NELLY MAIGRET

Statistique des chaînes contrôlées felleriennes

Astérisque, tome 68 (1979), p. 143-169
<http://www.numdam.org/item?id=AST_1979__68__143_0>

STATISTIQUE DES CHAÎNES CONTROLÉES FELLERIENNES

Nelly MAIGRET

Résumé :

On se pose ici, pour des chaînes contrôlées felleriennes, des problèmes d'estimation et de tests séquentiels. En s'inspirant du test de Chernoff dans le cas indépendant [1] , on construit une procédure. Grâce à un théorème de grandes déviations établi dans [4] , on montre que cette procédure est asymptotiquement la plus économique parmi les procédures de force au moins égale, pour le test de deux hypothèses simples.

Axiomatique et hypothèses

L'espace mesurable $(\Omega, \mathcal{A})$ est l'espace des épreuves, l'espace mesurable $(\mathcal{H}, \underline{\underline{\mathcal{H}}})$ celui des observations et $(\mathcal{A}, \underline{\underline{\mathcal{A}}})$ est l'espace des actions. Pour toute suite $(u_n)_{n\geqslant o}$, on note $u^{(n)} = (u_o, \ldots, u_n)$.

$(A_n)_{n\varepsilon N}$ est le processus des actions (l'évolution du système est contrôlée au temps n, par une action A_n mesurable de $(\Omega, \mathcal{A})$ dans $(\mathcal{A}, \underline{\underline{\mathcal{A}}}))$, et $(X_n)_{n\varepsilon N}$ est le processus des observations (X_n est l'état de la chaîne au temps n ; X_n est mesurable de $(\Omega, \mathcal{B}_n)$ dans $(\mathcal{H}, \underline{\underline{\mathcal{H}}}))$. On note $\mathcal{B}_n$ la tribu des évènements observables jusqu'au temps n : $\mathcal{B}_n = \sigma(X^{(n)}, A^{(n-1)})$.

A chaque instant n, on doit choisir l'action A_n au vu des observations faites jusqu'au temps n. Une stratégie aléatoire $\delta = (\delta_n)_{n\varepsilon N}$ est définie de la façon suivante : pour tout n, δ_n est une probabilité de transition de $\mathcal{H}^{n+1}$ dans $\mathcal{A}$. On note : $(x^{(n)} ; \delta) \longrightarrow \delta_n(x^{(n)} ; \delta)$. Seules seront utilisées les stratégies présentant certaines conditions de cohérence que l'on précisera plus loin. On notera $\mathcal{D}$ l'ensemble des stratégies cohérentes. $(\Theta, \underline{\underline{\Theta}})$ est l'espace des paramètres. On se donne une probabilité de transition π de $(\Theta \times \mathcal{H} \times \mathcal{A}, \underline{\underline{\Theta}} \times \underline{\underline{\mathcal{H}}} \times \underline{\underline{\mathcal{A}}})$ dans $(\mathcal{H}, \underline{\underline{\mathcal{H}}})$ définie de la manière suivante: à l'instant n, si $X_n = x$, $A_n = a$, alors X_{n+1} est dans Γ de $\underline{\underline{\mathcal{H}}}$, avec la probabilité $\pi(\theta, x, a ; \delta)$. Nous faisons les hypothèses suivantes :

Hypothèse 1

Pour toute stratégie $\delta = (\delta_n)_{n\varepsilon N}$, pour tout x de $\mathcal{H}$, pour toute valeur θ, il existe une probabilité $P^{\delta}_{\theta,x}$ sur $(\Omega, \mathcal{A})$ unique telle que pour tout n de N, pour tout x de $\mathcal{H}$, pour tout Γ de $\underline{\underline{\mathcal{H}}}$, on ait :

$$P^{\delta}_{\theta,x}(X_o = x) = 1$$

et pour $B \varepsilon \underline{\underline{\mathcal{H}}}$, pour $C \varepsilon \underline{\underline{\mathcal{A}}}$,

$$P^{\delta}_{\theta} (X \in B, A \in C) = \int \pi (\theta, X_n, a, du)\ 1_B(u)\ 1_C(a)\ \delta (X^{(n)}; da)$$

Le processus $(\Omega, \mathcal{A}, P^{\delta}_x, (X_n)_{n\varepsilon N}$ ainsi construit est une "chaine de Markov contrôlée".

Hypothèse 2

A tout x de $\mathcal{H}$, on associe un ensemble $D(x)$ non vide tel que l'ensemble $S = \{(x,a) \; ; \; a \in D(x)\}$ soit dans $\underline{\underline{\mathcal{H}}} \otimes \underline{\underline{\mathcal{A}}}$. Une stratégie aléatoire cohérente sera une stratégie aléatoire $\delta = (\delta_n)_{n\in N}$ telle que pour tout n , $\delta_n(X^{(n)} \; ; \; D(X_n)) = 1$.

Notations : ∘ Pour toute v.a. u sur $(\mathcal{H} \times \mathcal{A} \times \mathcal{H} \; , \underline{\underline{\mathcal{H}}} \otimes \underline{\underline{\mathcal{A}}} \otimes \underline{\underline{\mathcal{H}}})$, on définit la v.a. πu sur $(S \; , \underline{\underline{S}})$ par :

$$\pi \, u(\theta,x,a) = \int u(x,a,y) \; \pi(\theta,x,a,dy) \quad ,$$

si cette intégrale a un sens .

∘ Pour toute u sur $(\mathcal{H}, \underline{\underline{\mathcal{H}}})$, on définit la v.a. πu sur $(S \; , \underline{\underline{S}})$ par :

$$\pi \, u(\theta,x,a) = \int u(y) \, \Pi(\theta,x,a \; ; \; dy) \quad ,$$

toujours si cette intégrale a un sens.

Une partie importante de $\mathcal{D}$ est le sous ensemble $\mathcal{S}$ des stratégies aléatoires markoviennes stationnaires cohérentes : si s est une probabilité de transition de $\mathcal{H}$ dans $\mathcal{A}$ telle que pour tout x de $\mathcal{H}$, $s(x, D(x)) = 1$, la stratégie définie par $\delta_n(x^{(n)}, \; . \;) = s(x_n,.)$ est élément de $\mathcal{S}$, noté s .

Notations : Pour tout s de $\mathcal{S}$, pour tout x de $\mathcal{H}$, $\pi_\theta^{\mathcal{S}}$ désignera la transition $(\theta,x,C) \longrightarrow \pi(\theta,x,s(x) \; ; \; C)$ et on notera parfois par $\pi_\theta^{\mathcal{S}}$ la chaîne de Markov contrôlée par s .

Pour $C \in \underline{\underline{\mathcal{A}}}$, pour $x \in \mathcal{H}$,

$$\pi(\theta,x,s(x) \; ; \; C) = \int \pi(\theta,x,a \; ; \; C) \; s(x \; ; \; da) \quad .$$

Pour toute α probabilité sur $\mathcal{H}$, $\alpha \times \pi^{\mathcal{S}}$ désignera le probabilité :

$$\int 1_\Gamma(x,y)\ \alpha(dx)\ \pi(\theta,x,s(x)\ ;\ dy) = \int 1_\Gamma\ d(\alpha \times \pi^s)\ .$$

Hypothèse 3

$\mathcal{H}$, $\mathcal{A}$ et Θ sont métriques compacts et S est fermé.

Le noyau π est fellerien, c'est à dire que pour toute v.a. u bornée, continue sur $(S \times \mathcal{H}, \underline{\underline{S}} \times \underline{\underline{\mathcal{H}}})$, πu est continue.

Remarque : Sauf précision contraire, les espaces topologiques seront toujours munis de leur tribu borélienne et les espaces produits de la tribu produit.

