Séminaire de probabilités (Strasbourg)

Jean Jacod

Une application de la topologie d'Émery : le processus information d'un modèle statistique filtré

Séminaire de probabilités (Strasbourg), tome 23 (1989), p. 448-474

<http://www.numdam.org/item?id=SPS_1989__23__448_0>

UNE APPLICATION DE LA TOPOLOGIE D'EMERY: LE PROCESSUS INFORMATION D'UN MODELE STATISTIQUE FILTRE

Jean JACOD

Laboratoire de Probabilités
Université Pierre et Marie Curie
Tour 56 (3^e étage), 4 Place Jussieu
75252 PARIS Cedex 05

1 - INTRODUCTION

Soit $(\Omega,\mathcal{F},(P_\theta)_{\theta\in\Theta})$ un modèle statistique, avec Θ ouvert de $\mathbb{R}^d$ contenant 0. Rappelons, d'après LeCam [6] (voir aussi Strasser [9] ou Dacunha-Castelle et Duflo [2] pour une situation plus classique), comment on définit l'information de Fisher, sous des conditions "minimales" de régularité.

Supposons d'abord le modèle dominé par une probabilité Q (i.e. $P_\theta \ll Q$ pour tout $\theta\in\Theta$), et posons $z^\theta = dP_\theta/dQ$. Si $\theta \to \sqrt{z^\theta}$ est différentiable dans $L^2(Q)$ au point $\theta=0$, avec $W=(W^i)_{i\leq d}$ pour variable variable dérivée, alors W se met sous la forme $W=(\sqrt{z^0}/2)V$, chaque composante V^i de V étant centrée et de carré intégrable. La matrice d'information de Fisher en $\theta=0$ est alors définie par

1.1 $$I^{ij} = E_{P_0}(V^iV^j), \quad 1 \leq i,j \leq d.$$

De plus, soit $Z^\theta = z^\theta/z^0$ ($0/0=0$, $a/0=\infty$ si $a>0$) la dérivée (généralisée) dP_θ/dP_0; alors $\theta \to \sqrt{Z^\theta}$ est différentiable dans $L^2(P_0)$ en $\theta=0$, avec $V/2$ pour dérivée. Enfin si

1.2 $$H^{\theta\zeta} = E_Q(\sqrt{z^\theta z^\zeta})$$

désigne l'intégrale de Hellinger ("affinité") de P_θ et P_ζ (cette expression ne dépend pas de la mesure Q dominant P_θ et P_ζ), la différentiabilité de $\sqrt{z^\theta}$ en $\theta=0$ dans $L^2(Q)$ équivaut à

1.3 $$1 + H^{\theta\zeta} - H^{0\theta} - H^{0\zeta} = \theta.J.\zeta + o(|\theta||\zeta|) \quad \text{quand } \theta,\zeta \to 0,$$

où J est une certaine matrice, et dans ce cas $J=I/4$ (la notation $\theta.J.\zeta$ désigne $\sum_{i,j\leq d} \theta^i J^{ij} \zeta^j$, et de même $\theta.\zeta$ désignera le produit scalaire usuel de θ et ζ dans $\mathbb{R}^d$).

Lorsque le modèle n'est pas dominé, on peut toujours définir les intégrales de Hellinger $H^{\theta\zeta}$, et si 1.3 est satisfaite on appelle information de Fisher la matrice $I=4J$.

On va étudier ce qui se passe quand le modèle est muni d'une filtration $(\mathcal{F}_t)_{t\geq 0}$. Sous des hypothèses convenables, on peut définir la matrice d'information I_t en $\theta=0$ pour chaque modèle $(\Omega,\mathcal{F}_t,(P_\theta))$. Il semble plus judicieux de définir un processus information, déjà connu dans des cas particuliers, notamment lorsque la filtration est à temps discret: cf. Basawa et Prakasa Rao [1]. Ce processus $\Lambda = (\Lambda_t)_{t\geq 0}$, à valeurs matricielles $d\times d$, vérifie $I_t = E_{P_0}(\Lambda_t)$ dès que I_t existe, et dans de nombreux cas il est aisément calculable en fonction des caractéristiques du "processus de base" du modèle.

De même les intégrales de Hellinger sont remplacées par les processus de Hellinger, et le résultat principal de cet article est que le processus information est la "dérivée seconde" du processus de Hellinger, au sens où on a une relation analogue à 1.3. Il s'agit même d'une condition nécessaire et suffisante, dans le sens où la dérivabilité du processus de Hellinger entraine l'existence du processus information (voir les énoncés précis au §6).

Après des compléments sur la topologie d'Emery (§2), on introduit la dérivabilité des modèles filtrés (§3), puis la dérivabilité "locale" (§4). On rappelle quelques notions sur les processus de Hellinger au §5. Les résultats principaux sont énoncés et démontrés au §6, et le §7 est consacré à quelques exemples simples: en utilisant la forme "explicite" des processus de Hellinger donnée au §IV-3 de [5], le lecteur pourra écrire lui-même des exemples plus élaborés.

Nous utilisons les notations usuelles de la théorie générale. En particulier $H\bullet X$ désigne le processus intégrale stochastique du processus prévisible H par rapport à la semimartingale X, dont la partie martingale continue est notée X^c. On note aussi X^T le processus X arrêté en T: $X_t^T = X_{T\wedge t}$.

2 - LA TOPOLOGIE D'EMERY: QUELQUES COMPLEMENTS

On va étendre légèrement certains résultats de Mémin [8] sur la topologie des semimartingales qui a été introduite par Emery (et pour laquelle nous référons à [3]).

Soit X^n, X des processus définis sur une base stochastique $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},P)$. On utilise les notations $X_t^* = \sup_{s\leq t}|X_s|$, et

- $X^n \xrightarrow{P-U} X$ si $(X^n-X)_t^* \xrightarrow{P} 0$ $\forall t>0$;
- $X^n \xrightarrow{\mathcal{S}(P)} X$ si X^n converge vers X pour la topologie d'Emery (quand X^n, X sont des semimartingales);
- $X^n \xrightarrow{P-V} X$ si $\mathrm{Var}(X^n-X)_t \xrightarrow{P} 0$ $\forall t>0$ (quand X^n, X sont à variation finie; $\mathrm{Var}(Y)$ désigne le processus variation de Y).

Si X est une semimartingale spéciale, on appelle <u>compensateur</u> de X l'unique processus prévisible à variation finie A nul en 0, tel que $X-A-X_0$ soit une martingale locale.

Lorsque les X^n ont des sauts uniformément bornés, le résultat suivant est dû à Mémin.

2.1 THEOREME. <u>On suppose que les semimartingales</u> X^n <u>vérifient</u>

2.2 $\quad$ $\forall t>0$, la suite $\{(\Delta X^n)_t^*\}_{n\geq 1}$ est uniformément intégrable,

<u>et</u> $X^n \xrightarrow{\mathcal{S}(P)} X$. <u>Alors</u> X^n, X <u>sont spéciales, et leurs compensateurs</u> A^n <u>et</u> A <u>vérifient</u> $A^n \xrightarrow{P-V} A$.

<u>Preuve</u>. La première assertion découle de ce que toute semimartingale X vérifiant $\Delta X_t^* \in L^1(P)$ pour tout t est spéciale. Il reste à montrer que $\mathrm{Var}(A^n-A)_t \xrightarrow{P} 0$ pour tout t. On fixe donc t, et on peut

clairement se limiter à des sous-suites.

a) On va d'abord montrer qu'on peut trouver une suite croissante (R_p) de temps d'arrêt avec $\lim_{p\uparrow\infty} P(R_p<t) = 0$, et une sous-suite notée encore X^n, telles que $[X^n-X^m,X^n-X^m]_{R_p}^{1/2} \longrightarrow 0$ dans L^1 lorsque $n,m\uparrow\infty$, pour chaque entier p. Pour cela, on reprend essentiellement la fin de la preuve du lemme II.4 de Mémin [8]. D'après ce lemme, on sait d'ailleurs déjà que, quitte à prendre une sous-suite,

2.3 $$[X^n-X^m,X^n-X^m]_t \xrightarrow{P} 0, \quad \text{quand } n,m\uparrow\infty.$$

Posons $V^n = (\Delta X^n)^*_t$. D'après 2.2, pour tout entier N il existe $\varepsilon(N)\in]0,4^{-N}]$ tel que

2.4 $$P(A) \le \varepsilon(N) \quad \Rightarrow \quad E(V^n 1_A) \le 4^{-N} \quad \forall n.$$

D'après 2.3, il existe une sous-suite notée encore X^n telle que si $Z^n = [X^n-X^{n+1},X^n-X^{n+1}]^{1/2}$, on ait

2.5 $$P(Z^n_t > 4^{-n}) \le \varepsilon(n).$$

Soit $Y = \sum_{n\ge 1} 2^n Z^n$. Comme $\sum\varepsilon(n) < \infty$, le lemme de Borel-Cantelli et 2.5 impliquent $Y_t<\infty$ p.s., donc les $R_p = t\wedge\inf(s: Y_s>p)$ vérifient $P(R_p<t) \to 0$. On a aussi

$$E(Y_{R_p}) \le p + E(\Delta Y_{R_p}) \le p + \sum 2^n E(\Delta Z^n_{R_p})$$
$$\le p + \sum 2^n E[4^{-n} + \Delta Z^n_{R_p} 1_{\{Z^n_t>4^{-n}\}}].$$

Or $\Delta Z^n \le |\Delta X^n - \Delta X^{n+1}| \le V^n + V^{n+1}$, donc 2.4 et 2.5 entrainent

$$E(Y_{R_p}) \le p + \sum 2^{-n} + \sum 2^n E[(V^n_t+V^{n+1}_t) 1_{\{Z^n_t>4^{-n}\}}]$$
$$\le p + \sum 2^{-n} + \sum 2^n.2.4^{-n} \le p+3.$$

Comme $[X^n-X^m,X^n-X^m]_{R_p}^{1/2} \le Z^n_{R_p}+\dots+Z^{n+m-1}_{R_p} \le Y_{R_p}(2^n+\dots+2^{n+m-1})$, on en déduit le résultat annoncé.

b) Soit $X^n = X^n_0 + M^n + A^n$ et $X = X_0+ M + A$ les décompositions canoniques. D'après Lépingle [7] il existe une constante universelle c telle que $E([M^n-M^m,M^n-M^m]_S^{1/2}) \le cE([X^n-X^m,X^n-X^m]_S^{1/2})$ pour tout temps d'arrêt S. Vu (a), la suite $\{(M^n)^{R_p}\}_{n\ge 1}$ de martingales (arrêtées en R_p) est de Cauchy dans l'espace $\mathcal{H}^1(P)$ de martingales. Elle converge donc vers une limite $N(p)$ dans $\mathcal{H}^1(P)$, et a-fortiori dans $\mathcal{S}(P)$. Donc $(A^n)^{R_p}$ converge dans $\mathcal{S}(P)$ vers une limite $A(p)$; quitte à prendre encore une sous-suite on peut supposer que $(A^n)^{R_p} \to A(p)$ uniformément sur $[0,t]$, pour presque tout ω. Donc $A(p)$ est prévisible; comme limite dans $\mathcal{S}(0)$ de processus prévisibles à variation finie, $A(p)$ est aussi à variation finie. Donc $(A^n)^{R_p} \xrightarrow{P-V} A(p)$.

