Z. Wahrscheinlichkeitstheorie verw. Gebiete 32, 333–344 (1975)

Une formulation, et la finitude, du test séquentiel du rapport des vraisemblances dans le cas Markovien à temps continu

Colette Andrieu

I. Introduction et préliminaires

I.1. Soit $(\Omega, \mathscr{A}, \Pr)$ un espace probabilisé, soit $(\Theta, \mathscr{C})$ un espace mesurable, soit $\big((X_t^{(\theta)})_{t \geqq 0}\big)_{\theta \in \Theta}$ une famille de fonctions aléatoires markoviennes indicées par $\theta \in \Theta$, dont les probabilités de transition, de ce fait, dépendent elles aussi de θ. Supposons que θ soit la valeur d'une variable aléatoire Y, définie sur $(\Omega, \mathscr{A})$ et à valeur dans $(\Theta, \mathscr{C})$. La loi de Y est notée P_Y. (Dans le cas particulier où $P_Y = \delta_{\theta_0}$, probabilité de Dirac dont la masse est concentrée au point θ_0, on retrouve le cas où les probabilités dépendent d'un paramètre dont la vraie valeur est θ_0.) Le test de l'hypothèse $H_0\,(\theta = \theta_0)$ contre l'hypothèse $H_1\,(\theta = \theta_1)$ pose le problème du choix entre la fonction aléatoire $(X_t^{(\theta_0)})_{t \geqq 0}$ et la fonction aléatoire $(X_t^{(\theta_1)})_{t \geqq 0}$. Mais en pratique, la fonction aléatoire que nous observons et qui nous procure ainsi «l'échantillon» ne permet pas a priori de dire de laquelle des $\big((X_t^{(\theta)})_{t \geqq 0}\big)_{\theta \in \Theta}$ il s'agit, nous nous contenterons de parler de la donnée d'une fonction aléatoire $(X_t)_{t \geqq 0}$ dont les probabilités de transition dépendent d'un paramètre θ. Nous noterons $\Pr_\theta(A)$ à la place de $\Pr(A \mid Y = \theta)$ et nous dirons qu'il s'agit de la probabilité de A lorsque la valeur du paramètre est θ.

I.2. Cet article, qui développe la Note [2], a pour but de présenter une formulation du test séquentiel du rapport des vraisemblances dans le cas markovien à temps continu, de type de sauts et à un nombre quelconque d'états. Il existe déjà des articles traitant de ce genre de test dans le cas de processus à temps continu (cf. [5] et [11]), mais ils ne répondent pas directement à notre préoccupation. Il nous paraît donc nécessaire de souligner ici la motivation de notre article: (en ne mentionnant provisoirement que le cas d'un nombre fini d'états i pour mieux expliquer), ce ne sont pas les probabilités de transition $p_t(i, j; \theta)$ qui sont explicitement données dans le cas des processus à temps continu, mais ce sont les

$$q_i(\theta) = \lim_{t \to 0} \frac{1}{t}\,[1 - p_t(i, i; \theta)] \quad \text{et} \quad q_{ij}(\theta) = \lim_{t \to 0} \frac{1}{t}\, p_t(i, j; \theta)$$

qui sont données (à supposer que quel que soit $\theta \in \Theta$ et quel que soit i, $\lim_{t \to 0} p_t(i, i; \theta) = 1$). Ce ne sont donc pas avec les $p_t(i, j; \theta)$ que nous définissons le rapport des vraisemblances. D'autre part, selon [6], les $\tau_T(i)$ et $N_T(ij)$ ($\tau_T(i)$ et $N_T(ij)$ désignant respectivement la durée totale de séjour dans i et le nombre de passages directs de i à j, durant l'intervalle de temps $[0, T]$) constituent (avec l'état initial x_0) un résumé exhaustif. C'est pourquoi, exprimer le rapport des vraisemblances en fonction des $\tau_T(i)$, $N_T(ij)$, $q_i(\theta)$, $q_{ij}(\theta)$, ... est parfaitement justifié. Mais nous ne nous restreignons pas au seul cas d'un nombre fini d'états; nous étudions directement le cas d'un nombre quelconque d'états, et nous le faisons avec une expression du rapport des vraisemblances différente de celle formulée par [3]. Nous pensons donc que la formulation du test séquentiel telle que nous la concevons ci-dessous n'est pas classique et n'entre pas dans l'optique de [11].

Evidemment un test séquentiel n'a de sens que si sa finitude est prouvée, c'est donc par la preuve de la finitude du test, propriété essentielle, que nous commençons l'étude, et nous la démontrons explicitement dans cet article. Nous pouvons, bien sûr, dans une étape future envisager d'autre étude, comme celle sur les conditions équivalentes du genre données dans [11], ou comme celle sur l'optimalité (mais selon [10] et pourtant dans un cas bien plus simple, les tests optimaux semblent être trop compliqués pour être d'un intérêt pratique).

Enfin, à titre de comparaison, nous faisons quelques remarques sur le cas markovien à temps discret et à un nombre quelconque d'états, cas moins complexe, donc nécessitant des hypothèses plus faibles que celles du cas à temps continu. Citons en passant les articles [9] et [12] traitant du cas à temps discret.

I.3. Le résultat auxiliaire ci-dessous sera utile pour les paragraphes suivants.

