)$ et considérons un processus de Markov homogène, à temps continu, ayant comme espace des états un ensemble fini X et comme matrice (fonction) de transition $\mathbf{P}(t) = (p_{ij}(t))$, $i, j \in X$, définie pour tout $t > 0$. Supposons que $\mathbf{P}(t)$ est une matrice de transition standard, c'est-à-dire telle que

$$(1.1) \qquad \lim_{t \to 0} p_{ij}(t) = \delta_{ij}$$

pour tout $i, j \in X$, où δ_{ij} est le symbole de Kronecker; alors nous pouvons étendre la définition de chaque $p_{ij}(t)$ à R_+ en posant $p_{ij}(0) = \delta_{ij}$. Pour une matrice de transition standard $\mathbf{P}(t) = (p_{ij}(t))$ les dérivées (à droite) à l'origine de $p_{ij}(t)$ existent et sont finies et posons $q_{ij} = p'_{ij}(0)$ pour tout $i, j \in X$, $\mathbf{Q} = (q_{ij})$, de même que $q_i = -q_{ii}$ pour tout $i \in X$; dans ce cas $q_i = \sum_{j \neq i} q_{ij}$. Supposons de plus que

$$(1.2) \qquad \lim_{t \to \infty} p_{ij}(t) = \pi_j$$

pour tout $i, j \in X$. Dans ce cas $\pi = (\pi_i, i \in X)$ est l'unique probabilité absolue stationnaire, c'est-à-dire π est l'unique probabilité telle que $\pi\mathbf{P}(t) = \pi$ pour tout $t \in R_+$.

(*) Recherche supportée par la subvention A-7223 du Conseil National de Recherche du Canada.

(**) Nella seduta del 19 giugno 1973.

Soit maintenant $\{\xi_t , t \in R_+\}$ une fonction aléatoire de Markov homogène, séparable, définie sur un espace probabilisé $(\Omega , \mathcal{K} , \Pr)$ et attaché au processus de Markov homogène, à temps continu, considéré ci–dessus; évidemment dans ce cas $p_{ij}(t) = \Pr\{\xi_{s+t}(\omega) = j \mid \xi_s(\omega) = i\}$ pour tout $i , j \in X$ et $t > 0 , p_{ij}(t)$ étant la valeur commune de cette probabilité pour tout $s \in R_+$ pour lequel elle est définie. Prèsque toutes les trajectoires $\xi_t(\omega)$, $\omega \in \Omega$, de cette fonction aléatoire sont des fonctions de sauts. Il est donc commode de supposer que chaque trajectoire est une fonction de sauts continue à droite. Supposons de plus que

$$\min_{i \in X} q_i > 0 , \tag{1.3}$$

ce qui veut dire qu'il n'y a pas d'états absorbants.

Désignons ensuite par $\{\zeta_n(\omega) , n \in N^*\}$ les états successifs $(\zeta_n(\omega) \neq \zeta_{n+1}(\omega))$ pris par la trajectoire $\xi_t(\omega)$, et par $\{\rho_n(\omega) , n \in N^*\}$ les durées respectives où cette trajectoire se trouve dans ces états. Alors $\xi_t(\omega) = \zeta_1(\omega)$ si $0 \le t < \rho_1(\omega)$ et $\xi_t(\omega) = \zeta_n(\omega)$ si $\rho_1(\omega) + \cdots + \rho_{n-1}(\omega) \le t < \rho_1(\omega) + \cdots + \rho_n(\omega)$ pour tout $n > 1$. Ces nombres $\{\rho_n(\omega) , n \in N^*\}$ sont tous finis, d'après (1.3). Il existe donc deux suites de variables aléatoires réelles $\{\zeta_n , n \in N^*\}$ et $\{\rho_n , n \in N^*\}$, et $\{\zeta_n , n \in N^*\}$ est une fonction aléatoire de Markov homogène, à temps discret, dont l'espace des états est X et dont la matrice de transition est $\mathbf{M} = (m_{ij})$, où $m_{ij} = q_{ij}/q_i$ si $i \neq j$ et $m_{ij} = 0$ si $i = j$. Le processus de Markov homogène, à temps discret, dont l'espace des états est X et dont la matrice de transition est $\mathbf{M}$ est dit *extrait* du processus de Markov homogène, à temps continu, initial (voir, par exemple [6], p. 259).