Comme $X^{R_p} = X_0 + N(p) + A(p)$, on en déduit aussi $A(p)=A^{R_p}$, donc $(A^n)^{R_p} \xrightarrow{P-V} A^{R_p}$. Comme $\lim_p P(R_p<t) \to 0$, il vient $\mathrm{Var}(A^n-A)_t \xrightarrow{P} 0$. □

2.6 THEOREME. <u>Soit</u> X^n <u>une suite de semi-martingales d-dimensionnelles convergeant dans</u> $\mathcal{S}(P)$ <u>vers une limite</u> X^∞. <u>Soit</u> f^n <u>une suite de fonctions sur</u> $\mathbb{R}^d$ <u>de classe</u> C^2, <u>convergeant localement uniformément, ainsi que leurs dérivées partielles d'ordre</u> 1 <u>et</u> 2 <u>vers une fonction</u> f^∞ <u>et ses dérivées. Alors</u> $f^n(X^n) \xrightarrow{\mathcal{S}(P)} f^\infty(X^\infty)$.

<u>Preuve</u>. D'après la formule d'Ito, $f^n(X^n) = f^n(X_0^n)+\alpha^n+\beta^n+\gamma^n$, où
$\alpha^n = \sum_{i\le d} D_i f^n(X_-^n)\cdot X^{n,i}$, $\beta^n = \frac{1}{2}\sum_{i,j\le d} D_{ij} f^n(X_-^n)\cdot C^{n,ij}$, $\gamma^n = \sum_{s\le .} \delta_s^n$,
avec

$$C^{n,ij} = \langle X^{n,c,i}, X^{n,c,j}\rangle,$$

$$\delta_s^n = f^n(X_s^n) - f^n(X_{s-}^n) - \sum_{i\le d} D_i f^n(X_{s-}^n)\Delta X_s^{n,i}.$$

Si g^n converge localement uniformément vers une fonction continue g^∞, comme $X^n \xrightarrow{P-U} X^\infty$ il est clair que $g^n(X_-^n) \xrightarrow{P-U} g^\infty(X_-^\infty)$. On déduit alors de Mémin [8, III.13] que $g^n(X_-^n)\cdot X^{n,i} \xrightarrow{\mathcal{S}(P)} g^\infty(X_-^\infty)\cdot X^{\infty,i}$. Par suite, $\alpha^n \xrightarrow{\mathcal{S}(P)} \alpha^\infty$. Comme l'hypothèse entraine aussi que $C^{n,ij} \xrightarrow{\mathcal{S}(P)} C^{\infty,ij}$, on a $\beta^n \xrightarrow{\mathcal{S}(P)} \beta^\infty$.

Toujours parce que $X^n \xrightarrow{P-U} X^\infty$, on voit comme ci-dessus que $\delta^n \xrightarrow{P-U} \delta^\infty$. Par ailleurs il existe des constantes K_N (dépendant des f^n, mais pas de n) telles que $|\delta^n| \le K_N|\Delta X^n|^2$ si $|\Delta X^n|\le N$. On sait aussi que $\sum_{s\le .} |\Delta X_s^n|^2 \xrightarrow{P-V} \sum_{s\le .} |\Delta X_s^\infty|^2$. Par suite $\gamma^n \xrightarrow{P-V} \gamma^\infty$, et a-fortiori $\gamma^n \xrightarrow{\mathcal{S}(P)} \gamma^\infty$. Finalement $f^n(X_0^n) \xrightarrow{P} f^\infty(X_0^\infty)$. □

3 - DERIVABILITE DES MODELES STATISTIQUES FILTRES

On considère un modèle filtré $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},(P_\theta)_{\theta\in\Theta})$, où Θ est un ouvert de $\mathbb{R}^d$ contenant 0. On s'intéresse ici à la dérivabilité en $\theta=0$, et pour simplifier on écrit $P=P_0$. Sauf mention explicite du contraire, toutes les notions usuelles (martingales, semimartingales, etc...) sont relatives à P. Pour simplifier on suppose aussi que toutes les mesures P_θ coïncident sur $\mathcal{F}_0$.

On note Z^θ le processus densité de P_θ par rapport à $P=P_0$. C'est une surmartingale positive, avec $Z_0^\theta=1$. C'est une martingale si et seulement si $P_\theta \overset{loc}{\ll} P$. On a $Z^0=1$. Enfin si Q est une probabilité dominant P_0 et P_θ et si z^θ et z^0 désignent les processus densité de P_θ et P_0 par rapport à Q, on a $Z^\theta = z^\theta/z^0$ P-p.s.

Pour tout temps d'arrêt T on définit l'intégrale de Hellinger P_θ et P_ζ à l'instant T par

$$H_T^{\theta\zeta} = E_Q(\sqrt{z_T^\theta z_T^\zeta}), \tag{3.1}$$

où Q est une probabilité dominant P_θ et P_ζ et z^θ, z^ζ sont les processus densité correspondant. $H_T^{\theta\zeta}$ ne dépend pas du choix de Q.

Nous ne supposons pas le modèle dominé par une probabilité Q, si bien que la notion de différentiabilité est "du type 1.3":

3.2 DEFINITION. Soit T un temps d'arrêt. On dit que le modèle est <u>dérivable à l'instant</u> T (en $\theta=0$) s'il existe une matrice $d\times d$ (nécessairement symétrique non-négative) J_T telle que

$$1 + H_T^{\theta\zeta} - H_T^{0\theta} - H_T^{0\zeta} = \theta.J_T.\zeta + o(|\theta||\zeta|) \quad \text{quand} \quad \theta,\zeta \to 0. \quad \square \tag{3.3}$$

3.4 THEOREME. <u>Soit</u> T <u>un temps d'arrêt. Pour que le modèle soit dérivable à l'instant</u> T <u>en</u> $\theta=0$, <u>il faut et il suffit que</u>

$$E(1-Z_T^\theta) = o(|\theta|^2) \quad \text{quand} \quad \theta \to 0, \tag{3.5}$$

$$E[(\sqrt{Z_T^\theta} - 1 - \tfrac{1}{2}\,\theta.V_T)^2] = o(|\theta|^2) \quad \text{quand} \quad \theta \to 0, \tag{3.6}$$

<u>où</u> $V_T=(V_T^i)_{i\leq d}$ <u>est une variable aléatoire à valeurs dans</u> $\mathbb{R}^d$. <u>Dans ce cas,</u> V_T <u>est</u> $\mathcal{F}_T$<u>-mesurable, centrée, de carré intégrable, et la matrice</u> J_T <u>apparaissant dans 3.3 vaut</u> $J_T^{ij} = E(V_T^i V_T^j)/4$.

Comme $\sqrt{Z_T^\theta} \in L^2(P)$ pour tout θ, 3.6 signifie que $\theta \to \sqrt{Z_T^\theta}$ est différentiable dans $L^2(P)$ au point $\theta=0$, de dérivée $V/2$. Lorsque les P_θ sont toutes équivalentes (sur $\mathcal{F}_T$), 3.5 est automatique et ce résultat est classique.

Il y a une autre manière, souvent plus commode, d'exprimer 3.5 et 3.6. Posons d'abord:

3.7 $\mathcal{T}$ = ensemble des termes (θ_n,u_n,θ), où θ_n est une suite d'éléments de Θ tendant vers 0, avec $u_n=|\theta_n|$ et telle que θ_n/u_n converge vers une limite θ dans $\mathbb{R}^d$.

Alors, 3.5 et 3.6 sont respectivement <u>équivalents à</u>:

3.8 $$E(1-Z_T^{\theta_n})/u_n^2 \longrightarrow 0, \quad \forall(\theta_n,u_n,\theta)\in\mathcal{T},$$

3.9 $$[(Z_T^{\theta_n})^{1/2} - 1]/u_n \xrightarrow{L^2(P)} \frac{1}{2}\,\theta.V_T, \quad \forall(\theta_n,u_n,\theta)\in\mathcal{T}.$$

<u>Preuve</u>. a) Supposons d'abord la dérivabilité, et montrons 3.8 (=3.5). Soit $(\theta_n,u_n,\theta)\in\mathcal{T}$, et Q une probabilité dominant les P_ζ pour tout ζ dans $I = \{0,\theta_n,-\theta_n\colon n\geq 1\}$, z^ζ le processus densité de P_ζ par rapport à Q, et $Y^\zeta = (\sqrt{z_T^\zeta} - \sqrt{z_T^0})/|\zeta|$ pour $\zeta\in I\setminus\{0\}$. 3.1 implique que pour $\rho,\zeta\in I\setminus\{0\}$:

$$E_Q(Y^\rho Y^\zeta) = \frac{1}{|\rho||\zeta|}(1 + H_T^{\rho\zeta} - H_T^{0\rho} - H_T^{0\zeta}).$$

3.3 entraine $E_Q(Y^{\theta_n}Y^{\theta_m}) \to a := \theta.J_T.\theta$ si $n,m\uparrow\infty$, donc $Y^{\theta_n} \xrightarrow{L^2(Q)} W$ pour une variable W vérifiant $E_Q(W^2)=a$. De même $Y^{-\theta_n} \xrightarrow{L^2(Q)} W'$ avec $E_Q(W'^2)=a$. Enfin $E_Q(Y^{\theta_n}Y^{-\theta_n}) \to a$, donc $E_Q(WW') = -a$; par suite $W+W'=0$ Q-p.s. Comme $Y^{\theta_n}\geq 0$ et $Y^{-\theta_n}\geq 0$ sur l'ensemble $\{z_T^0=0\}$, on a $W\geq 0$ et $W'\geq 0$, d'où $W=W'=0$ sur $\{z_T^0=0\}$. En particulier,

3.10 $$\frac{1}{u_n^2}\, E_Q(z_T^{\theta_n}\, 1_{\{z_T^0=0\}}) = E_Q[(Y^{\theta_n})^2\, 1_{\{z_T^0=0\}}] \longrightarrow 0.$$

Mais pour tout $\zeta \in I$ on a

3.11 $$E_Q(z_T^\zeta 1_{\{z_T^0=0\}}) = 1 - E_Q(z_T^\zeta 1_{\{z_T^0>0\}}) = E(1-Z_T^\zeta),$$

donc 3.10 entraine 3.8.

b) Pour obtenir la condition nécessaire et suffisante, il suffit alors de montrer que, sous 3.5, la dérivabilité du modèle équivaut à la dérivabilité de $\theta \to \sqrt{Z_T^\theta}$ dans $L^2(P)$, en $\theta=0$. Si $C(\theta,\zeta) = E(\sqrt{Z_T^\theta Z_T^\zeta})$, cette dernière propriété équivaut classiquement à

3.12 $$C(0,0) + C(\theta,\zeta) - C(0,\theta) - C(0,\zeta) = \theta.J'.\zeta + o(|\theta||\zeta|)$$

pour une certaine matrice J', lorsque $\theta,\zeta \to 0$. Dans ce cas, $J'^{ij} = E(V_T^i V_T^j)/4$, où $V_T/2$ est la dérivée de $t \to \sqrt{Z_T^\theta}$, donc si $J'=J_T$ on aura aussi la dernière assertion de l'énoncé.