Rappelons, pour commencer, qu'étant donnés un espace probabilisé $(\Omega, \mathscr{A}, \Pr)$ et deux sous-tribus $\mathscr{K}_1$ et $\mathscr{K}_2$ de $\mathscr{A}$, le coefficient de dépendance φ, dit d'Ibragimov, de $\mathscr{K}_1$ et $\mathscr{K}_2$ est défini (voir [8]) par

$$\varphi(\mathscr{K}_1, \mathscr{K}_2) = \sup_{\substack{A_1 \in \mathscr{K}_1 : \Pr(A_1) \neq 0 \\ A_2 \in \mathscr{K}_2}} \left|\Pr(A_2 \mid A_1) - \Pr(A_2)\right|.$$

Ce coefficient s'exprime de façon équivalente par

$$\varphi(\mathscr{K}_1, \mathscr{K}_2) = \sup_{A \in \mathscr{K}_2} \left(\operatorname*{sup\,ess}_{\omega \in \Omega} \left|\Pr(A \mid \mathscr{K}_1)(\omega) - \Pr(A)\right|\right).$$

Le résultat auxiliaire signalé s'énonce comme suit:

I.4. **Proposition.** *Soit $(X_t)_{t \geqq 0}$ une fonction aléatoire markovienne définie sur $(\Omega, \mathscr{A}, \Pr)$. Quels que soient les instants $s_1 < \cdots < s_m < s < t < t_1 < \cdots < t_n$, on a*

$$\varphi(\mathscr{A}_{s_1, \ldots, s_m, s}, \mathscr{A}_{t, t_1, \ldots, t_n}) = \varphi(\mathscr{A}_s, \mathscr{A}_t),$$

$\mathscr{A}_{s_1, \ldots, s_m, s}$, $\mathscr{A}_{t, t_1, \ldots, t_n}$, $\mathscr{A}_s$ et $\mathscr{A}_t$ désignant ici les sous-tribus de $\mathscr{A}$ engendrées respectivement par $(X_{s_1}, \ldots, X_{s_m}, X_s)$, $(X_t, X_{t_1}, \ldots, X_{t_n})$, X_s et X_t.

Nous omettons la démonstration de ce résultat, qui figure de façon implicite dans [7], théorème 3, où il est question du coefficient de dépendance dit de Rosenblatt, mais dont la transposition au cas actuel est immédiate.

II. Resultats principaux

II.1. Soit un processus de Markov homogène à temps continu dont l'espace des états est $(\mathscr{X}, \mathscr{B})$, où $\mathscr{X}$ est un borélien de $\mathbb{R}^h$, et $\mathscr{B}$ la tribu borélienne de $\mathscr{X}$, et dont les probabilités de transition P_t, $t \in \mathbb{R}_+$, dépendent d'un paramètre θ,

$\theta \in \Theta$. Nous supposons que: quel que soit $\theta \in \Theta$, $\lim_{t\to 0} P_t(x, \{x\}; \theta) = 1$, uniformément en x. On sait, d'après [4] que, sous cette condition, quel que soit $\theta \in \Theta$,

$$\lim_{t\to 0} \frac{1 - P_t(x, \{x\}; \theta)}{t} = Q(x; \theta) \leqq U(\theta) < \infty$$

cette limite étant atteinte uniformément en x; pour $x \notin B$,

$$\lim_{t\to 0} \frac{1}{t} P_t(x, B; \theta) = Q(x, B; \theta) \leqq Q(x; \theta),$$

cette limite étant atteinte uniformément en $B \subset \mathscr{X} \setminus \{x\}$.

Puisque $Q(x, \mathscr{X} \setminus \{x\}; \theta) = Q(x; \theta)$, en posant $Q(x, \{x\}; \theta) = 0$, nous pouvons écrire $Q(x, \mathscr{X}; \theta) = Q(x; \theta)$. Nous nous plaçons dans l'hypothèse où quel que soit $x \in \mathscr{X}$, la mesure $Q(x, .; \theta)$ est absolument continue par rapport à une mesure σ-finie μ; la version $\bigotimes^2 \mathscr{B}$-mesurable de sa densité de Radon-Nikodym est notée $q(x, y; \theta)$. Nous supposons de plus que, quel que soit $x \in \mathscr{B}$, quel que soit $\theta \in \Theta$,

$$\lim_{t\to\infty} P_t(x, B; \theta) = \pi(B; \theta)$$

et que cette limite est atteinte uniformément en x et en B, et de façon exponentielle, c'est à dire que, pour tout $\theta \in \Theta$, il existe $K(\theta)$ et $\rho(\theta)$, avec $0 < \rho(\theta) < 1$, tels que pour tout $t \in \mathbb{R}_+$,

$$|P_t(x, B; \theta) - \pi(B; \theta)| \leqq K(\theta) \cdot \rho(\theta)^t,$$

et enfin que

$$\operatorname*{sup\,ess}_{(x, y) \in \mathscr{X}^2} \left| \operatorname{Log} \frac{q(x, y; \theta_1)}{q(x, y; \theta_0)} \right| < \infty.$$

Le sup ess étant relatif à la mesure $\pi(dx; \theta)\, Q(x, dy; \theta)$.

Soit maintenant une fonction aléatoire markovienne homogène, séparable $(X_t)_{t \in \mathbb{R}_+}$ définie sur un espace probabilisé $(\Omega, \mathscr{A}, \Pr)$ et à valeur, pour chaque t, dans $(\mathscr{X}, \mathscr{B})$, ayant $P_t(., .; \theta)$ pour probabilités de transition et $\pi(., \theta)$ pour loi initiale (où θ est le paramètre dont la valeur est à soumettre au test de l'hypothèse $H_0(\theta = \theta_0)$ contre l'hypothèse $H_1(\theta = \theta_1)$).

Nous supposons, sans que cela constitue une restriction (cf. [3]), que toutes les trajectoires soient continues à droite. Cela a sa raison d'être dans la procédure séquentielle exposée ci-dessous.

II.2. Relativement à un intervalle de temps $[t', t'']$ et au morceau de trajectoire $(X_t(\omega))_{t' \leqq t \leqq t''}$, nous notons: $\tau_{t', t''}(B; \omega)$ la durée totale de séjour dans B; $N_{t', t''}(B_1 \times B_2; \omega)$ le nombre de passages directs de B_1 à B_2. On sait que, pour $\omega \in \Omega$, $\tau_{t', t''}(.; \omega)$ est une mesure positive sur $\mathscr{B}$ et $N_{t', t''}(.; \omega)$ est une mesure positive sur $\bigotimes^2 \mathscr{B}$ à valeur entière, tandis que pour $B \in \mathscr{B}$, $B_1 \times B_2 \in \bigotimes^2 \mathscr{B}$, $\tau_{t', t''}(B; .)$ et $N_{t', t''}(B_1 \times B_2; .)$ sont des variables réelles que nous notons, pour simplifier, $\tau_{t', t''}(B)$ et $N_{t', t''}(B_1 \times B_2)$. Dans le cas particulier où $[t', t''] = [0, T]$, nous notons N_T à la place de $N_{0, T}$ et τ_T à la place de $\tau_{0, T}$; ainsi, $N_T(\mathscr{X} \times \mathscr{X})$ par exemple, désignera le nombre total de sauts pendant l'intervalle $[0, T]$.