On note $\mathcal{P}$ l'ensemble des probabilités sur $\mathcal{H}$, $\mathcal{P}_3$ celui des probabilités sur $\mathcal{H} \times \mathcal{A} \times \mathcal{H}$, concentrées sur S ; ils sont munis de la topologie de la convergence étroite. Pour λ de $\mathcal{P}_3$, on désigne par λ_1 et λ_3 ses premières et troisièmes marginales. Pour tout n de N, L_n est la fonction de répartition empirique : pour tout ω de Ω, pour tout Γ de $\underline{\underline{S}} \times \underline{\underline{\mathcal{H}}}$,

$$L_n(\omega,\delta) = \frac{1}{n} \sum_{k=o}^{n-1} 1_\Gamma(X_k\ ,\ A_k\ ,\ X_{k+1})\ (\omega)\ .$$

I - Estimation

1. Définitions - Exemples

1.1. Fonction de contraste

Une v.a. f sur $\Theta \times S \times \mathcal{H}$ est une fonction de contraste si pour tout (x,a) de S,

$$\int f(\theta,x,a,y)\ \pi(\theta,x,a\ ;\ dy)$$

est fini et
$$K(\theta,\phi,x,a) = \int (f(\phi,x,a,y) - f(\theta,x,a,y))\ \pi(\theta,x,a\ ;\ dy)$$

est positif et nul si et seulement si ϕ vaut θ .

Exemple 1

On suppose qu'il existe une v.a. p continue sur $\Theta \times S \times \mathcal{H}$ strictement positive et une transition λ de $\mathcal{H}$ dans $\mathcal{H}$, telles que pour tout Γ de $\mathcal{H}$, on ait :

$$\Pi(\theta,x,a\ ;\Gamma) = \int p(\theta,x,a,y)\ \lambda(x,dy)\ 1_\Gamma(y)\ .$$

On suppose aussi que pour tout (x,a) de S, $\Pi(\phi,x,a\ ;\cdot)$ est différent de $\Pi(\theta,x,a\ ;\cdot)$ dès que ϕ est différent de θ.

Sous ces hypothèses, la v.a. $f : f(\alpha,x,a,y) = -\operatorname{Log} p(\alpha,x,a,y)$ est une fonction de contraste, dès que les conditions d'intégrabilité sont vérifiées. De plus, la v.a. p étant strictement positive, pour tout (x,a) de S, l'ensemble

$$\{\Pi(\phi,x,a\ ;\cdot)\}_{\phi\varepsilon\Theta}$$

est un ensemble de probabilités équivalentes. Aussi, si p est bornée, on a :

$$K(\theta,\phi,x,a) = \int \operatorname{Log}\frac{p(\theta,x,a,y)}{p(\phi,x,a,y)}\ \Pi(\theta,x,a\ ;dy)$$

$$= I_{\Pi(\phi,x,a\ ;\cdot)}(\Pi(\theta,x,a\ ;\cdot))$$

où $I_{\Pi(\phi,x,a\ ;\cdot)}(\Pi(\theta,x,a\ ;\cdot))$ est l'information de Kullback de $\Pi(\theta,x,a\ ;\cdot)$ par rapport à $\Pi(\phi,x,a\ ;\cdot)$. On la note $I(\theta,\phi,x,a)$. Elle est strictement positive dès que ϕ est différent de θ.

Exemple 2

Soit g une v.a. sur $\mathcal{H}\times\mathcal{H}$, bornée. On définit :

$$m(\alpha,x,a) = \int g(x,y)\ \Pi(\alpha,x,a\ ;dy)$$

et on suppose que pour tout (x,a) de S, $m(\alpha,x,a)$ est différent de $m(\alpha',x,a)$ dès que α est différent de α'. La v.a. f définie par :

$$f(\alpha,x,a,y) = (g(x,y) - m(\alpha,x,a))^2$$

est une fonction de contraste.

1.2. Estimateur du minimum de contraste

Etant donnée f une fonction de contraste, on pose :

$$L_n(\alpha) = \sum_{k=o}^{n-1} f(\alpha, X_k, A_k, X_{k+1}) .$$

A l'instant n, un estimateur $\hat{\theta}_n$ est un estimateur du minimum de contraste si:

$$L_n(\hat{\theta}_n) = \inf_{\alpha\varepsilon\Theta} L_n(\alpha) .$$

Dans l'exemple 1, cet estimateur est l'estimateur du maximum de vraisemblance.

Le théorème de sélection de Brown assure l'existence d'une suite d'estimateurs du minimum de contraste si Θ est un espace métrique compact, et si l'application $\alpha \longrightarrow f(\alpha,x,a,y)$ est s.c.i de Θ dans R, pour tout (x,a,y) de $S \times \mathcal{H}$.

On posera :

$$h(\alpha,\phi,x,a,y) = f(\phi,x,a,y) - f(\alpha,x,a,y)$$

et

$$S_n(\alpha,\phi) = \sum_{k=o}^{n-1} h(\alpha,\phi,X_k, A_k, X_{k+1}) = \sum_{k=o}^{n-1} h_k(\alpha,\phi) .$$

2. Consistance exponentielle

Si λ_o est la distribution initiale de la chaîne, une suite $(\hat{\theta}_n)_{n\varepsilon N}$ d'estimateurs de θ est fortement consistante si pour tout voisinage $U(\theta)$ de θ, pour toute stratégie cohérente δ, la suite $(\hat{\theta}_n)$ tend vers θ $P^{\delta}_{\theta,x}$ p.s.

Le but est d'établir ici la consitance exponentielle de toute suite $(\hat{\theta}_n)_{n\varepsilon N}$ d'estimateurs du minimum de contraste lorsque la fonction de contraste f est continue sur $\Theta \times S \times \mathcal{H}$. (Le théorème de Brown assure l'existence d'une suite d'estimateurs du minimum de contraste lorsque f est continue. On améliore ainsi le résultat de Georgin [2] qui a montré sous les mêmes hypothèses, la consistance forte.

Pour tout voisinage $U(\theta)$ de θ, on définit la v.a. :

$$\Gamma(U(\theta)) = \inf\{n ; \quad p \geqslant n , \quad \hat{\theta}_p \varepsilon U(\theta)\}$$

Lorsque Θ sera fini, on notera T_θ la v.a. $T(\{\theta\})$.

<u>Théorème 1</u>

Lorsque la fonction de contraste f est continue sur $\Theta \times S \times \mathcal{H}$ toute suite $(\hat{\theta}_n)_{n\varepsilon N}$ d'estimateurs du minimum de contraste est exponentiellement consistante : pour tout $U(\theta)$ voisinage ouvert de θ, on a :

$$\overline{\lim_n} \frac{1}{n} \text{Log} \sup_{x\varepsilon \mathcal{H}} \sup_{\delta\varepsilon \mathcal{D}} P^\delta_{\theta,x}(P(U(\theta)) > n) < 0 .$$

<u>Démonstration</u>

Soit $U(\theta)$ un voisinage ouvert de θ. L'ensemble $\{T(U(\theta)) > n\}$ est inclus dans :

$$\bigcup_{m>n} \{ \inf_{\phi\varepsilon U^c(\theta)} \sum_{k=o}^{m-1} h_k(\theta,\phi) < 0\}$$

et pour tout $x \varepsilon \mathcal{H}$ et $\delta \varepsilon \mathcal{D}$,

$$P^\delta_{\theta,x}(T(U(\theta)) > n) \leqslant \sum_{m\geqslant n} P^\delta_{\theta,x}(\inf_{\phi\varepsilon U^c(\theta)} \sum_{k=o}^{m-1} h_k(\theta,\phi) \leqslant 0)$$

Soit $\Gamma = \{\mu \varepsilon \mathcal{P}_3 , \inf_{\phi\varepsilon U^c(\theta)} \int h(\theta,\phi,x,a,y)\, d\mu(x,a,y) \leqslant 0\}$. On a :

$$P^\delta_{\theta,x}(T(U(\theta)) > n) \leq \sum_{m\geqslant n} P^\delta_{\theta,x}(L_m \varepsilon \Gamma) .$$

Montrons que Γ est un fermé de $\mathcal{P}_3$.