Si Q est une probabilité dominant P, P_θ, P_ζ, on a avec les notations usuelles z^0, z^θ, z^ζ:

$$C(\theta,\zeta) = E_Q(\sqrt{z_T^\theta z_T^\zeta}\, 1_{\{z_T^0>0\}}) = H_T^{\theta\zeta} - E_Q(\sqrt{z_T^\theta z_T^\zeta}\, 1_{\{z_T^0=0\}}).$$

Donc, par l'inégalité de Schwarz, puis 3.11, puis 3.5:

$$|C(\theta,\zeta) - H_T^{\theta\zeta}| \le \{E_Q(z_T^\theta 1_{\{z_T^0=0\}})\, E_Q(z_T^\zeta 1_{\{z_T^0=0\}})\}^{1/2}$$
$$= [E(1-Z_T^\theta)\, E(1-Z_T^\zeta)]^{1/2} = o(|\theta||\zeta|).$$

Il est alors évident que la différence entre les premiers membres de 3.12 et 3.3 est $o(|\theta||\zeta|)$. On en déduit l'équivalence de 3.12 et 3.3, avec $J'=J_T$.

c) Il reste à montrer que V_T est centrée. Soit $\zeta_n=e_i/n$, où e_i est le $i^{\text{ième}}$ vecteur de base de $\mathbb{R}^d$. 3.9 implique $n[(Z_T^{\zeta_n})^{1/2}-1] \to V_T^i/2$ dans $L^2(P)$, donc a-fortiori $n(Z_T^{\zeta_n}-1) \to V_T^i$ dans $L^1(P)$, et $E(V_T^i)=0$ découle alors de 3.5. $\square$

Lorsqu'un modèle non filtré $(\Omega,\mathcal{F},(P_\theta)_{\theta\in\Theta})$ est dérivable il est bien connu (au moins lorsque les P_θ sont toutes équivalentes: cf. Ibragimov et Khashminski [4], I-7-2) que le modèle restreint à une sous-tribu de $\mathcal{F}$ est aussi dérivable. Le théorème suivant étend ce résultat, avec en plus une évaluation "uniforme".

3.13 PROPOSITION. Supposons le modèle dérivable à l'instant T pour un certain temps d'arrêt T. Pour tout temps d'arrêt $S \leq T$ on pose $V_S = E(V_T | \mathcal{F}_S)$ (où V_T vérifie 3.6). On a alors pour $(\theta_n, u_n, \theta) \in \mathcal{C}$:

3.14 $$\sup_{S \text{ temps d'arrêt, } S \leq T} E(1 - Z_S^{\theta_n}) / u_n^2 \longrightarrow 0,$$

3.15 $$\sup_{S \text{ temps d'arrêt, } S \leq T} E[|((Z_S^{\theta_n})^{1/2} - 1)/u_n - \tfrac{1}{2}\,\theta.V_S|^2] \longrightarrow 0.$$

Le modèle est donc dérivable en S pour tout temps d'arrêt $S \leq T$.

Par "recollement", on obtient le

3.16 COROLLAIRE. Supposons qu'il existe une suite (T_p) de temps d'arrêt croissant vers $+\infty$, et que le modèle soit dérivable en chaque instant T_p. Il existe alors une martingale locale, localement de carré intégrable V à valeurs dans $\mathbb{R}^d$, avec $V_0 = 0$, telle qu'on ait 3.6 pour tout temps d'arrêt T inférieur ou égal à l'un des T_p.

Avant de démontrer 3.13, énonçons un lemme dans lequel X_n, Y sont des variables sur $(\Omega, \mathcal{F}, P)$ avec $X_n \geq 0$ et u_n des réels positifs tendant vers 0.

3.17 LEMME. Si $(\sqrt{X_n} - 1)/u_n \to Y$ dans $L^2(P)$, on a

$$\sup_{\mathcal{G} \text{ sous-tribu de } \mathcal{F}} \| (\sqrt{E(X_n | \mathcal{G})} - 1)/u_n - E(Y | \mathcal{G}) \|_2 \longrightarrow 0.$$

Preuve. Posons $U_n = (\sqrt{X_n} - 1)/u_n$, $V_n = (X_n - 1)/u_n - 2Y$. Par hypothèse, $a_n := E[(U_n - Y)^2] \to 0$, ce qui implique a-fortiori $b_n := E(|V_n|) \to 0$.

Fixons la sous-tribu $\mathcal{G}$ de $\mathcal{F}$. Soit $X'_n = E(X_n | \mathcal{G})$, $Y' = E(Y | \mathcal{G})$, $U'_n = (\sqrt{X'_n} - 1)/u_n$, $V'_n = (X'_n - 1)/u_n - 2Y'$ et $W'_n = |U'_n - Y'|$. On a $E[|V'_n|] \leq b_n$, donc $P(|V'_n| > \varepsilon) \leq b_n/\varepsilon$.

Si $k \geq 1$, $\varepsilon \leq 1$, $|y| \leq k$ et $0 \leq u \leq \varepsilon/y^2$, il est facile de voir que

$|(\sqrt{x}-1)/u - y| > \varepsilon$ implique $|(x-1)/u - 2y| > \varepsilon$. Donc si $k\geq 1$, $\varepsilon\leq 1$ on obtient $P(W'_n>\varepsilon) \leq P(|V'_n|>\varepsilon) + P(|Y'|>k)$ dès que $u_n\leq\varepsilon/k^2$. Comme $E(Y'^2) \leq E(Y^2)$, il vient alors

3.18 $$u_n \leq \varepsilon/k^2 \Rightarrow P(W'_n > \varepsilon) \leq b_n/\varepsilon + E(Y^2)/k^2.$$

Par ailleurs $Y'^2 \leq E(Y^2|\mathcal{G})$ et $U'^2_n \leq E(U^2_n|\mathcal{G})$, donc

$$\begin{aligned}
E(W'^2_n) &\leq \varepsilon^2 + E(W'^2_n 1_{\{W'_n>\varepsilon\}}) \leq \varepsilon^2 + 2E[(Y'^2+U'^2_n)1_{\{W'_n>\varepsilon\}}]. \\
&\leq \varepsilon^2 + 2E(Y^2 1_{\{W'_n>\varepsilon\}}) + 2E[U^2_n 1_{\{W'_n>\varepsilon\}}] \\
&\leq \varepsilon^2 + 2E(Y^2 1_{\{W'_n>\varepsilon\}}) + 4E[(Y^2 + (U_n-Y)^2)1_{\{W'_n>\varepsilon\}}] \\
&\leq \varepsilon^2 + 6E(Y^2 1_{\{W'_n>\varepsilon\}}) + 4a_n \\
&\leq \varepsilon^2 + 6N^2P(W'_n>\varepsilon) + 6E(Y^2 1_{\{Y>N\}}) + 4a_n.
\end{aligned}$$

Si alors $u_n\leq\varepsilon/k^2$, on déduit de 3.18 que

$$E(W'^2_n) \leq \varepsilon^2 + 6N^2(b_n/\varepsilon + E(Y^2)/k^2) + 6E(Y^2 1_{\{Y>N\}}) + 4a_n.$$

Le membre de droite ci-dessus, qui ne dépend pas de $\mathcal{G}$, peut être rendu inférieur à $2\varepsilon^2$ pour tout n assez grand (choisir d'abord N puis k, puis utiliser le fait que $u_n\to 0$, $a_n\to 0$ et $b_n\to 0$). □

Preuve de 3.13. Pour tout temps d'arrêt $S\leq T$ on a $0 \leq E(1-Z^\zeta_S) \leq E(1-Z^\zeta_T)$, donc 3.14 découle de 3.8. Par ailleurs, posons $Z'^\zeta_S = E(Z^\zeta_T|\mathcal{F}_S)$. Comme $(\sqrt{x}-\sqrt{y})^2 \leq |x-y|$, il vient

$$\begin{aligned}
E[(\sqrt{Z'^\zeta_S} - \sqrt{Z^\zeta_S})^2] &\leq E[|Z'^\zeta_S - Z^\zeta_S|] \leq E(1-Z'^\zeta_S) + E(1-Z^\zeta_S) \\
&= E(1-Z^\zeta_T) + E(1-Z^\zeta_S) \leq 2\,E(1-Z^\zeta_T),
\end{aligned}$$

Donc si $(\theta_n,u_n,\theta)\in\mathcal{Z}$, on déduit de 3.8 que

3.19 $$\sup_{S \text{ temps d'arrêt}, S\leq T} E[|((Z'^{\theta_n}_S)^{1/2} - (Z^{\theta_n}_S)^{1/2})/u_n|^2] \longrightarrow 0.$$

Enfin 3.9 et le lemme 3.17 appliqué à $X_n = Z^{\theta_n}_T$ et $Y = \theta.V_T/2$ entrainent que

3.20 $$\sup_{S \text{ temps d'arrêt}, S\leq T} E[|((Z'^{\theta_n}_S)^{1/2} - 1)/u_n - \frac{1}{2}\,\theta.V_S|^2] \longrightarrow 0.$$

Comme 3.19 et 3.20 entrainent 3.15, on a le résultat. □

4 - DERIVABILITE LOCALE

Il peut sembler naturel d'appeler dérivabilité locale la propriété du corollaire 3.16. Toutefois cette définition ne permettrait pas d'obtenir un critère en termes de processus de Hellinger. Nous allons donc donner une définition un peu moins restrictive. Commençons par un point de terminologie.