Posons
$$\xi_T(\omega) = -\int_{\mathscr{X}} [Q(x;\theta_1) - Q(x;\theta_0)]\, d\tau_T(x;\omega) + \int_{\mathscr{X}^2} \operatorname{Log} \frac{q(x,y;\theta_1)}{q(x,y;\theta_0)}\, dN_T(x,y;\omega)$$

(ou encore, $\xi_T(\omega) = L_T(\theta_1;\omega) - L_T(\theta_0;\omega)$ d'après la notation de la Note [1]). Ce ξ_T joue le rôle du logarithme du rapport des vraisemblances.

Nous rappelons à ce propos (cf. [1]) que $L_T(\theta;\omega)$ constitue une extension de la notion de vraisemblance, car la vraisemblance proprement dite n'est pas définie pour le cas des processus de Markov à temps continu. L'introduction de $L_T(\theta;\omega)$ est motivée par le fait qu'avec $\theta \in \mathbb{R}$, l'équation dite de vraisemblance s'écrit
$$-\int_{\mathscr{X}} \frac{\partial}{\partial\theta} Q(x;\theta)\, d\tau_T(x;\omega) + \int_{\mathscr{X}^2} \frac{\partial}{\partial\theta} \operatorname{Log} q(x,y;\theta)\, dN_T(x,y;\omega) = 0.$$

II.3. **Formulation de la procédure séquentielle.** Nous proposons la procédure suivante:

Soient les nombres a et b tels que $0 < b < 1 < a$.

(i) Si $\xi_T(\omega) \leqq \operatorname{Log} b$, accepter l'hypothèse H_0.

(ii) Si $\xi_T(\omega) \geqq \operatorname{Log} a$, accepter l'hypothèse H_1.

(iii) Si $\operatorname{Log} b < \xi_T(\omega) < \operatorname{Log} a$, continuer l'observation.

Cette dernière situation (iii) nécessite une précision. En effet, si dans le cas du temps discret, «continuer l'observation» signifie poursuivre l'observation une unité de temps de plus puis procéder à la vérification pour voir laquelle des 3 situations se présente, le cas actuel du temps continu est plus complexe. Il est clair qu'on ne peut pas procéder à la vérification à n'importe quel instant. La question est donc: jusqu'à quel instant continue-t-on l'observation avant de procéder à une nouvelle vérification? Nous proposons, lorsqu'on se trouve dans la situation (iii) à l'instant T, *de continuer l'observation jusqu'au premier instant T' tel que $N_{T'}(\mathscr{X}\times\mathscr{X};\omega) = N_T(\mathscr{X}\times\mathscr{X};\omega)+1$*, c'est à dire jusqu'à ce que l'on rencontre un nouveau saut. On poursuivra ce procédé jusqu'à ce que (i) ou (ii) se produise. L'instant $S(\omega)$ où l'on arrête définitivement l'observation est celui où l'on a: ou bien $\xi_{S(\omega)} \leqq \operatorname{Log} b$, ou bien $\xi_{S(\omega)} \geqq \operatorname{Log} a$. $S(\omega) = +\infty$ si et seulement si quel que soit $T \in \mathbb{R}_+$,
$$\operatorname{Log} b < \xi_T(\omega) < \operatorname{Log} a.$$

II.4. Pour la commodité de l'exposé, posons
$$C(\theta) = \int_{\mathscr{X}^2} \left[\operatorname{Log} \frac{q(x,y;\theta_1)}{q(x,y;\theta_0)}\right]^2 \pi(dx;\theta)\, Q(x,dy;\theta) + \int_0^\infty G(t;\theta)\, dt$$

où
$$\tfrac{1}{2} G(t;\theta) = \int_{\mathscr{X}^2} [Q(x;\theta_1) - Q(x;\theta_0)]\,[Q(u;\theta_1) - Q(u;\theta_0)]\, \pi(dx;\theta) \cdot [P_t(x,du;\theta) - \pi(du;\theta)]$$
$$- \int_{\mathscr{X}^3} \operatorname{Log} \frac{q(x,y;\theta_1)}{q(x,y;\theta_0)} [Q(u;\theta_1) - Q(u;\theta_0)]\, \pi(dx;\theta) \cdot Q(x,dy;\theta)\,[P_t(y,du;\theta) - \pi(du;\theta)]$$

$$- \int_{\mathscr{X}^3} [Q(u;\theta_1)-Q(u;\theta_0)] \operatorname{Log} \frac{q(x,y;\theta_1)}{q(x,y;\theta_0)} \pi(du;\theta)$$
$$\cdot [P_t(u,dx;\theta)-\pi(dx;\theta)]\, Q(x,dy;\theta)$$
$$+ \int_{\mathscr{X}^4} \operatorname{Log} \frac{q(x,y;\theta_1)}{q(x,y;\theta_0)} \operatorname{Log} \frac{q(u,v;\theta_1)}{q(u,v;\theta_0)} \pi(dx;\theta)\, Q(x,dy;\theta)$$
$$\cdot [P_t(y,du;\theta)-\pi(du;\theta)]\, Q(u,dv;\theta).$$

Nous faisons remarquer que l'intégrale $\int_0^\infty G(t;\theta)\,dt$ est absolument convergente, car à l'aide d'une décomposition de Hahn de $P_t(x,.;\theta)-\pi(.;\theta)$, on a la majoration

$$|G(t;\theta)| \leqq W(\theta)\,\rho(\theta)^t,$$

où $W(\theta)$ est exprimable en fonction de $K(\theta)$, du

$$\sup_{(x,y)\in\mathscr{X}^2} \operatorname{ess} \left| \operatorname{Log} \frac{q(x,y;\theta_1)}{q(x,y;\theta_0)} \right|$$

relatif à la mesure $\pi(dx;\theta)\,Q(x,dy;\theta)$ et de la borne supérieure de

$$|Q(x;\theta_1)-Q(x;\theta_0)|.$$

Nous appelons *cas dégénéré*, le cas où

(II.4.1) $$C(\theta)=0$$

et $$\int_{\mathscr{X}} [Q(x;\theta_1)-Q(x;\theta_0)]\, \pi(dx;\theta)$$

(II.4.2) $$= \int_{\mathscr{X}^2} \operatorname{Log} \frac{q(x,y;\theta_1)}{q(x,y;\theta_0)} \pi(dx;\theta)\, Q(x,dy;\theta).$$

La signification de ce coefficient $C(\theta)$, ainsi que le sens du cas dégénéré, seront commentés plus loin.