Soit μ de $\overline{\Gamma}$. Il existe alors une suite $(\mu_n)_{n\varepsilon N}$ de Γ convergeant vers μ dans $\mathcal{P}_3$. Pour tout λ de $\mathcal{P}_3$, l'application :

$$\phi \longrightarrow \int h(\theta,\phi,x,a,y)\, d\lambda(x,a,y)$$

est continue de Θ dans R . On définit alors par $(\phi_n)_{n\varepsilon N}$ une suite de $U^c(\theta)$ qui vérifie :

$$\mu_n(\phi_n) = \inf_{\phi\varepsilon U^c(\theta)} (\int h(\theta,\phi,x,a,y)\, d\mu_n(x,a,y)) .$$

Pour tout n , $\mu_n(\phi_n)$ est négatif. On considère maintenant $(\phi_{n_k})_{k\varepsilon N}$ une sous suite de $(\phi_n)_{n\in N}$, qui converge vers ϕ_o de $U^c(\theta)$. La sous suite $h(\theta,\phi_{n_k},x,a,y)$ converge donc vers $h(\theta,\phi_o,x,a,y)$ uniformément sur $S \times \mathcal{H}$. Puisque $(\mu_{n_k})_{k\in N}$ tend vers dans $\mathcal{P}_3$, $(\mu_{n_k}(\phi_o))_{k\in N}$ tend vers $\mu(\phi_o)$ et $(\mu_{n_k}(\phi_{n_k}))_{k\in N}$ tend aussi vers $\mu(\phi_o)$, ce qui permet de conclure que $\mu(\phi_o)$ est négatif et μ appartient à Γ : Γ est un fermé de $\mathcal{P}_3$.

Puisque f est une fonction de contraste,

$$\int h(\theta,\phi,x,a,y)\,\Pi(\theta,x,a\,;\,dy)\,d\lambda_1(x)\,s(x,da)$$

est strictement positif pour toute λ_1 de $\mathcal{P}$ et pour toute s de $\mathcal{S}$, dès que ϕ est différent de θ. Comme h est continue sur le compact $\Theta \times \Theta \times S \times \mathcal{H}$, l'application :

$$\phi \longrightarrow \int h(\theta,\phi,x,a,y)\,\Pi(\theta,x,a\,;\,dy)\,d\lambda_1(x)\,s(x,da)$$

est continue de Θ dans R . Aussi dès que ϕ est différent de θ,

$$\inf_{\phi\in U^c(\theta)} \int h(\theta,\phi,x,a,y)\,\Pi(\theta,x,a\,;\,dy)\,d\lambda_1(x)\,s(x,da)$$

est strictement positif ; c'est à dire que $\lambda_1 \otimes s \otimes \Pi$ n'appartient pas à Γ, quelles que soient λ_1 de $\mathcal{P}$ et s de $\mathcal{S}$.

Le théorème des grandes déviations [4] entraîne :

$$\overline{\lim_n}\ \frac{1}{n}\ \text{Log}\ \sup_x\ \sup_\delta\ P^\delta_{\theta,x}(T(U(\theta)) > n) \lesssim -I(\Gamma \cap \Delta) < 0 .$$

II - Γ est séquentiel du rapport de vraisemblance

On se pose ici le même problème de test que dans [3] .

L'espace des paramètres est supposé fini (de cardinal q) , (Θ_1, Θ_2) est une partition de Θ et on veut savoir à quel élément de la partition appartient la vraie valeur du paramètre. Dans ce cadre contrôlé, il faudra en plus des règles d'arrêt et de décision, s'intéresser au choix d'un bon contrôle. Chaque expérience a un coût c, $0 < c < 1$, et on note r(c) la perte dûe à l'erreur de décision.

On se place dans le cadre de l'exemple 1, et on notera :

$$J(\theta,\phi) = \inf_{(x,a)\varepsilon S} I(\theta,\phi,x,a) , \qquad J(\theta) = \inf_{\phi\varepsilon a(\theta)} J(\theta,\phi) > 0$$

$$L(\theta,\phi) = \sup_{(x,a)\varepsilon S} I(\theta,\phi,x,a) , \qquad L(\theta) = \inf_{\phi\varepsilon a(\theta)} L(\theta,\phi) .$$

$(\hat{\theta}_n)_{n\varepsilon N}$ désigne désormais une suite d'estimateurs du maximum de vraisemblance, elle est exponentielle consistante.

Comme dans la partie sans contrôle, on note $h(\theta) = \Theta_i$ si θ est élément de Θ_i (i = 1 ou 2) , et $a(\theta) = \Theta - h(\theta)$, et on désigne par $(\hat{\hat{\theta}}_n)_{n\varepsilon N}$ une suite d'estimateurs de θ qui vérifient pour tout n :

$$\text{Log} \prod_{h=o}^{n-1} p(\hat{\hat{\theta}}_n, X_k, A_k, X_{k+1}) = \sup_{\alpha\varepsilon a(\hat{\theta}_n)} \text{Log} \prod_{k=o}^{n-1} p(\alpha, X_k, A_k, X_{k+1})$$

Les résultats obtenus sont des résultats asymptotiques.

Sans hypothèses supplémentaires, nous construisons un test séquentiel clos quelle que soit la stratégie ; nous obtenons une majoration (asymptotique) du risque, indépendante de la stratégie. Aucune stratégie n'est particularisée. Avec une hypothèse supplémentaire de récurrence, pour le test de 2 hypothèses simples, nous construisons une procédure (règles d'arrêt et de décision, stratégie) possédant une bonne propriété d'optimalité.

1. Etude d'un temps d'arrêt

1.1. Préliminaires

Lemme 1

$\mathcal{Y}$ étant un espace métrique compact et q uns transition de $\mathcal{Y}$ dans $\mathcal{Y}$ telle que l'application $x \longrightarrow q(x,\cdot)$ soit continue de $\mathcal{Y}$ dans $\mathcal{P}(\mathcal{Y})$ ensemble des probabilités sur $\mathcal{Y}$ muni de la topologie de la convergence étroite ; alors pour toute g v.a. sur $\mathcal{Y}\times\mathcal{Y}$ continue, l'application

$$x \longrightarrow q\ g(x) = \int g(x,y)\, q(x,dy)$$

est continue de $\mathcal{Y}$ dans $\mathcal{R}$.

Démonstration : Facile ■

Lemme 2

Soit $(Y_n)_{n\in N}$ une suite de v.a. définies sur un espace $(B, \underline{\underline{B}}, \lambda)$, intégrables adaptée à une famille croissante $(\sigma_n)_{n\in N}$ de sous tribus de $\underline{\underline{B}}$ vérifiant : il existe un entier positif tel que pour tout n, on ait :

$$E(Y_{n\wedge k}/\sigma_{n-1}) \geqslant m > 0.$$

On définit pour b réel, la v.a. ν'_b par $\inf\{n \,;\, y_1 + \dots + y_n \geqslant b\}$.

On a : $E(\nu'_b) \leqslant \frac{b+k}{m}$.

Démonstration

Il suffit de montrer le lemme pour des v.a. majorées par k.

On définit :

$$V_n = Y_1 + \dots + Y_n$$

$$A_n = \sum_{i=1}^{n} E(Y_i/\sigma_{i-1})$$

$$M_o = 0 \quad \text{et} \quad M_n = V_n - A_n$$

Puisque Y_n est intégrable, V_n et M_n sont intégrables, et $E(M_n/\sigma_{n-1}) = M_{n-1}$. $(M_n)_{n\geqslant o}$ est donc une martingale adaptée à la famille $(\sigma_n)_{n\geqslant o}$. Pour tout n entier naturel, on a d'après le théorème d'arrêt :

$$E(M_{\nu'_b\wedge n}) = 0 = E(V_{\nu'_b\wedge n}) - E(A_{\nu'_b\wedge n})$$

Comme : $$E(A_{\nu'_b\wedge n}) \geqslant m\, E(\nu'_b\wedge n)$$

et : $$E(V_{n\wedge\nu'_b}) = E(V_n\, 1_{\{n<\nu'_b\}}) + E((V_{\nu'_b-1} + Y_{\nu'_b})\, 1_{\{n\geqslant\nu'_b\}})$$

$$\leqslant b\,\lambda(n < \nu'_b) + (b+k)\,\lambda(n \geqslant \nu'_b) \leqslant (b+k)$$

On a : $$E(\nu'_b\wedge n) \leqslant \frac{b+k}{m}$$

$$E(\nu'_b) = \lim_{n\to\infty} \uparrow E(\nu'_b\wedge n) \leqslant \frac{b+k}{m} .$$