Notre but, dans cette Note, est de démontrer d'abord que le processus de Markov extrait possède une répartition de probabilité absolue stationnaire $\mu = \{\mu_i , i \in I\}$ unique, $\mu\mathbf{M} = \mu$, différente en général de π (cf. [4], p. 41), et ensuite d'établir les liens entre π et μ. Les formules obtenues peuvent aider à éclaircir certaines hypothèses employées dans des problèmes d'estimation concernant des processus de Markov à temps continu (cf., par exemple, [1–3]). Elles permettent aussi la comparaison entre les hypothèses posées lorsqu'on traite l'estimation à la manière de [5] et celles posées lorsqu'on traite l'estimation à la manière de [4].

2. Soit $\mathbf{M}^n = (m_{ij}(n))$.

2.1. *Supposons que les conditions* (1.1), (1.2) *et* (1.3) *sont satisfaites. Alors* $\mu = (\mu_i , i \in X)$, *donnée par*

$$\mu_i = \pi_i q_i \Big/ \sum_{k \in X} \pi_k q_k \tag{2.1}$$

pour tout $i \in X$, *est l'unique probabilité absolue stationnaire du processus de Markov homogène, à temps discret, extrait.*

En effet, puisque $\boldsymbol{\pi}$ est l'unique probabilité telle que $\boldsymbol{\pi}\mathbf{P}(t) = \boldsymbol{\pi}$ pour tout $t \in \mathrm{R}_+$, nous obtenons $\boldsymbol{\pi}\mathbf{Q} = \mathbf{0}$. Il s'en suit que $\sum_{i \neq j} \pi_i q_{ij} = \pi_j q_j$ pour tout $j \in \mathrm{X}$, et par conséquent $\sum_{i \neq j} \left(\pi_i q_i \Big/ \sum_{k \in \mathrm{X}} \pi_k q_k\right) (q_{ij}/q_i) = \pi_j q_j \Big/ \sum_{k \in \mathrm{X}} \pi_k q_k$ pour tout $j \in \mathrm{X}$, ce qui montre que $\mu = (\mu_i\,,\, i \in \mathrm{X})$, où μ_i pour tout $i \in \mathrm{X}$ est donné par (2.1), vérifie $\mu\mathbf{M} = \mu$. Donc l'existence de μ est établi.

La formule (2.1) peut être obtenue aussi d'une autre façon. Désignons par $\tau_t(i\,,\omega)$ la durée totale de séjour de $s \to \xi_s(\omega)$ dans l'état i pendant l'intervalle de temps $[0\,,t]$, et par $\mathrm{N}_t(ij,\omega)$ le nombre de fois que $s \to \xi_s(\omega)$ passe de l'état i à l'état j pendant le même intervalle de temps $[0\,,t]$. Notons aussi $\tau_t(i)$ (respectivement $\mathrm{N}_t(ij)$) à la place de la variable aléatoire réelle $\tau_t(i\,,\cdot)$ (respectivement $\mathrm{N}_t(ij\,,\cdot)$) et $\mathrm{N}_t(i\mathrm{X})$ (respectivement $\mathrm{N}_t(\mathrm{XX})$) à la place de $\sum_{j \in \mathrm{X}}' \mathrm{N}_t(ij)$ $\left(\text{respectivement } \sum_{i,j \in \mathrm{X}}' \mathrm{N}_t(ij)\right)$. D'après [5], p. 133, $\lim_{t \to \infty} \tau_t(i)/t = \pi_i$ p.s. pour tout $i \in \mathrm{X}$ et $\lim_{t \to \infty} \mathrm{N}_t(ij)/t = \pi_i q_{ij}$ p.s. pour tout $i,j \in \mathrm{X}$. Par conséquent $\lim_{t \to \infty} \mathrm{N}_t(\mathrm{XX})/t = \sum_{k \in \mathrm{X}} \pi_k q_k$ p.s. et $\lim_{t \to \infty} \mathrm{N}_t(i\mathrm{X})/\mathrm{N}_t(\mathrm{XX}) = \pi_i q_i \Big/ \sum_{k \in \mathrm{X}} \pi_k q_k$ p.s. pour tout $i \in \mathrm{X}$. D'autre part, d'après [4], p. 48, $\lim_{t \to \infty} \mathrm{N}_t(i\mathrm{X})/\mathrm{N}_t(\mathrm{XX}) = \mu_i$ p.s. pour tout $i \in \mathrm{X}$. Nous obtenons donc de nouveau (2.1).