Nous avons le résultat auxiliaire suivant, valable dans le cas non dégénéré:

II.5. **Proposition.** *Il existe un* $c>0$ *pour lequel* $\Pr_\theta[|\xi_c|<\gamma]<1$, *où*

$$\gamma = \operatorname{Log} a + |\operatorname{Log} b|.$$

Démonstration. En utilisant les résultats suivants, obtenus par calculs par discrétisation de l'axe du temps: quels que soient $B_1\in\mathscr{B}$, $B_2\in\mathscr{B}$, $B_3\in\mathscr{B}$, $B_4\in\mathscr{B}$,

$$E_\theta[\tau_T(B_1)\cdot\tau_T(B_2)]$$
$$= \int_0^T \Big[\int_{B_1} \pi(dx;\theta)\, P_t(x,B_2;\theta) + \int_{B_2} \pi(dx,\theta)\, P_t(x,B_1;\theta)\Big](T-t)\,dt,$$

$$E_\theta[N_T(B_1\times B_2)\cdot N_T(B_3\times B_4)]$$
$$= \int_{B_1\cap B_3} \pi(dx;\theta)\, Q(x,B_2\cap B_4;\theta)$$
$$+ \int_0^T \Big[\int_{B_1} \pi(dx;\theta) \int_{B_2} Q(x,dy;\theta) \int_{B_3} P_t(x,du;\theta)\, Q(u,B_4;\theta)$$
$$+ \int_{B_3} \pi(dx;\theta) \int_{B_4} Q(x,dy;\theta) \int_{B_1} P_t(y,du;\theta)\, Q(u,B_2;\theta)\Big](T-t)\,dt$$

et

$$\begin{aligned}E_\theta[\tau_T(B_1)\cdot N_T(B_2\times B_3)] \\ &= \int_0^T \Big[\int_{B_1} \pi(dx;\theta)\int_{B_2} P_t(x,du;\theta)\,Q(u,B_3;\theta) \\ &\quad + \int_{B_2}\pi(dx;\theta)\int_{B_3} Q(x,dy;\theta)\,P_t(y,B_1;\theta)\Big](T-t)\,dt,\end{aligned}$$

nous obtenons

$$\begin{aligned}E_\theta(\xi_T^2) = {} & T\int_{\mathscr{X}^2}\left[\operatorname{Log}\frac{q(x,y;\theta_1)}{q(x,y;\theta_0)}\right]^2 \pi(dx;\theta)\,Q(x,dy;\theta) \\ & +2\int_0^T\Big[\int_{\mathscr{X}^2}[Q(x;\theta_1)-Q(x;\theta_0)]\,[Q(u;\theta_1)-Q(u;\theta_0)]\,\pi(dx;\theta)\,P_t(x,du;\theta) \\ & -\int_{\mathscr{X}^3}\operatorname{Log}\frac{q(x,y;\theta_1)}{q(x,y;\theta_0)}[Q(u;\theta_1)-Q(u;\theta_0)]\,\pi(dx;\theta)Q(x,dy;\theta)\,P_t(y,du;\theta) \\ & -\int_{\mathscr{X}^3}[Q(u;\theta_1)-Q(u;\theta_0)]\operatorname{Log}\frac{q(x,y;\theta_1)}{q(x,y;\theta_0)}\pi(dx;\theta)P_t(u,dx;\theta)\,Q(x,dy;\theta) \\ & +\int_{\mathscr{X}^4}\operatorname{Log}\frac{q(x,y;\theta_1)}{q(x,y;\theta_0)}\operatorname{Log}\frac{q(u,v;\theta_1)}{q(u,v;\theta_0)}\pi(dx;\theta)\,Q(x,dy;\theta) \\ & \cdot P_t(y,du;\theta)\,Q(u,dv;\theta)\Big](T-t)\,dt.\end{aligned}$$

On peut vérifier que, pour T grand, $E_\theta(\xi_T^2)$ se comporte comme

$$\begin{aligned}T^2\Big[-\int_{\mathscr{X}}[Q(x;\theta_1)-Q(x;\theta_0)]\,\pi(dx;\theta)+\int_{\mathscr{X}^2}\operatorname{Log}\frac{q(x,y;\theta_1)}{q(x,y;\theta_0)}\pi(dx;\theta)\,Q(x,dy;\theta)\Big]^2 \\ +T[C(\theta)+o(1)].\end{aligned}$$

En effet, d'une part, l'espérance mathématique de ξ_T vaut

$$\begin{aligned}\frac{1}{T}E_\theta(\xi_T) = {} & -\int_{\mathscr{X}}[Q(x;\theta_1)-Q(x,\theta_0)]\,\pi(dx;\theta) \\ & +\int_{\mathscr{X}^2}\operatorname{Log}\frac{q(x,y;\theta_1)}{q(x,y;\theta_0)}\pi(dx;\theta)\,Q(x,dy;\theta)\end{aligned}$$

et d'autre part, la variance de ξ_T s'écrit

$$\sigma_\theta^2(\xi_T) = TC(\theta) - T\int_T^\infty G(t;\theta)\,dt - \int_0^T t\,G(t;\theta)\,dt$$

où on peut constater que

$$\left|\int_0^T t\,G(t;\theta)\,dt\right|$$

reste bornée quel que soit T.