1.2. Définition d'un temps d'arrêt. Propriété de finitude

Soit b réel positif et ϕ de Θ. On définit :

$$\tau_b = \inf\{n \,;\, S_n(\hat{\theta}_n , \hat{\hat{\theta}}_n) > b\}$$

$$\tau_{b,\phi} = \inf\{n \,;\, \forall\, p \geqslant n ,\ S_p(\theta,\phi) > b\}$$

On a : $\tau_b \leqslant \max(T_\theta , \max\limits_{\phi \varepsilon a(\theta)} \tau_{b,\phi})$ (1.2)

∘ Une application directe du lemme 2 permet d'écrire :

Lemme 3

Soit ϕ différent de θ. S'il existe un entier narurel k_o et un réel m, strictement positifs tels qu'on ait :

$$\inf_{(x,a)\varepsilon S} \psi_{k_o}(\theta,\phi,x,a) = \inf_{(x,a)\varepsilon S} \int \inf\left(k_o, \text{Log}\,\frac{p(\theta,x,a,y)}{p(\phi,x,a,y)}\right)\Pi(\theta,x,a\ ;\ dy) \geqslant m > 0$$

alors : $\sup\limits_{x\varepsilon \mathcal{H}} \sup\limits_{\delta\varepsilon \mathfrak{D}} E^{\delta}_{\theta,x}(\tau_{b,\phi}) \leqslant \dfrac{b+k_o}{m}$.

Montrons que les hypothèses du lemme 3 sont vérifiées.

Le théorème de convergence de Béppo-Lévi entraîne :

$$\lim_k \uparrow \psi_k(\theta,\phi,x,a) = \text{Log}\int \frac{p(\theta,x,a,y)}{p(\phi,x,a,y)}\, \Pi(\theta,x,a\ ;\ dy) = I(\theta,\phi,x,a)$$

Puisque p est continue et puisque le noyau Π est fellerien, le lemme 1, entraîne que $I(\theta,\phi,\cdot,\cdot)$ et $\psi_k(\theta,\phi,\cdot,\cdot)$ sont continues sur le compact S. Ce qui permet de conclure, grâce au théorème de Dini que la suite $(\psi_k(\theta,\phi,\cdot,\cdot))_{k\varepsilon N}$ converge uniformément vers $I(\theta,\phi,\cdot,\cdot)$ sur S. Donc, pour tout ϕ différent de θ, il existe un entier k_o, tel que pour tout $k \geqslant k_o$, $\inf\limits_{(x,a)\varepsilon S} \psi_k(\theta,\phi,x,a)$ soit strictement positif.

Grâce à la relation (1.2) et à la consitance exponentielle, on en déduit que les temps d'arrêt τ_b sont $P^{\delta}_{\theta,x}$ presque sûrement finis pour tout x et pour toute δ de $\mathfrak{D}$.

1.3. Définition d'une procédure

On note ν_c le temps d'arrêt $\tau_{-\mathrm{Log}\, c}$ et $\nu_{c,\phi}$ le temps d'arrêt $\tau_{-\mathrm{Log}\, c,\phi}$. Pour une stratégie δ de $\mathcal{D}$, on désigne par ϕ_c^δ la procédure suivante :

- sa règle d'arrêt est ν_c
- si $n = \nu_c$, on arrête et on décide $h(\hat{\theta}_n)$
- si $n < \nu_c$, on continue , on contrôle la chaîne avec la stratégie δ et on refait une expérience.

Les procédures ϕ_c^δ sont closes pour toute valeur du coût c et toute stratégie δ . $\beta_c^\delta(\theta)$ et $R_c^\delta(\theta)$ désigneront respectivement la probabilité d'erreur et le risque de la procédure ϕ_c^δ lorsque θ est le vrai paramètre.

Toute autre procédure séquentielle de temps d'arrêt T , de règle de décision D et de stratégie δ sera notée (T, D, δ) .

1.4. Propriétés

Théorème 2

Pour tout θ de Θ , on a :

$$\sup_{\delta \in \mathcal{D}} \beta_c^\delta(\theta) \leq q\, c\ .$$

Démonstration

Elle est analogue à celle faite dans [3] (théorème 3.a) ; il suffit de remarquer que pour tout δ de $\mathcal{D}$,

$$P_\theta^\delta = \prod_{i=o}^{n-1} \frac{p(\theta, X_i, A_i, X_{i+1})}{p(\phi, X_i, A_i, X_{i+1})}\ P_\phi^\delta \quad \text{sur } \mathcal{B}_n\ .$$

Théorème 3

Pour tout θ de Θ , on a :

$$\sup_{\delta \in \mathcal{D}} E_\theta^\delta(\nu_c) \leq \frac{-(1+\sigma(1))\,\mathrm{Log}\, c}{J(\theta)}\ .$$

Démonstration

Soit $\varepsilon > 0$ et $N_c = \dfrac{-(1 + \frac{\varepsilon}{2}) \text{Log } c}{J(\theta)}$.

Il suffit de prouver qu'il existe k et b strictement positifs, tels que pour tout $n > N_c$, on ait :

$$\sup_{\delta \varepsilon \mathcal{D}} P_\theta^\delta(\nu_c = n) \leqslant k\, e^{-bn}$$

Le reste de la démonstration est analogue à celle du théorème 3.b de [3] .

La relation : $\{\nu_c = n\} \subset \{T_\theta > n\} \cup \{ \bigcup_{\phi \varepsilon a(\theta)} \{\nu_{c,\phi} = n\}\}$ entraîne pour tout n de N :

$$\sup_{\delta \varepsilon \mathcal{D}} P_\theta^\delta(\nu_c = n) \leqslant \sup_{\delta \varepsilon \mathcal{D}} P_\theta^\delta(T_\theta \geqslant n) + \sum_{\phi \varepsilon a(\theta)} \sup_{\delta \varepsilon \mathcal{D}} P_\theta^\delta(\nu_{c,\phi} = n)$$

La suite $(\hat{\theta}_n)_{n \varepsilon N}$ étant exponentiellement consistante pour toute stratégie, il existe k_1 et b_1 strictement positifs tels que pour tout n de N , on ait :

$$\sup_{\delta \varepsilon \mathcal{D}} P_\theta^\delta(T_\theta \geqslant n) \leqslant k_1\, e^{-b_1 n} .$$

D'autre part, d'après le corollaire 1.3. du théorème des grandes déviations [4] , il existe k_2 et b_2 strictement positifs, tels que pour tout ϕ de Θ , pour tout n de N , on ait :

$$\sup_{\delta \varepsilon \mathcal{D}} P_\theta^\delta \left(S_n(\theta,\phi) \leqslant \frac{n\, J(\theta)}{1 + \frac{\varepsilon}{2}}\right) \leqslant k_2\, e^{-b_2 n}$$

c'est à dire :

$$\sup_{\delta \varepsilon \mathcal{D}} P_\theta^\delta(S_n(\theta,\phi) \leqslant -\text{Log } c) \leqslant k_2\, e^{-b_2 n} ,$$

dès que n est supérieur à N_c .

Comme : $\nu_{c,\phi}$ vaut $\inf\{n \,;\, \forall\, p \geqslant n ,\; S_p(\theta,\phi) > -\text{Log } c\}$, on a pour toute δ de $\mathcal{D}$:

$$P_\theta^\delta(\nu_{c,\phi} = n) \leqslant P_\theta^\delta(S_{n-1}(\theta,\phi) \leqslant -\text{Log } c)$$

c'est à dire, il existe k_2' et b_2' strictement positifs, tels que pour tout ϕ de Θ , pour tout $n > N_n$, on ait :

$$\sup_{\delta \in \mathfrak{D}} P_\theta^\delta(\nu_{c,\phi} = n) \leqslant k_2' \, e^{-b_2' n}$$

Ce qui permet de conclure.

Corollaire 1

Pour tout θ de Θ,

$$\sup_{\delta \in \mathfrak{D}} R_c^\delta(\theta) \leqslant \frac{-(1 + \sigma(1))\, c \operatorname{Log} c}{J(\theta)} .$$

Les lemmes suivantes seront utilisés pour demander une propriété d'optimalité.