On appellera suite localisante toute suite (T_p) de temps d'arrêt croissant P-p.s. vers $+\infty$. Une famille localisante est constituée d'une suite localisante (T_p), et pour chaque $p\in\mathbb{N}$ d'une suite $(T(n,p))_{n\geq 1}$ de temps d'arrêt tels que $T(n,p)\leq T_p$ et

4.1 $$P(T(n,p)<T_p) \longrightarrow 0 \quad \text{quand} \quad n\uparrow\infty.$$

4.2 DEFINITION. Le modèle est dit localement dérivable en $\theta=0$ s'il existe un processus càdlàg V à valeurs dans $\mathbb{R}^d$ tel que pour toute suite $(\theta_n,u_n,\theta)\in\mathcal{T}$ (cf. 3.7) il existe une famille localisante $(T_p,T(n,p))$ avec

4.3 $$E(1 - Z^{\theta_n}_{T(n,p)})/u_n^2 \longrightarrow 0, \quad \forall p\in\mathbb{N},$$

4.4 $$((Z^{\theta_n}_{t\wedge T(n,p)})^{1/2} - 1)/u_n \xrightarrow{L^2(P)} \frac{1}{2}\,\theta.V_{t\wedge T_p}, \quad \forall p\in\mathbb{N}, \forall t\geq 0. \ \square$$

Si le modèle est dérivable en chaque instant T_p pour une suite localisante (T_p), il est clairement localement dérivable (prendre $T(n,p)=T_p$ et utiliser 3.15, et 3.16 pour l'existence du processus V).

4.5 REMARQUE. On verra plus loin que la suite (T_p) peut être choisie indépendamment de (θ_n), et qu'on peut associer à chaque θ des temps d'arrêt $S(\theta,p)\leq T_p$, tels que dans 4.2 on puisse choisir $T(n,p) = S(\theta_n,p)$ pour tout $(\theta_n,u_n,\theta)\in\mathcal{T}$. Ceci s'applique aussi aux familles localisantes construites dans le théorème ci-dessous. $\square$

4.6 THEOREME. Supposons le modèle localement dérivable.

a) Le processus V est une martingale locale, localement de carré intégrable, avec $V_0=0$ p.s. (Comparer à 3.16).

b) Pour toute suite $(\theta_n,u_n,\theta)\in\mathcal{T}$ il existe une famille localisante $(T_p,T(n,p))$ telle qu'on ait 4.3 et les deux propriétés suivantes:

4.7 $$E[\sup_t|((Z^{\theta_n}_{t\wedge T(n,p)})^{1/2}-1)/u_n - \tfrac{1}{2}\,\theta.V_{t\wedge T_p}|^2] \longrightarrow 0, \quad \forall p\in\mathbb{N},$$

4.8 $$E[\sup_t|(Z^{\theta_n}_{t\wedge T(n,p)}-1)/u_n - \theta.V_{t\wedge T_p}|] \longrightarrow 0, \quad \forall p\in\mathbb{N}.$$

Preuve. (i) Soit $(\theta_n,u_n,\theta)\in\mathcal{T}$, à laquelle on associe une famille localisante $(T_p,S(n,p))$ vérifiant 4.3 et 4.4. D'après ces propriétés, on a 3.8 et 3.9, à condition de remplacer $Z^{\theta_n}_T$ par $Z^{\theta_n}_{S(n,p)}$ et $\theta.V_T$ par $\theta.V_{T_p}$. On peut alors reproduire la preuve de 3.13: si $V'_S = E(\theta.V_{T_p}|\mathcal{F}_S)$ pour tout temps d'arrêt $S\leq T_p$, on obtient:

4.9 $$\sup_{S \text{ temps d'arrêt},\ S\leq S(n,p)} E(1-Z^{\theta_n}_S)/u_n^2 \longrightarrow 0,$$

4.10 $$\sup_{S \text{ temps d'arrêt},S\leq S(n,p)} E[|((Z^{\theta_n}_S)^{1/2}-1)/u_n - \tfrac{1}{2}\,V'_S|^2] \longrightarrow 0.$$

(ii) Comparons 4.4 et 4.10 pour $S = t\wedge S(n,p)$: comme $P(S(n,p)<T_p) \to 0$, on en déduit que $\theta.V_{t\wedge T_p} = V'_{t\wedge T_p}$ p.s. Comme V est càdlàg et comme on peut évidemment choisir une version de V' càdlàg sur $[0,T_p]$, on peut donc remplacer V'_S par $\theta.V_S$ dans 4.10.

En particulier, $\theta.V$ est une martingale sur $[0,T_p]$, avec $\theta.V_{T_p}$ de carré intégrable par 4.4. Comme (T_p) est une suite localisante, $\theta.V$ est donc une martingale locale, localement de carré intégrable. Enfin θ est arbitraire sur la sphère unité de $\mathbb{R}^d$, et $V_0=0$ est trivialement vérifié, donc on a (a).

(iii) Posons $U^n_t = ((Z^{\theta_n}_{t\wedge S(n,p)})^{1/2}-1)/u_n - \frac{1}{2}\,\theta.V_{t\wedge T_p}$, et $T(n,p) = \inf(t\colon t\geq S(n,p)$ ou $|U^n_t|\geq\rho_n)$, où ρ_n est une suite de nombres positifs tendant vers 0, suffisamment lentement pour que si v_n désigne le premier membre de 4.10, on ait $v_n/\rho_n^2 \to 0$. Nous allons montrer que $(T_p,T(n,p))$ est une famille localisante vérifiant 4.3, 4.7 et 4.8. Comme $T(n,p)\leq S(n,p)$, 4.3 est d'ailleurs évident.

Comme $T(n,p)\leq S(n,p)$, 4.10 entraine $E[(U^n_{T(n,p)})^2] \leq v_n$. Mais si $T(n,p)<S(n,p)$ on a $|U^n_{T(n,p)}| \geq \rho_n$, d'où

$$P(T(n,p)<T_p) \;\leq\; P(S(n,p)<T_p) + v_n/\rho_n^2,$$

qui tend vers 0 quand $n\uparrow\infty$ d'après l'hypothèse faite sur ρ_n. Donc $(T_p,T(n,p))$ est une famille localisante.

Par ailleurs $(U^n)^{*2}_{T(n,p)} \leq \rho_n^2 + (U^n_{T(n,p)})^2$ par construction de $T(n,p)$, si bien que $E[(U^n)^{*2}_{T(n,p)}] \leq \rho_n^2 + v_n$, qui tend vers 0 par 4.10. Enfin si $W^n = ((Z^{\theta_n}_{t\wedge T(n,p)})^{1/2} - 1)/u_n - \frac{1}{2}\,\theta.V_{t\wedge T_p}$, on a $|W^n_t| \leq |U^n_{t\wedge T(n,p)}| + \frac{1}{2}|\theta.V_{t\wedge T(n,p)} - \theta.V_{t\wedge T_p}|$. Donc

$$E[|W^n|^{*2}_\infty] \;\leq\; 2E[|U^n|^{*2}_{T(n,p)}] + E[|\theta.V|^{*2}_{T_p}\,1_{\{T(n,p)<T_p\}}].$$

Comme $E[|U^n|^{*2}_{T(n,p)}] \to 0$, et $E(|\theta.V|^{*2}_{T_p}) \leq 4E[(\theta.V_{T_p})^2] < \infty$ (inégalité de Doob), et $P(T(n,p)<T_p) \to 0$, on en déduit que $E[|W^n|^{*2}_\infty] \to 0$, ce qui n'est autre que 4.7.

Enfin

$$(Z^{\theta_n}_{t\wedge T(n,p)} - 1)/u_n - \theta.V_{t\wedge T_p}$$
$$= 2[((Z^{\theta_n}_{t\wedge T(n,p)})^{1/2}-1)/u_n - \tfrac{1}{2}\,\theta.V_{t\wedge T_p}] + [(Z^{\theta_n}_{t\wedge T(n,p)})^{1/2}-1]^2/u_n,$$

ce qui permet de déduire facilement 4.8 de 4.7. □

4.11 REMARQUE. Même si on peut prendre $T(n,p)=T_p$ dans 4.2 (par exemple sous les hypothèses de 3.16), il n'en est pas de même dans 4.6. □

Le lemme suivant joue un rôle clé:

4.12 LEMME. <u>Supposons le modèle localement dérivable, et soit</u> $(\theta_n,u_n,\theta)\in\mathcal{T}$ <u>et</u> $U^\zeta = (Z^\zeta - 1)/|\zeta|$. <u>Alors</u> $U^{\theta_n} \xrightarrow{\mathcal{S}(P)} \theta.V$.

<u>Preuve</u>. Soit $(T_p,T(n,p))$ la famille localisante associée à (θ_n,u_n,θ) par 4.6, pour laquelle on peut toujours supposer que $T_p\leq p$. Soit $Z^\zeta = 1+M^\zeta-A^\zeta$ la décomposition canonique de la surmartingale Z^ζ: A^ζ

est croissant et $E(A_T^\zeta) = E(1-Z_T^\zeta)$ pour tout temps d'arrêt borné T. La décomposition canonique de $U^{\theta_n}-\theta.V$ est N^n+B^n, avec $B^n = A^{\theta_n}/u_n$ et $N^n = M^{\theta_n}/u_n - \theta.V$. D'après 4.3 et 4.8 il vient alors
$E(|N^n|^*_{T(n,p)} + Var(B^n)_{T(n,p)}) \leq E(|U^{\theta_n}-\theta.V|^*_{T(n,p)} + 2A^{\theta_n}_{T(n,p)}/u_n) \to 0$,

ce qui implique classiquement $(U^{\theta_n}-\theta.V)^{T(n,p)} \xrightarrow{\mathscr{S}(P)} 0$.

Par ailleurs, $(U^{\theta_n})_t^{T_p} = (U^{\theta_n})_t^{T(n,p)}$ pour tout t, sur l'ensemble $\{T(n,p)=T_p\}$, dont la probabilité tend vers 1 par 4.1; un autre résultat classique sur la topologie d'Emery permet alors de déduire que $(U^{\theta_n})^{T_p} \xrightarrow{\mathscr{S}(P)} \theta.V^{T_p}$. Ceci étant vrai pour tout p, $U^{\theta_n} \xrightarrow{\mathscr{S}(P)} \theta.V$.□

5 - LES PROCESSUS DE HELLINGER

D'abord, comme dans la preuve de 4.12, on considère la décomposition de Doob-Meyer de la P-surmartingale Z^θ:

5.1 $\qquad Z^\theta = 1 + M^\theta - A^\theta$, M^θ martingale, A^θ croissant prévisible.

Rappelons ensuite, d'après [5], la notion de processus de Hellinger d'ordre $\beta\in]0,1[$ entre P_θ et P_ζ. On appelle ainsi tout processus croissant prévisible $h(\beta)^{\theta\zeta}$ nul en 0, ayant la propriété suivante: pour toute probabilité Q dominant P_θ et P_ζ, si z^θ et z^ζ désignent les processus densité, alors

5.2 $\qquad (z^\theta)^\beta(z^\zeta)^{1-\beta} + [(z_-^\theta)^\beta(z_-^\zeta)^{1-\beta}]\bullet h(\beta)^{\theta\zeta}$ est une Q-martingale.