Ainsi donc, dans le cas non dégénéré, $E_\theta(\xi_T^2)\to\infty$ quand $T\to\infty$. Par conséquent, il existe un c tel que

$$\Pr_\theta[\xi_c^2<\gamma^2]<1.$$

Sinon, on aurait, quel que soit $T \in \mathbb{R}_+$, $\Pr_\theta[\xi_T^2 < \gamma^2] = 1$, et on aboutirait à $E_\theta(\xi_T^2) < \gamma^2$, c'est à dire à une contradiction.

II.6. Dans le cas des variables aléatoires indépendantes, il n'existe pas d'analogue au terme $o(1)$, ni d'analogue au terme $\int_0^\infty G(t;\theta)\,dt$ figurant dans l'expression de $C(\theta)$; c'est pourquoi la non dégénérescence γ est simple à exprimer.

Par contre dans le cas markovien actuel, on peut remarquer que $C(\theta) \neq 0$ n'est qu'une condition suffisante pour que $E_\theta(\xi_T^2) \to \infty$ quand $T \to \infty$. Il se peut que, même avec $E_\theta(\xi_T) = 0$ et $C(\theta) = 0$, $E_\theta(\xi_T^2)$ puisse tendre vers l'infini quand $T \to \infty$ (tel est le cas si $T \int_T^\infty G(t;\theta)\,dt \to \infty$ quand $T \to \infty$); mais c'est là une «concession» que nous devons faire à la dépendance markovienne (faible concession, à notre avis, car la nullité de $C(\theta)$ nous paraît exceptionnelle).

II.7. La proposition II.5 ayant montré l'existence d'un c pour lequel

$$\Pr_\theta[\xi_c^2 < \gamma^2] < 1,$$

considérons T assez grand. Soient $d(T)$ et $M(T)$ des entiers, fonctions de T, tendant vers l'infini avec T, choisis de sorte que $M(T)$ soit le plus grand entier tel que

$$M(T)\,(c + d(T)) \leqq T$$

et que

$$\lim_{T \to \infty} M(T)\,\rho(\theta)^{d(T)} = 0.$$

Cette dernière condition est analogue à celle dans Doob [4], p. 230, à propos de la démonstration d'un théorème limite centrale pour des processus de Markov à temps discret. On peut par exemple choisir $d(T)$ comme le plus grand entier positif tel que $d(T)^2 \leqq T$, et alors pour T grand

$$M(T) = \left[\frac{T}{c + d(T)}\right] \sim d(T).$$

Partageons maintenant l'intervalle $[0, T]$ en sous-intervalles disjoints de longueurs alternativement égales à c et à $d(T)$, c'est à dire en intervalles

$$\big[m(c+d(T)),\, m(c+d(T)) + c\big] \quad \text{et} \quad \big]m(c+d(T)) + c,\, (m+1)(c+d(T))\big[,$$

$m = 0, 1, \ldots, M(T) - 1$, et un intervalle résiduel $\big[M(T)(c+d(T)),\, T\big]$.

Posons maintenant, pour simplifier l'écriture:

$$\begin{aligned} {}^m\xi_c(\omega) = &- \int_{\mathfrak{X}} [Q(x;\theta_1) - Q(x;\theta_0)]\, d\tau_{t',t''}(x;\omega) \\ &+ \int_{\mathfrak{X}^2} \operatorname{Log} \frac{q(x,y;\theta_1)}{q(x,y;\theta_0)}\, dN_{t',t''}(x,y;\omega) \end{aligned}$$

lorsque $t' = m(c + d(T))$ et $t'' = m(c + d(T)) + c$.

Les variables aléatoires ${}^m\xi_c$, $m = 0, 1, \ldots, M(T) - 1$, sont de même loi que ξ_c (à cause de la stationnarité de $(X_t)_{t \in \mathbb{R}_+}$) et par conséquent $E_\theta({}^m\xi_c^2) = E_\theta(\xi_c^2)$.

II.8. **Proposition.**

$$\begin{aligned} &\left| \Pr_\theta\left[\bigcap_{m=0}^{M(T)-1} \{{}^m\xi_c^2 < \gamma^2\}\right] - \prod_{m=0}^{M(T)-1} \Pr_\theta[{}^m\xi_c^2 < \gamma^2] \right| \\ &\leqq \Pr_\theta[\xi_c^2 < \gamma^2] \cdot K(\theta) \cdot [M(T) - 1] \cdot \rho(\theta)^{d(T)}. \end{aligned}$$

Démonstration. Posons $D_m = \{{}^m\xi_c^2 < \gamma^2\}$, $m = 0, 1, \ldots, M(T)-1$. Si l'un des D_m est tel que $\Pr_\theta(D_m) = 0$, l'inégalité est vraie. Plaçons-nous dans le cas où aucun des D_m n'est nul. Nous avons

$$\begin{aligned}
&\left|\Pr_\theta\left(\bigcap_{m=0}^{M(T)-1} D_m\right) - \prod_{m=0}^{M(T)-1} \Pr_\theta(D_m)\right| \\
&\quad \leqq \left|\Pr_\theta\left(\bigcap_{m=0}^{M(T)-1} D_m\right) - \Pr_\theta(D_0)\cdot\Pr_\theta\left(\bigcap_{m=1}^{M(T)-1} D_m\right)\right| \\
&\qquad + \sum_{j=0}^{M(T)-3}\left|\left(\prod_{m=0}^{j}\Pr_\theta(D_m)\right)\cdot\Pr_\theta\left(\bigcap_{k=j+1}^{M(T)-1} D_k\right) - \prod_{m=0}^{j+1}\Pr_\theta(D_m)\cdot\Pr_\theta\left(\bigcap_{k=j+2}^{M(T)-1} D_k\right)\right| \\
&\quad \leqq \sum_{j=0}^{M(T)-2}\left|\Pr_\theta\left(\bigcap_{k=j}^{M(T)-1} D_k\right) - \Pr_\theta(D_j)\cdot\Pr_\theta\left(\bigcap_{k=j+1}^{M(T)-1} D_k\right)\right| \\
&\quad = \sum_{j=0}^{M(T)-2}\Pr_\theta(D_j)\cdot\left|\Pr_\theta\left(\bigcap_{k=j+1}^{M(T)-1} D_k \,\middle|\, D_j\right) - \Pr_\theta\left(\bigcap_{k=j+1}^{M(T)-1} D_k\right)\right| \\
&\quad \leqq \Pr_\theta(D_0)\cdot\sum_{j=0}^{M(T)-2}\varphi\left({}^j\mathcal{K}_c, \bigvee_{k=j+1}^{M(T)-1} {}^k\mathcal{K}_c\right),
\end{aligned}$$