Lemme 4

Soit b , $0 < b < 1$. Soit (T, D, δ) une procédure séquentielle, $e(\theta)$ la probabilité d'erreur associée. Si pour tout θ, $e(\theta)$ est un $\mathcal{O}(-b \operatorname{Log} b)$; on a pour tout $\varepsilon \in]0,1[$ et $(\theta,\phi) \in \Theta^2$,

$$P_\theta^\delta \, (S_T(\theta,\phi) < -(1 - \varepsilon) \operatorname{Log} b) = \mathcal{O}(b^\varepsilon \operatorname{Log} b) .$$

Démonstration

En se rappelant que :

$$P_\phi^\delta = \prod_{k=o}^{n-1} \frac{p(\phi, X_k, A_k, X_{k+1})}{p(\theta, X_k, A_k, X_{k+1})} \, P_\theta^\delta \quad \text{sur } \mathcal{B}_n ,$$

la démonstration se fait de la même façon que celle du lemme 3.a de [3]. ■

Lemme 5

Pour tout $\varepsilon > 0$, pour toute stratégie δ de $\mathfrak{D}$,

$$P_\theta^\delta(\max_{1 \leqslant m \leqslant n} \; \min_{\phi \varepsilon a(\theta)} S_m(\theta,\phi) \geqslant n(L(\theta) + \varepsilon)) \xrightarrow[n \uparrow \infty]{} 0$$

Démonstration

Soit δ une stratégie cohérente , et ε strictement positif.

Pour tout m de N , pour tout ϕ ,

$$S_m(\theta,\phi) = (S_m(\theta,\phi) - \sum_{k=0}^{m-1} I(\theta,\phi,X_k,A_k)) + \sum_{k=0}^{m-1} I(\theta,\phi,X_k,A_k)$$

$$S_m(\theta,\phi) = Z_{m,1}(\theta,\phi) + Z_{m,2}(\theta,\phi)$$

$$Z_{m,2}(\theta,\phi) \leqslant m\ L(\theta,\phi)\ P_\theta^\delta \text{ ps}$$

$$\min_{\phi\varepsilon a(\theta)} Z_{m,2}(\theta,\phi) \leq m\ L(\theta)\ ,\ P_\theta^\delta \text{ ps}$$

Ce qui entraîne :

$$P_\theta^\delta(\max_{1<m<n} \min_{\phi\varepsilon a(\theta)} S_m(\theta,\phi) \geqslant n(L(\theta)+\varepsilon)) \leqslant P_\theta^\delta(\max_{1\leqslant m\leqslant n} \min_{\phi\varepsilon a(\theta)} Z_{m,1}(\theta,\phi) > n\ \varepsilon)$$

$$\leqslant \sum_{\phi\varepsilon a(\theta)} P_\theta^\delta(\max_{1\leqslant m\leqslant n} Z_{m,1}(\theta,\phi) > n\ \varepsilon)$$

$\{Z_{m,1}(\theta,\phi)\}$ est une martingale adaptée aux tribus $(\mathcal{B}_n)_{n\varepsilon N}$ sur $(\Omega,\mathcal{A},P_\theta^\delta)$. D'après l'inégalité de Doob , on a :

$$P_\theta^\delta(\max_{1\leqslant m\leqslant n} \min_{\phi\varepsilon a(\theta)} S_m(\theta,\phi) \geqslant n(L(\theta)+\phi)) \leqslant \sum_{\phi\varepsilon a(\theta)} \frac{E_\theta^\delta(Z_{n,1}^2)}{n^2\ \varepsilon^2}$$

et on vérifie facilement que pour tout $n\ \varepsilon\ N$, $\frac{1}{n} E_\theta^\delta(Z_{n,1}^2)$ est majorée par une constante finie M .

$$P_\theta^\delta(\max_{1\leqslant m\leqslant n} \min_{\phi\varepsilon a(\theta)} S_m(\theta,\phi) \geqslant n(L(\theta)+\varepsilon) \leqslant \frac{qM}{n\varepsilon}\ . \blacksquare$$

<u>Lemme 6</u>

Soit b , $0 < b < 1$. Soit (T, D, δ) une procédure séquentielle et $e(\theta)$ la probabilité d'erreur associée lorsque θ est le paramètre. On suppose que pour tout θ, $e(\theta)$ est un $\mathcal{O}(-b \text{ Log } b)$. On a :

$$E_\theta^\delta(T) \geqslant \frac{-(1+\sigma(1)) \text{ Log } b}{L(\theta)}\ .$$

<u>Démonstration</u>

On suit la démonstration du lemme 3.c de [3] en utilisant les 2 lemmes ci-dessus.

Conséquence

Une procédure (T, D, δ) telle que pour tout θ, la probabilité d'erreur $e(\theta)$ soit un $\mathcal{O}(-c \text{ Log } c)$, vérifie : pour tout θ,

$$E_\theta^\delta(T) \geqslant \frac{-(1 + \sigma(1)) \text{ Log } c}{L(\theta)} .$$

En se rappelant que l'erreur de ϕ_c^δ est un $\mathcal{O}(c)$, on obtient :

Théorème 4

Toute procédure de force au moins égale aux procédures $\{\phi_c^\delta\}_{\delta\varepsilon\mathfrak{D}}$ dure en moyenne plus longtemps que $\frac{-(1+\sigma(1))\text{Log } c}{L(\theta)}$, qui est du même ordre de grandeur que $\frac{-(1 + \sigma(1))\text{Log } c}{J(\theta)}$, qui majore la durée moyenne des procédures $\{\phi_c^\delta\}_{\delta\varepsilon}$.

III - Test de deux hypothèses simples pour des chaînes contrôlées "récurrentes"

Ces résultats ne particularisent pas une stratégie ; valables pour toute stratégie, ils ne sont pas très précis. On n'obtient pas de stratégie meilleure que les autres. En ajoutant une hypothèse de récurrence, nous construisons une procédure ayant une bonne propriété d'obtimalité pour le test de 2 hypothèses simples : le temps d'arrêt et la règle de décision seront les mêmes que précédemment, mais nous particularisons une stratégie.

1.5. Rappels de quelques résultats

Soit s une stratégie markoviennes stationnaire, telle que la chaîne associée à Π_θ^s soit récurrente Doeblin, apériodique, de probabilité invariante μ_θ^s . L'information de Kullback de $\Pi(\theta,x,s(x) ; \cdot)$ par rapport à $\Pi(\phi,x,s(x) ; \cdot)$ est :

$$I_s(\theta,\phi,x) = \int \text{Log } \frac{p(\theta,x,s(x) ; y)}{p(\phi,x,s(x) ; y)} \Pi(\theta,x,s(x) ; dy)$$

On définit :

$$I_s(\theta,\phi) = \int I_s(\theta,\phi,x) \, d\mu_\theta^s(x)$$

et
$$I_s(\theta) = \min_{\phi \in a(\theta)} I_s(\theta,\phi)$$

$I_s(\theta)$ est strictement positif.

Les théorème 3.a et 3.b |3| s'écrivent ici :

$$\beta_c^s(\theta) \leqslant q\,c \qquad \text{et}$$

$$E_\theta^s(\nu_c) \leqslant \frac{-(1+\sigma(1))\operatorname{Log} c}{I_s(\theta)} \quad .$$

La stratégie s semble d'autant meilleure qu'elle est d'information maximum. Nous exploitons cette idée pour le test de 2 hypothèses simples. Donc ici , $\mathcal{O} = \{\theta_1, \theta_2\}$, (on notera θ la vraie valeur , ϕ_θ l'autre valeur) , et on fait les hypothèses suivantes :

(1) : Toutes les stratégies markoviennes cohérentes stationnaires sont récurrentes Doeblin, apériodiques. (Pour une telle s , on note μ_θ^s la probabilité invariante) .

(2) : Il existe une stratégie markovienne stationnaire déterministe s^θ qui permet d'obtenir le gain moyen : $\overline{\lim_n} \frac{S_n(\theta,\phi)}{n}$, optimal . (On dit que s^θ est optimale). On notera $I(\theta) = I_{s^\theta}(\theta) = \sup_{s\varepsilon} I_s(\theta)$.