Alors, $h(\beta)^{\theta\zeta}$ est déterminé de manière Q-p.s. unique sur l'intervalle stochastique $\{z_-^\theta>0,\ z_-^\zeta>0\}$. Dans la suite $h(\beta)^{\theta\zeta}$ désigne une version <u>quelconque</u> de ce processus, et pour simplifier $h^{\theta\zeta} = h(\frac{1}{2})^{\theta\zeta}$. Noter que $h(\beta)^{\zeta\theta}$ est une version de $h(1-\beta)^{\theta\zeta}$, donc on peut supposer (et on supposera) que $h^{\theta\zeta} = h^{\zeta\theta}$.

Pour des raisons techniques, on va introduire une série d'autres processus, et étudier certaines relations entre ceux-ci. D'abord, le processus $(Z^\theta)^\beta(Z^\zeta)^{1-\beta}$ est clairement une P-surmartingale positive, donc il existe un processus croissant prévible $\overline{h}(\beta)^{\theta\zeta}$ nul en 0, avec

5.3 $\quad (Z^\theta)^\beta(Z^\zeta)^{1-\beta} + [(Z^\theta_-)^\beta(Z^\zeta_-)^{1-\beta}]\bullet\overline{h}(\beta)^{\theta\zeta}$ est une P-martingale.

Là encore $\overline{h}(\beta)^{\theta\zeta}$ est P-p.s. unique sur $\{Z^\theta_->0, Z^\zeta_->0\}$, et sur cet ensemble on a $\overline{h}(\beta)^{\zeta\theta} = \overline{h}(1-\beta)^{\theta\zeta}$. On écrit $\overline{h}^{\theta\zeta} = \overline{h}(\frac{1}{2})^{\theta\zeta}$.

Ensuite, on pose:

5.4 $$k(\beta)^{\theta\zeta} = [(Z^\theta_-)^\beta(Z^\zeta_-)^{1-\beta}]\bullet\overline{h}(\beta)^{\theta\zeta} - \beta A^\theta - (1-\beta)A^\zeta,$$

et $k^{\theta\zeta} = k(\frac{1}{2})^{\theta\zeta}$ (ces processus sont P-p.s. définis de manière unique). Puis, soit

5.5 $$\begin{cases} H(\beta)^{\theta\zeta} = h(\beta)^{\theta 0} + h(\beta)^{0\zeta} - h(\beta)^{\theta\zeta}, \\ K(\beta)^{\theta\zeta} = k(\beta)^{\theta 0} + k(\beta)^{0\zeta} - k(\beta)^{\theta\zeta}, \end{cases}$$

5.6 $$\begin{cases} Y(\beta)^{\theta} = \beta Z^\theta + 1 - \beta - (Z^\theta)^\beta, \\ W(\beta)^{\theta\zeta} = [1-(Z^\theta)^\beta][1-(Z^\zeta)^{1-\beta}]. \end{cases}$$

D'après 5.1 et 5.3, on a

5.7 $\quad Y(\beta)^\theta - k(\beta)^{\theta 0}$ et $W(\beta)^{\theta\zeta} - K(\beta)^{\theta\zeta}$ sont des P-martingales.

5.8 LEMME. a) <u>On a</u> $\overline{h}(\beta)^{\theta 0} = h(\beta)^{\theta 0}$ P-<u>p.s. sur</u> $\{Z^\theta_->0\}$.

b) <u>Le processus</u> $[(Z^\theta_-)^\beta(Z^\zeta_-)^{1-\beta}]\bullet[\overline{h}(\beta)^{\theta\zeta} - h(\beta)^{\theta\zeta}]$ <u>est</u> P-<u>p.s.</u> <u>croissant et majoré au sens fort par le processus</u> $wA^\theta + w'A^\zeta$ <u>dès que</u>

5.9 $$w^\beta\, w'^{1-\beta} \geq \beta^\beta\,(1-\beta)^{1-\beta}.$$

<u>Preuve</u>. Soit Q une probabilité dominant $P=P_0$, P_θ et P_ζ, et soit $z=z^0$, z^θ, z^ζ les processus densité respectifs par rapport à Q. On écrit $X\sim Y$ si $X-Y$ est une Q-martingale locale. Soit $R=\inf(t:z_t=0)$.

Comme $Z^\theta+A^\theta$ est une P-martingale, on a $z(Z^\theta+A^\theta)\sim 0$. On a aussi $zA^\theta \sim z_-\bullet A^\theta$ et $zZ^\theta = z^\theta 1_{\{z>0\}}$. Donc $z_-\bullet A^\theta \sim -z^\theta 1_{\{z>0\}} \sim z^\theta 1_{\{z=0\}} = z^\theta-(z^\theta)^R+z^\theta_R 1_{[\![R,\infty[\![} \sim B^\theta := z^\theta_R 1_{[\![R,\infty[\![}$, et par suite

5.10 $$z_-\bullet A^\theta \text{ est le Q-compensateur de } B^\theta.$$

Pour simplifier, posons $h = h(\beta)^{\theta\zeta}$, $\overline{h} = \overline{h}(\beta)^{\theta\zeta}$, $Y = (z^\theta)^\beta(z^\zeta)^{1-\beta}$ et $\overline{Y} = (Z^\theta)^\beta(Z^\zeta)^{1-\beta}$. D'après 5.3, on a $z(\overline{Y} + \overline{Y}_-\bullet\overline{h}) \sim 0$, et donc $-z_-\overline{Y}_-\bullet\overline{h} \sim z\overline{Y} = Y1_{\{z>0\}} = Y^R-Y_R 1_{[\![R,\infty[\![}$. Mais $Y \sim -Y_-\bullet h$ par 5.2, de sorte que si $C = Y_R 1_{[\![R,\infty[\![}$ et si $\tilde{C}$ est son Q-compensateur, on a

5.11 $$z\_\bar{Y}\_\bullet\bar{h} = Y\_\bullet h^R + \tilde{C} \quad \text{Q-p.s.}$$

Si $\zeta=0$ on a $Y = (z^\theta)^\beta(z)^{1-\beta}$, donc $z\bar{Y}=Y$ et $Y_R=0$, d'où $C=\tilde{C}=0$ et 5.11 entraine $Y\_\bullet h^R = Y\_\bullet h$ Q-p.s. Comme $R=\infty$ P-p.s., on en déduit que $\bar{h}=h$ P-p.s. sur $\{Y\_>0\}$, d'où (a).

Passons au cas général. Un calcul simple montre que $x^\beta y^{1-\beta} \leq wx + w'y$ pour tous $x,y\geq 0$ si et seulement si on a 5.9. Donc sous 5.9 on a $Y_R \leq wz_R^\theta + w'z_R^\zeta$, donc le processus croissant C est fortement dominé par $wB^\theta+w'B^\zeta$, et le processus croissant $\tilde{C}$ est fortement dominé par $z\_\bullet(wA^\theta+w'A^\zeta)$ par 5.10. Par ailleurs $R=\infty$ et $Y=z\bar{Y}$ P-p.s., donc 5.11 entraine que P-p.s. $\bar{Y}\_\bullet(\bar{h}-h) = (1/z\_)\bullet\tilde{C}$, et on en déduit (b). □

6 - PROCESSUS INFORMATION ET PROCESSUS DE HELLINGER

Commençons par définir le processus information:

6.1 DEFINITION. Si le modèle est localement dérivable en $\theta=0$, on appelle <u>processus information</u> (de Fisher, en $\theta=0$), le processus $\Lambda = (\Lambda^{ij})_{i,j\leq d}$ défini par $\Lambda^{ij}=\langle V^i,V^j\rangle$ (covariation quadratique prévisible de la martingale localement de carré intégrable V, cf. 4.6). □

Si le modèle vérifie les conditions de 3.16, pour tout temps d'arrêt T inférieur ou égal à l'un des T_p le modèle non-filtré restreint à la tribu $\mathcal{F}_T$ est dérivable en $\theta=0$, et $I_T^{ij} = E(\Lambda_T^{ij})$ en est la matrice d'information de Fisher (cf. 3.4).

Voici alors le résultat essentiel de cet article:

6.2 THEOREME. a) <u>Pour que le modèle soit localement dérivable en</u> $\theta=0$ <u>il faut et il suffit qu'on ait [A] et [B] ci-dessous</u>:

[A] <u>Il existe un processus</u> $\Lambda=(\Lambda^{ij})_{i,j\leq d}$ <u>nul en</u> 0, <u>croissant dans</u>

<u>l'espace des matrices symétriques non-négatives, tel que pour tous</u> (θ_n, u_n, θ) <u>et</u> (ζ_n, v_n, ζ) <u>dans</u> $\mathcal{T}$ <u>on ait</u>

6.3 $$(h^{0,\theta_n} + h^{0,\zeta_n} - h^{\theta_n,\zeta_n})/u_n v_n \xrightarrow{P-V} \frac{1}{4}\,\theta.\Lambda.\zeta\ .$$

[B] <u>Pour tout</u> $(\theta_n, u_n, \theta)\in\mathcal{T}$ <u>on a</u>

6.4 $$A_t^{\theta_n}/u_n^2 \xrightarrow{P} 0, \qquad \forall t\geq 0.$$

(Noter que si $P_\theta \overset{loc}{\ll} P$ pour tout θ, [B] est automatique).

b) <u>Dans ce cas,</u> Λ <u>est le processus information, et pour tous</u> $\beta\in]0,1[$, $(\theta_n, u_n, \theta)\in\mathcal{T}$, $(\zeta_n, v_n, \zeta)\in\mathcal{T}$, <u>avec</u> $u_n/v_n \to a\in]0,\infty[$ <u>lorsque</u> $\beta\neq 1/2$, <u>on a</u>:

6.5 $$h(\beta)^{\theta_n,0}/u_n^2 \xrightarrow{P-V} \frac{\beta(1-\beta)}{2}\,\theta.\Lambda.\theta,$$

6.6 $$[h(\beta)^{\theta_n,0} + h(\beta)^{0,\zeta_n} - h(\beta)^{\theta_n,\zeta_n}]/u_n v_n \xrightarrow{P-V} \beta(1-\beta)\theta.\Lambda.\zeta.$$

6.7 REMARQUES. 1) Il est immédiat que [A] équivaut à la condition apparemment plus forte suivante: il existe un processus Λ comme dans [A] tel que

$$\frac{1}{|\theta||\zeta|}\,\mathrm{Var}(h^{0\theta} + h^{0\zeta} - h^{\theta\zeta} - \frac{1}{4}\,\theta.\Lambda.\zeta)_t \xrightarrow{P} 0 \quad \text{si} \quad \theta,\zeta\to 0,\ \forall t\geq 0.$$

On pourrait écrire 6.4, 6.5 et 6.6 de manière analogue.