où ${}^j\mathcal{K}_c$ désigne la tribu engendrée par $(X_t)_{j(c+d(T))\leqq t\leqq j(c+d(T))+c}$ et où $\bigvee_{k=j+1}^{M(T)-1} {}^k\mathcal{K}_c$ désigne la tribu engendrée par $\bigcup_{k=j+1}^{M(T)-1} {}^k\mathcal{K}_c$. Par conséquent, en utilisant la proposition I.4, nous pouvons écrire

$$\begin{aligned}
\left|\Pr_\theta\left(\bigcap_{m=0}^{M(T)-1} D_m\right) - \prod_{m=0}^{M(T)-1}\Pr_\theta(D_m)\right| &\leqq \Pr_\theta(D_0)\cdot\sum_{j=0}^{M(T)-2}\varphi(\mathcal{A}_{j(c+d(T))+c}, \mathcal{A}_{(j+1)(c+d(T))}) \\
&= \Pr_\theta(D_0)\cdot[M(T)-1]\cdot K(\theta)\cdot\rho(\theta)^{d(T)}.
\end{aligned}$$

Ce résultat et les deux précédents nous conduisent au résultat suivant, lorsque l'on ne se trouve pas dans le cas dégénéré:

II.9. **Proposition.** $\Pr_\theta[S < \infty] = 1$ *et* $\Pr_\theta[N_S(\mathcal{X}\times\mathcal{X}) < \infty] = 1$.

Démonstration. Nous remarquons pour commencer, que si $E_\theta(\xi_T) \neq 0$, c'est à dire si

$$-T\int_{\mathcal{X}}[Q(x;\theta_1) - Q(x;\theta_0)]\,\pi(dx;\theta) + T\int_{\mathcal{X}^2}\operatorname{Log}\frac{q(x,y;\theta_1)}{q(x,y;\theta_0)}\,\pi(dx;\theta)\,Q(x,dy;\theta)$$

n'est pas nulle, la démonstration de la finitude de S découle de ce que $\frac{1}{T}\xi_T$ converge $\Pr_\theta$-presque-surement vers

$$-\int_{\mathcal{X}}[Q(x;\theta_1) - Q(x;\theta_0)]\,\pi(dx;\theta) + \int_{\mathcal{X}^2}\operatorname{Log}\frac{q(x,y;\theta_1)}{q(x,y;\theta_0)}\,\pi(dx;\theta)\,Q(x,dy;\theta),$$

lorsque $T\to\infty$. (Cette convergence presque-sûre, qui n'est pas évidente, est dûe à celle de $\frac{1}{T}\tau_T(B)$ et à celle de $\frac{1}{T}N_T(B_1\times B_2)$ quand $T\to\infty$.)

En incluant la possibilité que $E_\theta(\xi_T)$ prenne la valeur 0, la démonstration suivante est donnée:

Si quel que soit $T\in\mathbb{R}_+$,

$$\operatorname{Log} b < \xi_T(\omega) < \operatorname{Log} a,$$

alors, en écrivant pour tout $m=0,1,\ldots,M(T)-1$,

$$\operatorname{Log} b<\xi_{m(c+d(T))+c}(\omega)<\operatorname{Log} a,$$
$$\operatorname{Log} b<\xi_{m(c+d(T))}(\omega)<\operatorname{Log} a,$$

nous obtenons par soustraction

$$-\gamma<{}^m\xi_c(\omega)<\gamma$$

où $\gamma=\operatorname{Log} a+|\operatorname{Log} b|$.

Par conséquent, en nous servant de II.8, nous avons

$$\begin{aligned}\operatorname{Pr}_\theta[S=+\infty] &\leqq \lim_{T\to\infty} \operatorname{Pr}_\theta\left[\bigcap_{m=0}^{M(T)-1}\{{}^m\xi_c^2<\gamma^2\}\right]\\ &\leqq \lim_{T\to\infty}(\operatorname{Pr}_\theta[\xi_c^2<\gamma^2])^{M(T)}\\ &\quad +K(\theta)\cdot \operatorname{Pr}_\theta[\xi_c^2<\gamma^2]\cdot \lim_{T\to\infty}[M(T)-1]\cdot\rho(\theta)^{d(T)}\\ &=0.\end{aligned}$$

Par conséquent, $\operatorname{Pr}_\theta[S<\infty]=1$. Et comme dans tout intervalle fini, il n'y a qu'un nombre fini de sauts, nous avons

$$\operatorname{Pr}_\theta[N_S(\mathscr{X}\times\mathscr{X})<\infty]=1.$$

II.10. Dans le cas particulier d'un nombre fini d'états, par exemple, le logarithme du rapport des vraisemblances s'écrit

$$\xi_T(\omega)=-\sum_{i\in\mathscr{X}}[q_i(\theta_1)-q_i(\theta_0)]\,\tau_T(i;\omega)+\sum_{(i,j)\in\mathscr{X}^2}\operatorname{Log}\frac{q_{ij}(\theta_1)}{q_{ij}(\theta_0)}\,N_T(ij;\omega).$$

Et lorsque quels que soient $i\in\mathscr{X}$ et $j\in\mathscr{X}$, $q_{ij}(\theta_1)=q_{ij}(\theta_0)$ (ce qui ne permet pas de distinguer l'hypothèse H_1 de l'hypothèse H_0, car on a alors aussi $q_i(\theta_1)=q_i(\theta_0)$), on se trouve alors dans un sous-cas du cas dégénéré.