(1) est vérifié par exemple, s'il existe une mesure positive non nulle sur $\mathcal{H}$, telle que $\inf_{(x,a)\varepsilon S} \Pi(\theta,x,a\,;\,\cdot\,) > \nu(\cdot)$. Si en plus les espaces d'états $\mathcal{H}$ et d'actions $\mathcal{A}$ sont finis, alors (2) est vérifiée. (Ces résultats on été montrés lors de la recherche du gain optimal sous contraintes, dans [2] .

On connait des critères plus fins sous lesquels la condition (2) est vérifiée ([2]) .

<u>1.6.</u> : Désormais s^θ désigne une telle stratégie optimale :

$$\lim_n \frac{1}{n} S_n(\theta,\phi_\theta) \xrightarrow[n\uparrow\infty]{P_\theta^{s^\theta}\ ps} I(\theta)$$

et P_θ^δ ps , pour toute stratégie δ cohérente : $\overline{\lim_n} \frac{1}{n} S_n(\theta,\phi_\theta) < I(\theta)$.

De plus, on a :

<u>a</u> $\beta_c^{s^\theta}(\theta) \leqslant 2c$

<u>b</u> $E_\theta^{s^\theta}(\nu_c) \leqslant \dfrac{-(1+\sigma(1))\text{Log } c}{I(\theta)}$.

Nous allons montrer que s^θ possède une "bonne" propriété d'optimalité.

<u>Lemme 7</u>

Pour tout $\varepsilon > 0$, pour toute δ de $\mathfrak{D}$,

$$P_\theta^\delta\left(\max_{1\leqslant m\leqslant n} S_m(\theta,\phi_\theta) \geqslant n(I(\theta)+\varepsilon)\right) \xrightarrow[n\uparrow\infty]{} 0 .$$

<u>Démonstration</u>

Soit δ une stratégie de $\mathfrak{D}$, ε et η quelconques fixés strictement positifs. La relation : $\overline{\lim_n} \frac{1}{n} S_n(\theta,\phi_\theta) \leqslant I(\theta)$ P_θ^δ ps entraîne qu'il existe un entier n_o vérifiant : $P_\theta^\delta\left(\max_{n_o\leqslant m\leqslant n} S_m(\theta,\phi_\theta) \geqslant n(I(\theta)+\varepsilon)\right) \leqslant \eta$.

De plus il existe n_1 tel que pour tout n supérieur à n_1 , on ait :

$$P_\theta^\delta\left(\max_{0\leqslant m\leqslant n_o} S_m(\theta,\phi_\theta) \geqslant n(I(\theta)+\varepsilon)\right) = 0 .$$

Ce qui implique qu'il existe $n_2(\varepsilon,\eta)$, tel que pour tout n supérieur à n_2 , on ait :

$$P_\theta^\delta\left(\max_{0\leqslant m\leqslant n} S_m(\theta,\phi_\theta) \geqslant n(I(\theta)+\varepsilon)\right) \leqslant \eta .$$

<u>Lemme 8</u>

Soit b , $0 < b < 1$, et soit (T, D, δ) une procédure telle que la probabilité d'erreur associée $e(\phi)$ soit un $\mathcal{O}(-b \log b)$ pour tout ϕ.

On a :

$$E_\theta(T) \geqslant \frac{-(1+\sigma(1)) \text{ Log } b}{I(\theta)} .$$

Démonstration

Elle découle du lemme précédent et du lemme 4. Elle est la même que celle du lemme 6 .

De ce dernier résultat, et en rappelant ceux de a) et b) , on conclut :

Théorème 5

La procédure $\phi_c^{s^\theta}$ est la plus économique parmi les procédures de force au moins égale.

Mais la stratégie s^θ n'est pas calculable puisqu'elle dépend du paramètre θ . On considère la suite $(\hat{\theta}_n)_{n\varepsilon N}$ consistante exponentiellement pour toute stratégie. L'idée est alors de prendre la stratégie $\hat{s}$ obtenues en remplaçant θ dans s^θ par son estimateur : $\hat{s} = (s_n)_{n\varepsilon N}$ où pour tout n , on a :

$$s_n(X^{(n)}) = s^{\hat{\theta}_n}(X_n) \quad .$$

Le but est de montrer que $\hat{s}$ est aussi bonne sur s^θ . Nous allons prouver que :

$$E_\theta^{\hat{s}}(\nu_c) \leqslant \frac{-(1+\sigma(1)) \text{ Log } c}{I(\theta)}$$

ce qui permettra de conclure, (puisqu'on a aussi $\beta^{\hat{s}}(\theta) \leqslant 2c)$.
Pour cela, nous introduisons la notion de stratégies "exponentiellement voisines".

1.7. Stratégies exponentiellement voisines

On suppose ici Θ réduit à un point et on omet le paramètre. Les hypothèses topologique sur $\mathcal{H}$, $\mathcal{A}$ et Π sont inutiles.

Définition

Soient d et δ 2 stratégies. On définit la v.a. :

$$T(d,\delta) = \inf \{n \,;\forall\, m \geqslant n \,,\ d_m = \delta_m\}$$

d et δ seront dites exponentiellement voisines si il existe k et b

strictement positifs tels que pour tout n ,

$$P^{d}(T(d,\delta) > n) \leqslant k\, e^{-bn}$$

et
$$P^{\delta}(T(d,\delta) > n) \leqslant k\, e^{-bn}$$

Exemple : Revenant au cadre général de cet article, pour toute stratégie d^{θ}, on note $\hat{d}$ l'estimée de d^{θ} obtenue en remplaçant θ par son estimateur, c'est à dire, $\hat{d}$ est définie par :

$$\hat{d} = \hat{d}_n(X^{(n)}) \ , \quad \text{où pour tout } n \ , \quad d_n(X^{(n)}) = d^{\hat{\theta}_n}(X^{(n)}) \ .$$

La consistance exponentielle de la suite $(\hat{\theta}_n)_{n\in N}$ pour toute stratégie entraîne que d^{θ} et $\hat{d}$ sont exponentiellement voisines.

Théorème 6

Soit s une stratégie markovienne stationnaire récurrente Doeblin, apériodique, de probabilité invariante μ^{s}. Soit δ une stratégie de $\mathcal{D}$ telle que δ et s soient exponentiellement voisines, et soit g une v.a. bornée sur $S \times \mathcal{H}$, non nulle. Pour tout $\varepsilon > 0$, on a :

$$\overline{\lim_{n}}\ \frac{1}{n} \operatorname{Log} \sup_{x\varepsilon} P_x^{\delta}\left(\frac{1}{n}\sum_{p=0}^{n-1} g(X_p, A_p, X_{p+1}) \leq \left(\left(\int gd\ \mu^{s} \otimes s \otimes \Pi\right) - \varepsilon\right) < 0\right.$$

Démonstration

s étant récurrente Doeblin apériodique, il existe $a > 0$ et ρ, $0 < \rho < 1$, tels que pour tout n de N , on ait :

$$\sup_{x\varepsilon \mathcal{H}} ||\Pi_n(x,s(x),\cdot) - \mu^{s}(\cdot)|| \leqslant a_{\rho}^{n} \ .$$