2) Sous [A] on a $h_t^{0,\theta_n} \xrightarrow{P} 0$ si $\theta_n \to 0$, donc par [5, V.4.31] il y a convergence en variation des restrictions des P_{θ_n} à $\mathcal{F}_t$ vers la restriction de P à $\mathcal{F}_t$. Donc les convergences 6.3, 6.4, 6.5, 6.6 sont vraies non seulement en P-probabilité, mais aussi en P_{η_n}-probabilité lorsque $\eta_n \to 0$: cela signifie, pour 6.3 par exemple, que pour tous $\varepsilon>0$, $t\geq 0$:

$$P_{\eta_n}[\mathrm{Var}((h^{0,\theta_n} + h^{0,\zeta_n} - h^{\theta_n,\zeta_n})/u_n v_n - \frac{1}{4}\,\theta.\Lambda.\zeta)_t > \varepsilon] \longrightarrow 0.\ \square$$

Nous commençons par plusieurs lemmes.

6.8 LEMME. <u>Supposons le modèle localement dérivable, et soit</u> $(\theta_n, u_n, \theta)\in\mathcal{T}$, $(\zeta_n, v_n, \zeta)\in\mathcal{T}$, <u>et supposons de plus que</u> $u_n/v_n \to a\in]0,\infty[$

<u>si</u> $\beta\neq 1/2$. <u>Alors, si</u> Λ <u>est le processus information on a</u>:

6.9 $$k(\beta)^{\theta_n,0}/u_n^2 \xrightarrow{P-V} \frac{\beta(1-\beta)}{2}\theta.\Lambda.\theta \qquad \text{(cf. 5.4)},$$

6.10 $$K(\beta)^{\theta_n,\zeta_n}/u_n v_n \xrightarrow{P-V} \beta(1-\beta)\theta.\Lambda.\zeta \qquad \text{(cf. 5.5)}.$$

<u>Preuve</u>. a) D'après 4.6 on peut trouver une famille localisante $(T_p,T(n,p))$ telle qu'on ait 4.7 pour (θ_n,u_n,θ) et pour (ζ_n,v_n,ζ), et aussi $E(|V|^{*2}_{T_p})<\infty$. On a alors $\sup_t|(Z^{\theta_n}_{t\wedge T(n,p)})^{1/2} - 1|/u_n \longrightarrow \sup_t|\theta.V_{t\wedge T_p}|$ dans L^2, et de même pour (ζ_n). Donc

6.11 les suites $\{\sup_t[(Z^{\theta_n}_{t\wedge T(n,p)})^{1/2}-1]^2/u_n^2\}_{n\geq 1}$ et $\{\sup_t[(Z^{\zeta_n}_{t\wedge T(n,p)})^{1/2}-1]^2/v_n^2\}_{n\geq 1}$ sont uniformément intégrables.

b) Les fonctions $f^n(x) = [\beta(1+u_n x)+1-\beta-(1+u_n x)^\beta]/u_n^2$ convergent localement uniformément ainsi que leurs dérivées d'ordre 1 et 2 vers la fonction $f(x) = \frac{\beta(1-\beta)}{2}x^2$ et ses dérivées, quand $n\uparrow\infty$. Donc par 4.12 et 2.6 on a $f^n(U^{\theta_n}) \xrightarrow{\mathcal{L}(P)} f(\theta.V)$, d'où a-fortiori (cf 4.1):

6.12 $$[f^n(U^{\theta_n})]^{T(n,p)} \xrightarrow{\mathcal{L}(P)} f(\theta.V)^{T_p}.$$

Par ailleurs $|f^n(x)| \leq K[1 - (1+u_n x)^{1/2}]^2/u_n^2$ pour une certaine constante K, donc $|f^n(U^{\theta_n})| \leq K[(Z^{\theta_n})^{1/2}-1]^2/u_n^2$. Par suite 6.11, 6.12 et 2.1 entrainent que les compensateurs du premier membre de 6.12 convergent en variation, en P-probabilité, vers celui du second membre. Comme $f^n(U^{\theta_n}) = Y(\beta)^{\theta_n}/u_n^2$ (cf. 5.6) et $f(\theta.V) = \frac{\beta(1-\beta)}{2}(\theta.V)^2$ il vient d'après 5.7 et la définition de Λ:

$$[k(\beta)^{\theta_n,0}]^{T(n,p)}/u_n^2 \xrightarrow{P-V} \frac{\beta(1-\beta)}{2}(\theta.\Lambda.\theta)^{T_p}.$$

Comme $(T_p,T(n,p))$ est une famille localisante, on en déduit 6.9.

c) Passons maintenant à 6.10. Les fonctions

$$f^n(x',x'') = [1-(1+u_n x')^\beta][1-(1+v_n x'')^{1-\beta}]/u_n v_n,$$

convergent localement uniformément ainsi que leurs dérivées partielles d'ordre 1 et 2 vers la fonction $f(x',x'') = \beta(1-\beta)x'x''$ et ses dérivées. Donc $f^n(U^{\theta_n},U^{\zeta_n}) \xrightarrow{\mathcal{L}(P)} f(\theta.V,\zeta.V)$. On a $f^n(U^{\theta_n},U^{\zeta_n}) =$

$W(\beta)^{\theta_n,\zeta_n}$ et $f(\theta.V,\zeta.V) = \beta(1-\beta)(\theta.V)(\zeta.V)$, dont les compensateurs sont respectivement $K(\beta)^{\theta_n,\zeta_n}$ (cf. 5.7) et $\beta(1-\beta)\theta.\Lambda.\zeta$. Exactement comme en (b), pour obtenir 6.10 il suffit alors de montrer que chaque suite $\{\sup_t |f^n(U^{\theta_n},U^{\zeta_n})_{t\wedge T(n,p)}|\}_{n\geq 1}$ est uniformément intégrable.

Pour cela, on remarque qu'il existe une constante C_β telle que $|1-(x)^\beta| \leq C_\beta|1-\sqrt{x}|^{2\beta}$, de sorte que $|f^n(U^{\theta_n},U^{\zeta_n})| \leq$ $C_\beta C_{1-\beta}\ |((Z^{\theta_n})^{1/2}-1)/u_n|^{2\beta}\ |((Z^{\zeta_n})^{1/2}-1)/v_n]^{2(1-\beta)}\ (u_n/v_n)^{2\beta-1}$, et comme $u_n/v_n \to a$ si $\beta\neq 1/2$, on voit que

$$\begin{aligned} f^n(U^{\theta_n},U^{\zeta_n}) &\leq C\ |((Z^{\theta_n})^{1/2}-1)/u_n|^{2\beta}\ |((Z^{\zeta_n})^{1/2}-1)/v_n|^{2(1-\beta)} \\ &\leq C\ [\beta|((Z^{\theta_n})^{1/2}-1)/u_n|^2 + (1-\beta)|((Z^{\zeta_n})^{1/2}-1)/v_n|^2] \end{aligned}$$

pour une certaine constante C (car $x^\beta y^{1-\beta} \leq \beta x + (1-\beta)y$). On déduit alors de 6.11 l'uniforme intégrabilité cherchée. □

6.13 LEMME. <u>Soit</u> $(\theta_n,u_n,\theta)\in\mathcal{T}$, $(\zeta_n,v_n,\zeta)\in\mathcal{T}$, $\beta\in]0,1[$.

a) <u>Sous 6.4 et</u>

6.14 $$Z^{\theta_n} \xrightarrow{P-U} 1,$$

<u>il y a équivalence entre 6.5 et 6.9</u>.

b) <u>Soit</u> $S(\eta,q) = \inf(t: |((Z_t^\eta)^{1/2} - 1)/|\eta| \geq q)$. <u>Si on a 6.4 pour</u> (θ_n,u_n,θ) <u>et</u> (ζ_n,v_n,ζ), <u>si</u> $u_n/v_n \to a\in]0,\infty[$ <u>lorsque</u> $\beta\neq 1/2$, <u>et si</u>

6.15 $$\lim_{q\uparrow\infty} \limsup_{n\uparrow\infty} P(S(\theta_n,q)\wedge S(\zeta_n,q) \leq t) = 0, \quad \forall t\geq 0,$$

<u>6.5 et 6.6 sont équivalents à 6.9 et 6.10</u>.

<u>Preuve</u>. a) D'après 5.4 et 5.8(a) on a $k(\beta)^{\rho 0} = (Z_-^\rho)^\beta \bullet h(\beta)^{\rho 0} - \beta A^\rho$ P-p.s. L'équivalence cherchée est donc immédiate.

b) D'après 5.8(a), un calcul simple montre que P-p.s. (cf. 5.5):

$$K(\beta)^{\rho\eta} = (Z_-^\rho)^\beta (Z_-^\eta)^{1-\beta} \bullet H(\beta)^{\rho\eta} + B(1)^{\rho\eta} + B(2)^{\rho\eta} + B(3)^{\rho\eta},$$

où

$$B(1)^{\rho\eta} = (Z_-^\rho)^\beta(1-(Z_-^\eta)^{1-\beta})\bullet h(\beta)^{\rho 0}, \quad B(2)^{\rho\eta} = (Z_-^\eta)^{1-\beta}(1-(Z_-^\rho)^\beta)\bullet h(\beta)^{0\eta},$$

$$B(3)^{\rho\eta} = -\ (Z_-^\rho)^\beta (Z_-^\eta)^{1-\beta}\bullet[\bar{h}(\beta)^{\rho\eta} - h(\beta)^{\rho\eta}].$$

Comme 6.15 implique 6.14 pour (θ_n) et pour (ζ_n), le résultat découlera de ce que, sous les conditions équivalentes 6.5 ou 6.9, on a pour $i=1,2,3$:

6.16 $$B(i)^{\theta_n,\zeta_n}/u_n v_n \xrightarrow{P-V} 0.$$

Commençons par $i=1,2$. Il existe une constante c_q telle que si $|\sqrt{x}-1| \le qu$ et $0<u\le 1$, alors $|x^{1-\beta}-1| \le c_q u$ et $x^\beta \le c_q$. On a $\mathrm{Var}(B(1)^{\theta_n,\zeta_n})_{S(\theta_n,q)\wedge S(\zeta_n,q)} \le c_q^2 \, v_n \, h(\beta)^{\theta_n,0}$, donc sous 6.5 et 6.15 on a 6.16 pour $i=1$, et de même pour $i=2$.

Il reste à examiner le cas de $B(3)^{\theta_n,\zeta_n}$, qui d'après 5.8 est majoré en variation par $w_n A^{\theta_n} + w'_n A^{\zeta_n}$ dès que w_n, w'_n vérifient 5.9. Lorsque $\beta=1/2$ on prend $w_n=v_n/u_n$ et $w'_n=u_n/v_n$; lorsque $\beta\neq 1/2$, on prend $w_n=w'_n=1$: dans les deux cas on a 5.9. Etant donné 6.4, il est alors clair que 6.16 est satisfait pour $i=3$ (rappelons que $u_n/v_n \to a\in]0,\infty[$ si $\beta\neq 1/2$), et la démonstration est achevée. □

<u>Condition nécessaire et partie (b) de 6.2</u>. On suppose le modèle localement dérivable. D'abord, 6.4 découle immédiatement de 4.1 et 4.3 (car $E(A_T^\rho) = E(1-Z_T^\rho)$ pour tout temps d'arrêt borné T).