III. Remarques sur le cas du temps discret

III.1. Nous voudrions, à titre de comparaison, terminer par quelques remarques sur le cas markovien à temps discret, mais à un nombre quelconque d'états. Par sa nature moins complexe, il exige des conditions moins restrictives que celles du cas à temps continu. Soit donc un processus de Markov homogène à temps discret dont l'espace des états est un espace mesurable $(\mathscr{X},\mathscr{B})$ et dont la probabilité de transition P dépend d'un paramètre θ, $\theta\in\Theta$. Nous supposons que, quel que soit $\theta\in\Theta$ et quel que soit $x\in\mathscr{X}$, $P(x,.;\theta)$ soit absolument continue par rapport à une mesure σ-finie et nous notons $p(x,y;\theta)$ sa densité de Radon-Nikodym. Nous supposons de plus, comme dans le paragraphe II, que quel que soit $x\in\mathscr{X}$, quel que soit $B\in\mathscr{B}$ et quel que soit $\theta\in\Theta$, $\lim_{n\to\infty}P_n(x,B;\theta)=\pi(B;\theta)$ et que cette limite soit atteinte uniformément en x et en B et de façon exponentielle, c'est à dire que, pour tout $\theta\in\Theta$, il existe $K(\theta)$ et $\rho(\theta)$ avec $0<\rho(\theta)<1$ tels que, pour tout $n\in\mathbb{N}^*$,

$$|P_n(x,B;\theta)-\pi(B;\theta)|\leqq K(\theta)\cdot\rho(\theta)^n.$$

Nous supposons, enfin, que quel que soit $\theta \in \Theta$,

$$\int_{\mathscr{X}^2} \left[\operatorname{Log} \frac{p(x, y; \theta_1)}{p(x, y, \theta_0)}\right]^2 \pi(dx; \theta)\, P(x, dy; \theta) < \infty.$$

III.2. La procédure du test séquentiel s'exprime, ici, à l'aide de la donnée d'une fonction aléatoire markovienne $(X_t)_{t \in \mathbb{N}}$ définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathrm{Pr})$ et à valeur dans $(\mathscr{X}, \mathscr{B})$, ayant $P(.,.; \theta)$ pour probabilité de transition, et $\pi(.; \theta)$ pour loi initiale, et où θ est justement le paramètre dont la valeur est à tester.

Le rapport des vraisemblances, (en négligeant un premier terme) peut être exprimé sous l'une des deux formes suivantes:

$$\zeta_n(\omega) = \int_{\mathscr{X}^2} \operatorname{Log} \frac{p(x, y; \theta_1)}{p(x, y; \theta_0)}\, dN_n(x, y; \omega)$$

$N_n(B_1 \times B_2; \omega)$ désignant le nombre de passages directs de B_1 à B_2 du temps 0 au temps n, ou encore:

$$\zeta_n(\omega) = \sum_{k=0}^{n-1} \operatorname{Log} \frac{p(X_k(\omega), X_{k+1}(\omega); \theta_1)}{p(X_k(\omega), X_{k+1}(\omega); \theta_0)}.$$

C'est cette deuxième forme que nous utiliserons de préférence.

La procédure séquentielle s'exprime de la façon suivante: Soient les nombres a et b tels que $0 < b < 1 < a$,

(i) *Si* $\zeta_n(\omega) \leqq \operatorname{Log} b$, accepter l'hypothèse H_0.

(ii) *Si* $\zeta_n(\omega) \geqq \operatorname{Log} a$, accepter l'hypothèse H_1.

(iii) *Si* $\operatorname{Log} b < \zeta_n(\omega) < \operatorname{Log} a$, continuer l'observation une unité de temps de plus, avant de procéder à une vérification pour savoir laquelle des situations se présente.

L'instant $\nu(\omega)$ où l'on arrête définitivement l'observation est celui où l'on a:

$$\text{ou bien } \zeta_{\nu(\omega)}(\omega) \leqq \operatorname{Log} b, \qquad \text{ou bien } \zeta_{\nu(\omega)}(\omega) \geqq \operatorname{Log} a.$$

$\nu(\omega) = +\infty$ si et seulement si quel que soit $n \in \mathbb{N}^*$,

$$\operatorname{Log} b < \zeta_n(\omega) < \operatorname{Log} a.$$

III.3. Posons

$$\begin{aligned} \tilde{G}_k(\theta) = 2 \int_{\mathscr{X}^4} \operatorname{Log} \frac{p(x, y; \theta_1)}{p(x, y; \theta_0)} \operatorname{Log} \frac{p(u, v; \theta_1)}{p(u, v; \theta_0)}\, \pi(dx; \theta)\, P(x, dy; \theta) \\ \cdot [P_k(y, du; \theta) - \pi(du; \theta)]\, P(u, dv; \theta). \end{aligned}$$

La série $\sum_{k=0}^{\infty} \tilde{G}_k(\theta)$ est une série absolument convergente.

En effet, en posant $V_i = \operatorname{Log} \dfrac{p(X_i, X_{i+1}; \theta_1)}{p(X_i, X_{i+1}; \theta_0)}$ la stationnarité nous permet d'écrire

$$\tfrac{1}{2} \tilde{G}_k(\theta) = E(V_0 V_k) - E(V_0) \cdot E(V_k).$$

Le lemme 1.1.7, p. 10 de [8] nous permet de majorer $|\frac{1}{2}\tilde{G}_k(\theta)|$ sous une condition moins restrictive que celle imposée dans le cas du temps continu. Nous avons

$$\begin{aligned}|\tfrac{1}{2}\tilde{G}_k(\theta)| &\leqq 2E(V_0^2)\,\varphi^{1/2}(\mathscr{A}_{0,1},\mathscr{A}_{k,k+1})\\ &= 2E(V_0^2)\,\varphi^{1/2}(\mathscr{A}_1,\mathscr{A}_k), \qquad \text{d'après I.3},\\ &\leqq 2E(V_0^2)\cdot K(\theta)\cdot\rho(\theta)^{k-1}.\end{aligned}$$