δ et s étant exponentiellement voisines de s , il existe b et k strictement positifs tels que pour tout n de N , on ait :

$$P_{\theta}^{s}(T(s,\delta) > n) \leqslant k\, e^{-bn}$$

et
$$P_{\theta}^{\delta}(T(s,\delta) > n) \leqslant k\, e^{-bn}$$

Dans la démonstration, on écrira T au lieu de $T(s,\delta)$ et on notera θ_n l'opérateur de translation de $(\mathcal{H}\times\mathcal{A})^N$: $\theta_n((x_p,a_p)_{p\in N})=((x_{p+n},a_{p+n}))_{p\in N}$

$$\sup_{A\in\mathcal{a}} |P_x^\delta(\theta_{n+p}(A)) - P_x^s(\theta_{n+p}(A))| =$$

$$= \sup_{A\in\mathcal{a}} |P_x^\delta(\theta_{n+p}(A)\cap (T > n)) + P_x^\delta(\theta_{n+p}(A)\cap (T \leqslant n)$$

$$- P_x^s(\theta_{n+p}(A)\cap (T > n)) - P_x^s(\theta_{n+p}(A)\cap (T < n))|$$

$$\leqslant P_x^\delta(T > n) + P_x^s(T > n) + \sup_{A\in\mathcal{a}} |P_x^s(\theta_{n+p}(A) \cap (T < n)) - P_x^\delta(\theta_{n+p}(A)\cap (T < n))|$$

$$\leqslant 2ke^{-bn} + \sup_{A\in\mathcal{a}} |E_x^s(1_{\{T<n\}} E_x^s(1_{\{\theta_{n+p}(A)\}} /\mathcal{B}_n)) - E_x(1_{(T<n)}E_n^\delta(1_{\{\theta_{n+p}(A)\}} /\mathcal{B}_n))|$$

$$\leqslant 2ke^{-bn} + \sup_{A\in\mathcal{a}} |E_x^s(1_{\{T<n\}} \int \Pi_p^s(X_n,dy)P_y^s(A)) - E_x^\delta(1_{\{T<n\}} \int \Pi_p^s(X_n;dy)P_y^s(A))|$$

$$\leqslant 2ke^{-bn} + \sup_{A\in\mathcal{a}} |E_x^s(\int \Pi_p^s(X_n,dy)P_y^s(A))$$

$$- E_x^\delta(\int \Pi_p^s(X_n,dy)P_y^s(A))| + \sup(P_x^\delta(T > n), P_x^s(T > n))$$

$$\leqslant 3ke^{-bn} + \sup_{A\in\mathcal{a}} |E_x^s(\int \Pi_p^s(X_n;dy)P_y^s(A)) - E_x^\delta(\int \Pi_p^s(X_n,dy)P_y^s(A))|$$

$$\leqslant 3ke^{-bn} + \sup_{A\in\mathcal{a}} \sup_{\substack{z\in\mathcal{H}\\ x\in\mathcal{H}}} |\int \Pi_p^s(x,dy)P_y^s(A) - \int \Pi_p^s(z,dy)P_y^s(A)|$$

$$\leqslant 3ke^{-bn} + 2a\rho^p .$$

On a donc montré que pour tout entier n et tout entier P,

$$\sup_{x\in\mathcal{H}} \sup_{A\in\mathcal{a}} |P_x^\delta(\theta_{n+p}(A)) - P_x^s(\theta_{n+p}(A))| \leqslant 3ke^{-bn} + 2a\rho^p .$$

En particulier, pour tout p supérieur à n :

$$\sup_{x\in\mathcal{H}} \sup_{A\in\mathcal{a}} |P_x^\delta(\theta_{n+p}(A) - P_x^s(\theta_{n+p}(A))| \leqslant 3ke^{-bn} + 2a\rho^n$$

c'est à dire, il existe 2 constantes k_1 et b_1 strictement positives, telles que pour tout p supérieur à n , on ait :

$$\sup_{x\in\mathcal{H}} \sup_{A\in\mathcal{A}} |P_x^\delta(\theta_{n+p}(A)) - P_x^s(\theta_{n+p}(A))| \leqslant k_1 e^{-b_1 n} .$$

En particulier ,

$$\text{pour} \quad p = n , \quad \sup_{x\in\mathcal{H}} \sup_{A\in\mathcal{A}} |P_x^\delta(\theta_{2n}(A) - P_x^s(\theta_{2n}(A))| \leqslant k_1 e^{-\frac{b}{2}2n}$$

$$\text{pour} \quad p = n+1, \quad \sup_{x\in\mathcal{H}} \sup_{A\in\mathcal{A}} |P_x^\delta(\theta_{2n+1}(A)) - P_x^s(\theta_{2n+1}(A))| \leqslant k_1 e^{+\frac{b_1}{2}} e^{-\frac{b_1}{2}(2n+1)}$$

D'où il existe k_2 et b_2 2 constante strictement positives telle que pour tout n de N , on ait :

$$\sup_{x\in\mathcal{H}} \sup_{A\in\mathcal{A}} |P_x^\delta(\theta_n(A)) - P_x^s(\theta_n(A))| \leqslant k_2 e^{-b_2 n}$$

Soit alors n un entier naturel. Pour tout m inférieur à n , pour tout x , on a :

$$P_x^\delta(\frac{1}{n} \sum_{p=0}^{n-1} g(X_p,A_p,X_{p+1}) \leqslant \int gd\, \mu^s \otimes s \otimes \Pi - \varepsilon)$$

$$\leqslant P_x^\delta(\frac{1}{n} \sum_{p=m}^{n-1} g(X_p,A_p,X_{p+1}) \leqslant \int gd\, \mu^s \otimes s \otimes \Pi - \varepsilon + \frac{||g||m}{n}$$

En utilisant le résultat démontré ci-dessus, on obtient pour tout x ,

$$P_x^\delta(\frac{1}{n} \sum_{p=0}^{n-1} g(X_p,A_p,X_{p+1}) \leqslant \int gd\, \mu^s \otimes s \otimes \Pi - \varepsilon)$$

$$\leqslant k_2 e^{-b_2 n} + P_x^s(\frac{1}{n} \sum_{p=m}^{n-1} g(X_p,X_p,X_{p+1}) \leqslant \int gd\, \mu^s \times s \times - \varepsilon + \frac{||g||m}{n})$$

$$\leqslant k_2 e^{-b_2 n} + P_x^s(\frac{1}{n} \sum_{p=0}^{n-1} g(X_p,A_p,X_{p+1}) \leqslant \int gd\, \mu^s \times s \times - \varepsilon + \frac{2||g||m}{n})$$

en particulier prenons $m = \inf (\frac{\varepsilon}{4||g||} , 1)n$, on a :

$$P_x^\delta\left(\frac{1}{n}\sum_{p=0}^{n-1} g(X_p,A_p,X_{p+1}\right) \leq \int gd(\mu^s \times s \times \Pi)-\varepsilon) < (\int gd\ \mu^s \times s \times \Pi - \frac{\varepsilon}{2}) .$$

Puisque la chaîne Π^s est récurrente Doeblin, grâce au théorème 1 de [3] , il existe k_3 et b_3 strictement positifs tels que pour tout x de $\mathcal{H}$ et pour tout n de N , on ait :

$$P_x^s\left(\frac{1}{n}\sum_{p=0}^{n-1} g(X_p,A_p,X_{p+1}\right) \leq \int gd\ \mu^s \times s \times \Pi - \frac{\varepsilon}{2}) \leq k_3\, e^{-b_3 n}$$

Ce qui permet de conclure qu'il existe k et b strictement positifs, tels que pour tout n , on ait :

$$\sup_{x\in\mathcal{H}} P_x^\delta\left(\frac{1}{n}\sum_{p=0}^{n-1} g(X_p,A_p,X_{p+1}\right) \leq \int gd\ \mu^s \times s \times \Pi - \varepsilon) \leq k\, e^{-bn}$$

c'est à dire :

$$\overline{\lim_n}\ \frac{1}{n} \text{Log} \sup_{x\in\mathcal{H}} P_x^\delta\left(\frac{1}{n}\sum_{p=0}^{n-1} g(X_p,A_p,X_{p+1}\right) \leq \int gd\ \mu^s \times s \times \Pi - \varepsilon) \leq 0 .$$

1.8. Optimalité de la stratégie $\hat{s}$

Théorème 7

Pour tout θ, $E_\theta^{\hat{s}}(\nu_c) \leq \dfrac{-(1+\sigma(1))\text{Log } c}{I(\theta)}$.

Démonstration

On suit celle du théorème 3.b de [3] .

Soit ε strictement positif et $N_c = \dfrac{-(1+\frac{\varepsilon}{2})\text{Log } c}{I(\theta)}$

$$E_\theta^{\hat{s}}(\nu_c) \leq E_\theta^{\hat{s}}(\nu_c\, 1_{(\nu_c<N_c)}) + E_\theta^{\hat{s}}(\nu_c\, 1_{(\nu_c>N_c)})$$

$$E_\theta^{\hat{s}}(\nu_c) \leq N_c + \sum_{n>N_c} n\, P_\theta^{\hat{s}}(\nu_c = n) .$$

Montrons qu'il existe k' et b' strictement positifs tels que pour tout $n > N_c$, on ait : $P_\theta^{\hat{s}}(\nu_c = n) \leq k' e^{-b'n}$.