Prouvons maintenant l'unicité de μ. En vertu de (1.2), il existe une seule classe ergodique $\mathrm{E}_0 = \{i \in \mathrm{X} : \pi_i > 0\}$ pour le processus de Markov homogène, à temps continu, initial. Nous allons montrer que le processus de Markov, à temps discret, extrait, étant de même nature, ne peut avoir lui aussi qu'une seule classe ergodique qui n'est autre que E_0. Pour cela, soit $p_{ij}^{\langle n \rangle}(t) = \Pr\{\xi_{s+t}(\omega) = j \mid \xi_s(\omega) = i$ et le passage de i à j s'effectue en n sauts$\}$. Alors (cf. [7], p. 250)

$$p_{ij}^{\langle 0 \rangle}(t) = \delta_{ij} \exp(-q_i t),$$

$$p_{ij}^{\langle 1 \rangle}(t) = \int_0^t \exp(-q_i t_0 - q_j(t - t_0))\, q_{ij}\, \mathrm{d}t_0,$$

$$p_{ij}^{\langle n \rangle}(t) = \sum_{\substack{k_l \neq i \\ 1 \leq l \leq n-1}} \int \cdots \int_\Lambda \exp\Bigg(-q_i t_0 - q_{k_1} t_1 - \cdots - q_{k_{n-1}} t_{n-1} - $$
$$- q_j\left(t - \sum_{k=0}^{n-1} t_k\right)\Bigg) \times q_{ik_1} q_{k_1 k_2} \cdots q_{k_{n-1} j}\, \mathrm{d}t_0 \cdots \mathrm{d}t_{n-1}, \qquad n \geq 2,$$

où $\Lambda = \left\{(t_0\,,\cdots,t_{n-1}) : t_k \in \mathrm{R}_+\,,\ 1 \leq k \leq n-1\,,\ \sum_{k=0}^{n-1} t_k \leq t\right\}$. De plus, pour tout $i\,,j \in \mathrm{X}$ nous avons $p_{ij}(t) = \sum_{n=0}^{\infty}' p_{ij}^{\langle n \rangle}(t)$, ce qui nous donne

$$p_{ij}(t) = \delta_{ij} \exp(-q_i t) + \tag{2.2}$$
$$+ \sum_{n \in \mathrm{N}^*} \sum_{\substack{k_l \neq i \\ 1 \leq l \leq n-1}} m_{ik_1} m_{k_1 k_2} \cdots m_{k_{n-1} j}\, \mathrm{A}_{ij}^{k_1 \cdots k_{n-1}}(t),$$

où

$$A_{ij}^{k_1\cdots k_{n-1}}(t) = q_i\, q_{k_1}\cdots q_{k_{n-1}} \int\cdots\int_{\Lambda} \exp\Big(-q_i t_0 - q_{k_1} t_1 \cdots - q_{k_{n-1}} t_{n-1} -$$

$$- q_j\Big(t - \sum_{k=0}^{n-1} t_k\Big)\Big)\, \mathrm{d}t_0 \cdots \mathrm{d}t_{n-1},$$

et $A_{ij}^{k_1\cdots k_{n-1}}(t) > 0$ pour tout $t > 0$ en vertu de (1.3). De même, puisque $m_{ii} = 0$, nous avons

$$m_{ij}(n) = \sum_{\substack{k_l \neq i \\ 1 \leq l \leq n-1}} m_{ik_1} m_{k_1 k_2} \cdots m_{k_{n-1} j}.$$