Par conséquent,

$$\sum_{k=1}^{\infty}|\tilde{G}_k(\theta)| \leqq 4E(V_0^2)\cdot K(\theta)\cdot\sum_{k=1}^{\infty}\rho(\theta)^k<\infty.$$

III.4. Posons maintenant

$$\begin{aligned}\tilde{C}(\theta) = &\int_{\mathscr{X}^2}\left[\operatorname{Log}\frac{p(x,y;\theta_1)}{p(x,y;\theta_0)}\right]^2\pi(dx;\theta)\,P(x,dy;\theta)\\ &-\left[\int_{\mathscr{X}^2}\operatorname{Log}\frac{p(x,y;\theta_1)}{p(x,y;\theta_0)}\pi(dx;\theta)\,P(x,dy;\theta)\right]^2\\ &+\sum_{k=0}^{\infty}\tilde{G}_k(\theta).\end{aligned}$$

Nous appelons cas dégénéré, le cas où

$$\tilde{C}(\theta)=0$$

et

$$\int_{\mathscr{X}^2}\operatorname{Log}\frac{p(x,y;\theta_1)}{p(x,y;\theta_0)}\pi(dx;\theta)\,P(x,dy;\theta)=0.$$

La finitude du test séquentiel s'énonce comme suit, valable dans le cas non dégénéré:

III.5. **Proposition.** $\Pr_\theta[\nu<\infty]=1$.

La démonstration s'effectue de façon analogue à celle du cas du temps continu, mais plus simplement.

1°. Un calcul (voir [8], proposition 1.1.20, p. 23) donne:

$$E_\theta(\zeta_n^2)=n^2\left[\int_{\mathscr{X}^2}\operatorname{Log}\frac{p(x,y;\theta_1)}{p(x,y;\theta_0)}\pi(dx;\theta)\,P(x,dy;\theta)\right]^2+n[\tilde{C}(\theta)+o(1)].$$

Ce qui entraîne l'existence d'un c pour lequel

$$\Pr_\theta[\zeta_c^2<\gamma^2]<1.$$

2°. Nous procédons comme dans le cas du temps continu. Nous nous intéressons aux sous-intervalles disjoints

$$\big[m(c+d(n)),\,m(c+d(n))+c-1\big],\qquad m=0,1,\ldots,M(n)-1$$

de l'intervalle $[0,n]$.

Nous posons

$${}^m\zeta_c(\omega)=\sum_{k=m(c+d(n))}^{m(c+d(n))+c-2}\operatorname{Log}\frac{p(X_k(\omega),X_{k+1}(\omega);\theta_1)}{p(X_k(\omega),X_{k+1}(\omega);\theta_0)}.$$

En nous servant du résultat I.3, nous obtenons l'inégalité

$$\left|\Pr_\theta\left[\bigcap_{m=0}^{M(n)-1}\{{}^m\zeta_c^2<\gamma^2\}\right]-\prod_{m=0}^{M(n)-1}\Pr_\theta[{}^m\zeta_c^2<\gamma^2]\right|$$
$$\leqq \Pr_\theta[\zeta_c^2<\gamma^2]\cdot K(\theta)\cdot[M(n)-1]\cdot\rho(\theta)^{d(n)}.$$

3°. La démonstration s'achève comme dans le cas du temps continu, avec

$$\begin{aligned}\Pr_\theta[\nu=\infty] &\leqq \lim_{n\to\infty}(\Pr_\theta[\zeta_c^2<\gamma^2])^{M(n)}\\ &\quad + K(\theta)\cdot\Pr_\theta[\zeta_c^2<\gamma^2]\cdot\lim_{n\to\infty}[M(n)-1]\cdot\rho(\theta)^{d(n)}\\ &=0.\end{aligned}$$

Remerciements. Je tiens à exprimer ici toute ma gratitude à Monsieur le Professeur M. Iosifescu pour les conseils précieux qu'il m'a donnés au cours de la rédaction de cet article.

References

1. Andrieu, C.: La méthode du quasi-maximum de vraisemblance concernant des processus de Markov à temps continu. C. R. Acad. Sci. Paris, Ser. A, **272**, 334–336 (1971)
2. Andrieu, C.: Une formulation d'un test séquentiel dans le cas markovien à temps continu. C. R. Acad. Sci. Paris, Ser. A, **274**, 148–151 (1974)
3. Billingsley, P.: Statistical inference for Markov processes. University of Chicago Press 1961
4. Doob, J. L.: Stochastic processes. New York: Wiley 1953
5. Dvoretzky, A., Kiefer, J., Wolfowitz, J.: Sequential decision problems for processes with continuous time parameter. Testing Hypotheses. Ann. Math. Statist. **24**, 254–264 (1953)
6. Fortet, R.: Résumés exhaustifs pour un processus de Markov. C. R. Acad. Sci. Paris **247**, 28–29 (1958)
7. Iosifescu, M.: On finite tail σ-algebras. Z. Wahrscheinlichkeitstheorie verw. Gebiete **24**, 159–166 (1972)
8. Iosifescu, M., Theodorescu, R.: Random processes and learning. Berlin-Heidelberg-New York: Springer 1969
9. Kraft, Ch.: Some conditions for consistency and uniform consistency and statistical procedures. Univ. Calif. Publ. Statist. **2**, 125–142 (1955)
10. Schmitz, N.: Likelihoodquotienten-Sequenztests bei homogenen Markov'schen Ketten. Biometrische Z. **10**, 231–347 (1968), (10 bis) Math. Reviews **44**, #6112
11. Schmitz, N.: Existenz von sequentiellen Tests zu vorgegebenen Niveaus. Arch. Math. **21**, 617–628 (1970)
12. Schmitz, N.: Stichprobenumfänge des LQST bei homogenen Markov-Ketten. Operations Research-Verfahren **VIII**, 271–279 (1969/1970)
13. Wald, A.: Sequential analysis. New York: Wiley 1947

Colette Andrieu
29 rue J. Fouriaux
F-92160-Antony
France

(Reçu le 10 juin 1974, en forme revisée le 16 avril 1975)