On a : $\{\nu_c = n\} \subset \{T_\theta > n\} \cup \{\nu_{c,\phi_\theta} = n\}$.

s^θ est récurrente doeblin, apériodique, et "exponentiellement voisine" de $\hat{s}$.
Le théorème 1.7.1. appliqué à la v.a. :

$$g(\theta,x,a,y) = \text{Log}\,\frac{P(\theta,x,a,y)}{P(\phi_\theta,x,a,y)}$$

entraîne qu'il existe k_2' et b_2' strictement positifs tels que pour tout n de N , on ait :

$$P_\theta^{\hat{s}}(S_n(\theta,\phi_\theta) \leq n \frac{I(\theta)}{1+\frac{\varepsilon}{2}}) \leq k_2' e^{-b_2'n} .$$

La relation : $P_\theta^{\hat{s}}(\nu_{c,\phi_\theta} = n) < P_\theta^{\hat{s}}(S_{n-1}(\theta,\phi_\theta) < -\log c)$ permet d'écrire qu'il existe k_2'' et b_2'' strictement positifs, tels que pour tout $n > N_c$, on ait :

$$P_\theta^{\hat{s}}(\nu_{c,\phi_\theta} = n) \leq k_2'' e^{-b_2''n}$$

Cette dernière relation et la consistance exponentielle de la suite $(\hat{\theta}_n)_{n\in N}$ permettant de conclure qu'il existe k' et b' strictement positifs tels que pour tout $n > N_c$, on ait :

$$E_\theta^{\hat{s}}(\nu_c) \leq N_c + k' \sum_{n>N_c} n e^{-b'n}$$

On trouve alors facilement la démonstration, comme par exemple au théorème 3.2 de [3].

Le lemme 8 et ce dernier résultat impliquent :

Théorème 8

La procédure $\phi_c^{\hat{s}}$ est asymptotiquement la plus économique parmi les procédures de force au moins égale, pour le test de 2 hypothèses simples.

1.9. Exemple : Bandit à 2 bras markovien

On dispose de 2 machines A et B ; les observations possibles après la mise en marche de A ou B sont 0 et 1. Pour chaque machine, l'ensemble $\{0,1\}$ est une chaîne markovienne. Pour l'une, $\Pi(0,1) = \alpha_0$ et $\Pi(1,0) = \alpha_1$; pour l'autre, $\Pi(0,1) = \beta_0$ et $\Pi(1,0) = \beta_1$. Les α_i et β_i (pour i valant 0 ou 1) sont connus, mais les machines ne sont pas identifiées (les fiches techniques sont mélangées ...) : on se pose le problème de savoir à quelle machine correspond le couple (α_0, α_1). Le problème de test que l'on se pose est donc un test de 2 hypothèses simples. L'espace des paramètres est un sous ensemble de R^2 (la première coordonnée correspondra à la machine A) :

$$\Theta = \{(\alpha,\beta), (\alpha,\beta)\} .$$

On note "A" |"B"| l'action de mettre la machine A|B| en marche.

a) <u>Les transitions</u> : Pour i,j valant 0 ou 1 ,

$$\begin{aligned} \Pi((\alpha,\beta), i, \text{"A"}, j) &= \alpha_i \quad \text{si} \quad i \neq j \\ &= 1-\alpha_i \quad \text{si} \quad i = j \end{aligned}$$

$$\begin{aligned} \Pi((\alpha,\beta), i, \text{"B"}, j) &= \beta_i \quad \text{si} \quad i \neq j \\ &= 1-\beta_i \quad \text{si} \quad i = j \end{aligned}$$

On pose pour i valant 0 ou 1

$$m_i = \alpha_i \operatorname{Log} \frac{\alpha_i}{\beta_i} + (1 - \alpha_i) \operatorname{Log} \frac{1-\alpha_i}{1-\beta_i}$$

$$p_i = \beta_i \operatorname{Log} \frac{\beta_i}{\alpha_i} + (1 - \beta_i) \operatorname{Log} \frac{1-\beta_i}{1-\alpha_i}$$

b) <u>L'information</u>

$$I((\alpha,\beta), (\beta,\alpha), i, \text{"A"}) = m_i = I((\beta,\alpha), (\alpha,\beta), i, \text{"B"})$$

$$I((\alpha,\beta), (\beta,\alpha), i, \text{"B"}) = p_i = I((\beta,\alpha), (\alpha,\beta), i, \text{"A"})$$

c) Les stratégies déterministes sont au nombre de 4 : On les note :

* $s_1 : s_1(0) = s_1(1) = $ "A"

$s_2 : s_2(0) = s_2(1) = $ "B"

$s_3 : s_3(0) = $ "A", $s_3(1) = $ "B"

$s_4 : s_4(0) = $ "B", $s_4(1) = $ "A"

* On obtient pour les probabilités invariantes des chaînes Π^{s_i} :

$$\mu^{s_1}_{(\alpha,\beta)}(1) = \frac{\alpha_o}{\alpha_o+\alpha_1} \quad , \quad \mu^{s_1}_{(\beta,\alpha)}(1) = \frac{\beta_o}{\beta_o+\beta_1}$$

$$\mu^{s_2}_{(\alpha,\beta)}(1) = \frac{\beta_o}{\beta_o+\beta_1} \quad , \quad \mu^{s_2}_{(\beta,\alpha)}(1) = \frac{\alpha_o}{\alpha_o+\alpha_1}$$

$$\mu^{s_3}_{(\alpha,\beta)}(1) = \frac{\alpha_o}{\alpha_o+\alpha_1} \quad , \quad \mu^{s_3}_{(\beta,\alpha)}(1) = \frac{\beta_o}{\beta_o+\alpha_1}$$

$$\mu^{s_4}_{(\alpha,\beta)}(1) = \frac{\beta_o}{\beta_o+\alpha_1} \quad , \quad \mu^{s_4}_{(\beta,\alpha)}(1) = \frac{\alpha_o}{\alpha_o+\beta_1}$$

* On note I_i l'information $I((\alpha,\beta), (\beta,\alpha), s_i)$ et I^*_i l'information $I((\beta,\alpha), (\alpha,\beta), s_i)$. On obtient :

$$I_1 = \frac{\alpha_o}{\alpha_o+\alpha_1} m_1 + \frac{\alpha_1}{\alpha_o+\alpha_1} m_o \qquad \text{et} \qquad I_3 = \frac{\alpha_o}{\alpha_o+\beta_1} p_1 + \frac{\beta_1}{\alpha_o+\beta_1} m_o$$

$$I^*_1 = \frac{\alpha_o}{\alpha_o+\alpha_1} p_1 + \frac{\alpha_1}{\alpha_o+\alpha_1} p_o \qquad \text{et} \qquad I^*_3 = \frac{\alpha_o}{\alpha_o+\beta_1} m_1 + \frac{\beta_1}{\alpha_o+\beta_1} p_o$$

BIBLIOGRAPHIE

[1] CHERNOFF H. : Sequential design of experiments - Ann. Math. Stat., Vol. 30, 755-770 (1959) .

[2] GEORGIN J.P. : Estimation et contrôle des chaînes de Markov sur des espaces arbitraires - Journées de statistique des processus stochastiques - Proceedings, Grenoble - Lecture notes in Mathematics, 636 (juin 1977) .

[3] MAIGRET N. : Majorations de Chernoff et statistique séquentielle pour des chaînes de Markov récurrentes au sens de Doeblin - Séminaires d'Orsay - Astérisque.

[4] MAIGRET N. : Majorations de Chernoff pour des chaînes de Markov contrôlées. A paraître , Zeitchrift, 1979 .

Nelly MAIGRET
Université Paris Nord
Mathématiques
ERA CNRS 532 "Statistique Appliquée"
Avenue J.B. Clément
93430 VILLETANEUSE