Si $i, j \in X$ sont tels qu'il existe $t > 0$ pour lequel $p_{ij}(t) > 0$, (2.2) montre qu'il existe $n_0 \in N^*$ tel que $m_{ij}(n_0) > 0$. En effet, si $i \neq j$, d'après (2.2) il existe $n \in N^*$ et $k_1, \cdots, k_{n-1} \in X$ avec $k_1 \neq \cdots \neq k_{n-1}$, tel que $m_{ik_1} m_{k_1 k_2} \cdots m_{k_{n-1} j} > 0$. Autrement dit, il existe $n \in N^*$ tel que $m_{ij}(n) > 0$. Si $i = j$, nous nous servons du cas $i \neq j$ en écrivant $m_{ii}(n + n') > m_{ik}(n)\, m_{ki}(n') > 0$, où $k \neq i$ et où n et n' sont respectivement tels que $m_{ik}(n) > 0$ et $m_{kj}(n') > 0$.

Réciproquement d'ailleurs, si $i, j \in X$ sont tels qu'il existe $k_1, \cdots, k_{n-1} \in X$ avec $k_1 \neq \cdots \neq k_{n-1}$, pour lesquels $m_{ik_1} m_{k_1 k_2} \cdots m_{k_{n-1} j} > 0$, c'est-à-dire s'il existe $n \in N^*$ tel que $m_{ij}(n) > 0$, (2.2) montre alors que pour ces états $i, j \in X$ il existe $t > 0$ tel que $p_{ij}(t) > 0$.

Nous déduisons que E_0 qui est l'unique classe ergodique du processus de Markov, à temps continu, initial, est aussi l'unique classe ergodique du processus de Markov, à temps discret, extrait, et par conséquent μ existe et est unique.

2.2. *π s'exprime en fonction de μ par la formule*

$$\pi_i = \Big(\mu_i \prod_{k \neq i} q_k\Big) \Big/ \Big(\sum_{l \in X} \mu_l \prod_{k \neq l} q_k\Big) \tag{2.3}$$

pour tout $i \in X$.

En effet, supposons que $X = \{1, \cdots, r\}$. Alors dans la relation $\sum_{k=1}^{r} \pi_k = 1$ nous remplaçons d'abord π_r par $1 - \sum_{k=1}^{r-1} \pi_k$. D'après (2.1) nous avons $\mu_i \sum_{k=1}^{r} \pi_k q_k - \pi_i q_i = 0$ pour tout $1 \leq i \leq r$, donc

$$\mu_i \Big(\sum_{k=1}^{r-1} \pi_k q_k + \Big(1 - \sum_{k=1}^{r-1} \pi_k\Big) q_r\Big) - \pi_i q_i = 0 \tag{2.4}$$

pour tout $1 \leq i \leq r$. En considérant les π_i, $1 \leq i \leq r-1$, comme des inconnues dans les premières $r-1$ équations linéaires (2.4) (la $r^{\text{ième}}$ n'étant pas indépendante des $r-1$ premières), nous obtenons un système de $r-1$

équations linéaires à $r-1$ inconnues

$$\sum_{\substack{k=1 \\ k\neq i}}^{r-1} \pi_k (q_k - q_r) \mu_i + \pi_i ((q_i - q_r) \mu_i - q_i) = - \mu_i q_r \tag{2.5}$$

pour tout $1 \leq i \leq r-1$. La résolution du système (2.5) nous conduit alors à la formule (2.3).

Examinons maintenant le cas particulier $\pi = \mu$.

2.3. $\pi = \mu$ *si et seulement si* $q_i = \sum_{k \in X} \pi_k q_k$ *pour tous les* $i \in X$ *tels que* $\pi_i \neq 0$. *En particulier, si* $q_i = \sum_{k \in X} \pi_k q_k$ *pour tout* $i \in X$, *alors* $\pi = \mu$.