Ensuite si Λ est le processus information, d'après le lemme 6.8 on a 6.9 et 6.10. On a aussi 6.15, qui découle aisément de 4.7, donc le lemme 6.13 implique 6.5 et 6.6, donc a-fortiori 6.3. □

<u>Condition suffisante de 6.2</u>. On suppose [A] et [B]. La preuve va comporter plusieurs étapes.

1) On fixe $(\theta_n,u_n,\theta)\in\mathcal{T}$. On a 6.4 et 6.5, et aussi (par 6.6):

6.17 $$H(\tfrac{1}{2})^{\theta_n,\theta_m}/u_n u_m \xrightarrow{P-V} \tfrac{1}{4}\,\theta.\Lambda.\theta \qquad \text{si } n,m\uparrow\infty.$$

Montrons d'abord 6.14. D'après 6.5, $h_t^{\theta_n,0} \xrightarrow{P} 0$, donc [5,V.4.31] implique que la distance en variation de P à P_{θ_n} sur $(\Omega,\mathcal{F}_t)$ tend vers 0. Soit $Q_n=(P+P_{\theta_n})/2$, et z^n, z'^n les processus densité de P et P_{θ_n} par rapport à Q. On a donc $E_{Q_n}[|z_t^n-z_t'^n|] \to 0$, et comme le processus $Z^n - Z'^n$ est une martingale bornée par 2 on a aussi $E_{Q_n}[|Z^n - Z'^n|_t^*] \to 0$. Comme $Z^n + Z'^n = 2$ et $Z^{\theta_n} = Z'^n / Z^n$, il en découle que $Q_n(|Z^{\theta_n} - 1|_t^* > \varepsilon) \to 0$ pour tout $\varepsilon > 0$. Enfin $P \leq 2Q_n$, d'où 6.14.

2) Posons $K^n = k^{\theta_n,0}/u_n^2$ et $K^{nm} = K(\frac{1}{2})^{\theta_n,\theta_m}/u_n u_m$. D'après 5.4 et 5.5 on a $k(\beta)^{\theta\theta} = 0$, d'où $K^{nn} = 2K^n$. Par 5.7 il existe des martingales M^n et M^{nm} telles que, si $W^n = ((Z^{\theta_n})^{1/2} - 1)/u_n$, on ait

6.18 $$W^n = M^n - u_n K^n - A^{\theta_n}/2u_n, \qquad W^n W^m = M^{nm} + K^{nm}.$$

Comme on a 6.4, 6.5 et 6.14, on déduit du lemme 6.13(a) que

6.19 $$K^n \xrightarrow{P-V} K/8,$$

où $K = \theta.\Lambda.\theta$. Posons

$$R_p = p \wedge \inf(t: K_t \geq p), \qquad R(n,p) = R_p \wedge \inf(t: \mathrm{Var}(K^n)_t \geq p \text{ ou } A_t^{\theta_n} \geq u_n^2).$$

(R_p) est une suite localisante, et $P(R(n,p) < R_p) \to 0$ quand $n \uparrow \infty$ par 6.19 et 6.4. De plus ces temps d'arrêt sont prévisibles, strictement positifs. Il existe donc des temps d'arrêt $T_p < R_p$ et $T(n,p) < R(n,p)$, avec $T(n,p) \leq T_p$, $P(T_p \leq R_p - 1) \leq 2^{-p}$ et $P(T(n,p) < T_p) \leq P(R(n,p) < R_p)$. Donc $(T_p, T(n,p))$ est une famille localisante, et

6.20 $$\mathrm{Var}(K^n)_{T(n,p)} \leq p, \qquad K_{T_p} \leq p, \qquad A_{T(n,p)}^{\theta_n} \leq u_n^2.$$

3) Pour tout temps d'arrêt borné S on a $E[(W_S^n)^2] = 2E(K_S^n)$ par 6.18 et $K^{nn} = 2K^n$. En appliquant ceci à $S = S(\theta_n, q) \wedge T(n,p)$, où $S(\theta_n, q) = \inf(t: |W_t^n| \geq q)$, comme en 6.13 , on déduit de 6.20 que

$$P(S(n,q) \leq T(n,p)) \leq 2p/q^2,$$

Comme $(T_p, T(n,p))$ est une famille localisante, il découle qu'on a 6.15. On peut alors appliquer 6.13(b) (avec une suite indicée par le couple n,m au lieu de n, mais la preuve est la même): comme on a 6.4, 6.5 et 6.17, il vient

6.21 $$K^{nm} \xrightarrow{P-V} K/4 \quad \text{si } n,m \uparrow \infty.$$

Soit alors $R'_{nmp} = \inf(t: \mathrm{Var}(K^{nm})_t \geq p)$, qui d'après 6.21 et la définition de R_p vérifie $P(R'_{nmp} < R_p) \to 0$ quand $n,n \uparrow \infty$. Exactement comme ci-dessus, R'_{nmp} est prévisible strictement positif, tandis que $T_p<R_p$, donc il existe des temps d'arrêt $R_{nmp}<R'_{nmp}$ vérifiant

6.22 $$P(R_{nmp} < T_p) \longrightarrow 0 \quad \text{si} \quad n,m\uparrow\infty.$$

Comme de plus $\mathrm{Var}(K^{nm})_{R_{nmp}} \leq p$ on déduit de 6.21 que pour tout t:

6.23 $$K^{nm}_{t\wedge R_{nmp}} \xrightarrow{L^1(P)} \frac{1}{4} K_{t\wedge T_p} \quad \text{si} \quad n,m\uparrow\infty.$$

4) Fixons $p\in\mathbb{N}^*$. Posons $Y^n_t = W^n_{t\wedge T(n,p)}$, et $U^{nm}_t = Y^n_tY^m_t - Y^n_{t\wedge R_{nmp}}Y^m_{t\wedge R_{nmp}}$. On a

6.24 $$E(|U^{nm}_t|) \leq E[|Y^m_t||Y^n_t - Y^n_{t\wedge R_{nmp}}|] + E[|Y^m_t - Y^m_{t\wedge R_{nmp}}||Y^n_{t\wedge R_{nmp}}|].$$

6.18, 6.20 et $K^{nn}=2K^n$ entrainent que pour tout temps d'arrêt S

6.25 $$E[|Y^n_S|^2] = 2E(K^n_{S\wedge T(n,p)}) \leq 2p.$$

Par ailleurs si T est un autre temps d'arrêt tel que $S\leq T$, il vient $(Y^n_T-Y^n_S)^2 = (Y^n_T)^2 - (Y^n_S)^2 - 2Y^n_S(Y^n_T-Y^n_S)$, donc 6.20 et 6.18 entrainent:

$$\begin{aligned} E[(Y^n_T-Y^n_S)^2] &= E[2(K^n_{T\wedge T(n,p)} - K^n_{S\wedge T(n,p)})(1 + u_nY^n_S) \\ &\qquad + (A^{\theta_n}_{T\wedge T(n,p)} - A^{\theta_n}_{S\wedge T(n,p)})Y^n_S/u_n] \\ &\leq 3pE[(1 + u_nY^n_S)\, 1_{\{S\wedge T(n,p)<T\wedge T(n,p)\}}] \\ &\leq 3p\, \{6p\, P(S\wedge T(n,p) < T\wedge T(n,p))\}^{1/2} \end{aligned}$$

dès que $u_n\leq 1$ (utiliser 6.25 et l'inégalité de Schwarz). En appliquant ceci à $S=t\wedge R_{nmp}$ et $T=t\wedge T(n,p)$ (donc $Y^n_t=Y^n_T$), et 6.25 à Y^m, on voit que le premier terme de droite de 6.24 est majoré par

$$\{E[|Y^m_t|^2]\, E[|Y^n_t - Y^n_{t\wedge R_{nmp}}|^2]\}^{1/2} \leq \sqrt{2p}\, \{3p[6p\, P(R_{nmp}<T_p)]^{1/2}\}^{1/2}$$

(car $\{S<T\}\subset\{R_{nmp}<T_p\}$). On a une majoration analogue pour le second terme de droite de 6.24, de sorte que 6.22 implique

6.26 $$E(|U^{mn}_t|) \longrightarrow 0 \quad \text{quand} \quad n,m\uparrow\infty.$$

Par ailleurs 6.18 entraine que $E(Y^n_{t\wedge R_{nmp}}Y^m_{t\wedge R_{nmp}}) = E(K^{nm}_{t\wedge R_{nmp}})$, donc 6.23 et 6.26 impliquent:

6.27 $$E[Y^n_tY^m_t] \longrightarrow \frac{1}{4} E[K_{t\wedge T_p}] \quad \text{quand} \quad n,m\uparrow\infty.$$

5) 6.27 entraine la convergence dans L^2 des variables Y_t^n vers une limite $Y(p)_t$. D'après 6.18, puis 6.4 et 6.20 on a donc

$$M^n_{t\wedge T(n,p)} \xrightarrow{L^2} Y(p)_t .$$

On en déduit d'abord que $Y(p)$ est une martingale, puis, par l'inégalité de Doob (avec $t=\infty$) et de nouveau 6.18 et 6.20, que

$$E[\sup_t |W^n_{t\wedge T(n,p)} - Y(p)_t|^2] \longrightarrow 0.$$

Etant donné 4.1, il est clair que $Y(p)_t = Y_{t\wedge T_p}$ où Y est une martingale locale, localement de carré intégrable. On a aussi

6.28 $$E[\sup_t |W^n_{t\wedge T(n,p)} - Y_{t\wedge T_p}|^2] \longrightarrow 0.$$

6) Remarquons que 4.3 découle immédiatement de 6.4 et 6.20, et aussi que les $(T(n,p))_{p\geq 1}$ ne dépendent pas de la suite (θ_m), mais seulement de θ_n (cf. remarque 4.5). Il reste donc à montrer qu'on a un processus càdlàg V vérifiant 4.4, et étant donné 6.28 cela se ramène à montrer que Y ci-dessus se met sous la forme $Y=\theta.V$.

Le processus V est construit ainsi: la $i^{\text{ième}}$ composante V^i est le processus Y obtenu comme ci-dessus, avec la suite $\theta_n=e_i/n$.

Soit (θ_n,u_n,t), $(\zeta_n,v_n,\zeta)\in\mathcal{T}$, avec les processus associés Y et Y' ci-dessus. En utilisant 6.28, on montre comme en 4.12 que $U^{\theta_n} \xrightarrow{\mathcal{L}(P)} Y$ et $U^{\zeta_n} \xrightarrow{\mathcal{L}(P)} Y'$. On peut alors reprendre la preuve de 6.8 (pour $\beta=\frac{1}{2}$), en remplaçant $\theta.V$ et $\zeta.V$ par Y et Y'. Dans 6.9 et 6.10 il convient alors de remplacer $\theta.\Lambda.\theta$ et $\theta.\Lambda.\zeta$ par $\langle Y,Y\rangle$ et $\langle Y,Y'\rangle$. En particulier, $K(\frac{1}{2})^{\theta_n,\zeta_n}/u_nv_n \xrightarrow{P-V} \frac{1}{4}\langle Y,Y'\rangle$.

Par ailleurs, exactement comme pour 6.21, on montre que $K(\frac{1}{2})^{\theta_n,\zeta_n}/u_nv_n \xrightarrow{P-V} \frac{1}{4}\,\theta.\Lambda.\zeta$, donc $\langle Y,Y'\rangle = \theta.\Lambda.\zeta$. En particulier si $\theta=e_i$ et $\zeta=e_j$ on obtient $\langle V^i,V^j\rangle = \Lambda^{ij}$, et pour θ quelconque on calcule aisément que $\langle Y-\theta.V,Y-\theta.V\rangle = 0$. Donc $Y=\theta.V$. $\square$

7 - EXEMPLES

Dans ce paragraphe on suppose que le modèle statistique est engendré par un processus de base X. Cela signifie qu'on a un processus X défini sur Ω et qui engendre la filtration $(\mathcal{F}_t)_{t\geq 0}$.

<u>a) Chaînes de Markov</u>. On considère ici un processus à temps discret $X=(X_n)_{n\in\mathbb{N}}$, donc $\mathcal{F}_t=\sigma(X_0,..,X_n)$ si $n\leq t<n+1$. On va bien-sûr retrouver les notions introduites en [1] ou [2].

Nous allons nous restreindre en fait au cas où, sous chaque mesure P_θ, le processus X est une chaîne de Markov homogène, dont on note $Q_\theta(x,dy)$ la probabilité de transition, et où $X_0=x_0$ est indépendant de θ et déterministe. L'espace d'état $(E,\mathcal{E})$ de X est quelconque.

<u>Hypothèse [H]</u>. Pour chaque $x\in E$ le modèle statistique (non filtré) $\mathcal{E}_x = (E,\mathcal{E},(Q_\theta(x,.))_{\theta\in\Theta})$ admet une information de Fisher $I(x)$ en $\theta=0$, au sens de 1.1. □

7.1 THEOREME. <u>Sous l'hypothèse [H], le modèle statistique filtré est localement fortement dérivable en</u> $\theta=0$, <u>et son processus information est donné par</u>

7.2 $$\Lambda_t = \sum_{0\leq i\leq n-1} I(X_i) \qquad \text{si} \quad n\leq t<n+1.$$

<u>Preuve</u>. Notons $H^{\theta\zeta}(x)$ la famille des intégrales de Hellinger du modèle $\mathcal{E}_x$ de [H]: si $\overline{Q}=(Q_\theta+Q_\zeta)/2$ et si $y^\theta(x,y) = Q_\theta(x,dy)/\overline{Q}(x,dy)$ et $y^\zeta(x,dy) = Q_\zeta(x,dy)/\overline{Q}(x,dy)$, on a d'après 1.2:

7.3 $$H^{\theta\zeta}(x) = \int\overline{Q}(x,dy)\ \sqrt{y^\theta(x,y)\ y^\zeta(x,y)}.$$

L'hypothèse [H] entraine (cf. 1.3) que, lorsque $\theta,\zeta\to 0$:

7.4 $$1 + H^{\theta\zeta}(x) - H^{0\theta}(x) - H^{0\zeta}(x) = \frac{1}{4}\,\theta.I(x).\zeta + o(|\theta||\zeta|).$$

Par ailleurs si $\zeta=0$, soit $Y^\theta=y^\theta/y^0$ et $a^\theta(x) = \int Q_\theta(x,dy)[1 - Y^\theta(x,y)]$ (on a $a^\theta\geq 0$, et $a^\theta(x)=0$ si et seulement si $Q_\theta(x,.) \ll Q_0(x,.)$. On peut considérer le modèle $\mathcal{E}_x$ comme filtré par $\mathcal{F}_t = \{\emptyset,\Omega\}$ si $t<1$, $\mathcal{F}_t = \mathcal{E}$ si $t\geq 1$. D'après 7.4, le modèle est alors dérivable à l'instant 1 en $\theta=0$ (cf. 3.2), et 3.4 entraine:

7.5 $$a^\theta(x) = o(|\theta|^2) \qquad \text{quand} \quad \theta\to 0.$$

Passons maintenant au modèle $(\Omega,\mathcal{F},(\mathcal{F}_t),(P_\theta))$. En appliquant IV.1.63 de [5] avec $P=P_\theta$, $P'=P_\zeta$, et pour Q la probabilité faisant de X une chaîne de Markov de transition $\overline{Q}$, et comme $z_t^\theta = \prod_{1\leq i\leq [t]} y^\theta(X_{i-1},X_i)$ et de même pour z^ζ, on voit que

$$h_t^{\theta\zeta} = \sum_{1\leq i\leq [t]} E_Q[1 - \sqrt{y^\theta(X_{i-1},X_i)\, y^\zeta(X_{i-1},X_i)}\,|\mathcal{F}_{t_{i-1}})$$
$$= \sum_{1<i\leq [t]} [1 - H^{\theta\zeta}(X_{i-1})].$$

Donc 7.4 entraine que la condition [A] est satisfaite avec Λ défini par 7.2. Par ailleurs, on a $Z_t^\theta = \prod_{1\leq i\leq [t]} Y^\theta(X_{i-1},X_i)$, donc le processus A^θ de 5.1 est

$$A_t^\theta = \sum_{1\leq i\leq [t]} Z_{i-1}^\theta\, a^\theta(X_{i-1}).$$

Comme $E(\sup_{i\leq n} Z_{i-1}^\theta) \leq 1$, on déduit alors [B] de 7.5, et le résultat découle du théorème 6.2. □

b - Diffusions. On suppose maintenant que, sous P_θ, X est une diffusion uni-dimensionnelle s'écrivant

$$X_t := x_0 + \int_0^t b_s^\theta ds + W_t^\theta,$$

où W^θ est un brownien standard et b^θ est un processus prévisible tel que $\int_0^t (b_s^\theta)^2 ds < \infty$ pour tout $t<\infty$. On sait qu'alors $P_\theta \overset{loc}{\sim} P_\zeta$ pour tout θ,ζ, et d'après [5, IV,.4.23] on a

$$h_t^{\theta\zeta} = \frac{1}{8}\int_0^t (b_s^\theta - b_s^\zeta)^2\, ds.$$

Donc

$$H(\tfrac{1}{2})^{\theta\zeta} = \frac{1}{4}\int_0^t (b_s^\theta - b_s^0)(b_s^\zeta - b_0^0)\, ds.$$

D'après le théorème 6.2 on sait qu'alors le modèle est localement dérivable, dès qu'il existe un processus $\lambda=(\lambda_s)_{s\geq 0}$ à valeurs matricielles symétriques, tel que si $\theta,\zeta \to 0$:

7.6 $$\frac{1}{|\theta||\zeta|}\int_0^t [(b_s^\theta-b_s^0)(b_s^\zeta-b_s^0) - \theta.\lambda_s.\zeta]\, ds \xrightarrow{P} 0 \qquad \forall t\geq 0,$$

et dans ce cas, le processus information est

7.7 $$\Lambda_t = \int_0^t \lambda_s ds.$$

Remarquons d'ailleurs qu'on a 7.6 dès que pour tout ω et tout t l'application $\theta \to b^{\theta}_{.}(\omega)$ est différentiable dans $L^2([0,t],ds)$ en $\theta=0$; si la dérivée est $u_{.}(\omega) = (u^i_{.}(\omega))_{i\leq d}$, on a alors $\lambda^{ij}_s = u^i_s u^j_s$.

c - Processus ponctuels. On suppose enfin que, sous P_θ, X est un processus ponctuel de compensateur $\int_0^t a^\theta_s ds$, et pour simplifier on suppose que $a^\theta_s > 0$ identiquement. D'après [5, IV.4.2] on a alors

$$h^{\theta\zeta}_t = \frac{1}{2}\int_0^t (\sqrt{a^\theta_s} - \sqrt{a^\zeta_s})^2\, ds.$$

On a donc exactement les mêmes résultats qu'au §b, à condition de poser $b^\theta = 2\sqrt{a^\theta}$.

BIBLIOGRAPHIE

[1] BASAWA I.V., PRAKASA RAO B.L.S.: Statistical inference for stochastic processes. Academic Press, New York, 1980.

[2] DACUNHA-CASTELLE D., DUFLO M.: Probabilités et statistiques. Masson, Paris, 1982 (tome 1) et 1983 (tome 2).

[3] EMERY M.: Une topologie sur l'espace des semimartingales. Sém. de Probabilités XIII. Lect. Notes in Math. 721, 260-281. Springer Verlag: Berlin, 1979.

[4] IBRAGIMOV I.A., HAS'MINSKII R.Z.: Statistical Estimation. Springer Verlag: Berlin, 1981.

[5] JACOD J., SHIRYAEV A.N.: Limit theorems for stochastic processes. Springer Verlag: Berlin, 1987.

[6] LECAM L.: Asymptotic methods in statistical decision theory. Springer Verlag: Belin, 1986.

[7] LEPINGLE D.: Une inégalité de martingales. Sém. de Probabilités XII. Lect. Notes in Math. 649, 134-137. Springer Verlag: Berlin, 1978.

[8] MEMIN J.: Espaces de semimartingales et changements de probabilités. Z. Wahrsch. Verw. Geb. 52, 9-40, 1980.

[9] STRASSER H.: Mathematical theory of statistics. De Gruyter: Berlin, 1985.

Notes ajoutées sur les épreuves:

1) On pourrait démontrer le théorème 2.1 plus simplement, en utilisant les remarques qui suivent le théorème 1.11 de l'article suivant:

[10] STRICKER C.: Quelques remarques sur la topologie des semimartingales, applications aux intégrales stochastiques. Sém. de Probabilités XV, Lect. Notes in Math. 850, 499-522, Springer Verlag, Berlin: 1981.

2) Il semble que la notion de processus d'information apparaisse pour les première fois dans les articles suivants:

[11] FEIGIN P.D.: Maximum likelihood estimation for continuous-time stochastic processes. Adv. Appl. Prob. 8, 712-736, 1976.

[12] HEYDE C.C.: Remarks on efficiency in estimation for branching processes. Boimetrika 62, 49-55, 1975.