La condition est suffisante. En effet, si $\pi_i = 0$ (2.1) montre que $\mu_i = 0$; si $\mu_i \neq 0$ (2.1) nous donne $\mu_i = \pi_i$. La condition est nécessaire; puisque $\mu_i = \pi_i$ pour tout $i \in X$, nous avons ou bien $\pi_i = \mu_i = 0$, ou bien $q_i \Big/ \sum_{k \in X} \pi_k q_k = 1$, d'après (2.1).

Citons un exemple. Soit $X = \{1, 2, 3\}$ et

$$\mathbf{P}(t) = \begin{pmatrix} \exp(-t/2) & (1-\exp(-t/2))/2 & (1-\exp(-t/2))/2 \\ 0 & (1+\exp(-2t))/2 & (1-\exp(-2t))/2 \\ 0 & (1-\exp(-2t))/2 & (1+\exp(-2t))/2 \end{pmatrix}.$$

Il est clair que cette matrice de transition est standard. D'autre part,

$$\mathbf{Q} = \begin{pmatrix} -1/2 & 1/4 & 1/4 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$

et $\min_{i \in X} q_i = 1/2 > 0$. Nous avons aussi $\pi = (0 \;\; 1/2 \;\; 1/2)$. Donc les conditions (1.1), (1.2) et (1.3) sont vérifiées. De même,

$$\mathbf{M} = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix};$$

il est clair que μ existe et est unique. Puisque $\pi_2 \neq 0$, que $\pi_3 \neq 0$ et que $q_2 = q_3 = \sum_{k=1}^{3} \pi_k q_k = 1$, **2.3** confirme que $\mu = \pi = (0 \;\; 1/2 \;\; 1/2)$.

3. Nous avons examiné en détail le cas où X est fini. Faisons maintenant quelques remarques rapides sur les autres cas.

Si X est dénombrable, sous des conditions analogues à (1.1) (avec uniformité en plus), (1.2) et (1.3) et supposant que le produit $\prod_{k \in X} q_k$ est convergent, nous obtenons encore les formules (2.1) et (2.3).

Soit ensuite X un borélien quelconque de $R = (-\infty, \infty)$. Supposons que la probabilité de transition $P(t; x, A)$ est permanente discontinue de noyau $q(x, A)$, et posons $q(x) = -q(x, \{x\})$ pour tout $x \in X$. De plus, supposons que $\lim_{t \to \infty} P(t; x, A) = \pi(A)$ pour tout borélien $A \subset X$ et que $\inf_{x \in X} q(x) > 0$. Alors l'analogue de la formule (2.1) est

$$\mu(A) = \int_A d\pi(x)\, q(x) \Big/ \int_X d\pi(x)\, q(x)$$

pour tout borélien $A \subset X$.

BIBLIOGRAPHIE

[1] C. ANDRIEU, *La méthode du quasi-maximum de vraisemblance concernant des processus de Markov à temps continu*, « C. R. Acad. Sci. Paris », *272*, A334–A336 (1971).

[2] C. ANDRIEU, *Existence et convergence des estimateurs du quasi-maximum de vraisemblance pour des processus de Markov finis à temps continu*, « Rev. Roumaine Math. Pures Appl. », *17*, 169–182 (1972).

[3] C. ANDRIEU et CL. LANGRAND, *Estimateurs du quasi-maximum de vraisemblance pour des processus de Markov à temps continu*, « Rev. Roumaine Math. Pures Appl. » (à paraître).

[4] P. BILLINGSLEY, *Statistical inference for Markov processes*, Chicago, The University of Chicago Press (1961).

[5] BUI-TRONG-LIEU, *Estimations pour des processus de Markov*, « Publ. Inst. Statist. Univ. Paris », *11*, 75–188 (1962).

[6] K. L. CHUNG, *Markov chains with stationary transition probabilities*, 2e edition, New York. Springer-Verlag (1967).

[7] J. L. DOOB, *Stochastic processes*, New York, Wiley (1953).