ANNALES DE L'I. H. P., SECTION B

JEAN JACOD

Sur le processus de vraisemblance partielle

Annales de l'I. H. P., section B, tome 26, nº 2 (1990), p. 299-329

<http://www.numdam.org/item?id=AIHPB_1990__26_2_299_0>

Ann. Inst. Henri Poincaré,

Vol. 26, n° 2, 1990, p. 299-329. *Probabilités et Statistiques*

Sur le processus de vraisemblance partielle

par

Jean JACOD

Laboratoire de Probabilités, Université Pierre-et-Marie-Curie,
tour n° 56, 3[e] étage, 4, place Jussieu, 75252 Paris Cedex 05, France

---

1 RÉSUMÉ. — Nous proposons une définition du processus de vraisemblance partielle pour des semi-martingales générales, qui étend la définition donnée en [10] et qui nous semble mieux adaptée aux problèmes statistiques « semi-paramétriques » : cette assertion est illustrée par diverses applications statistiques. Nous montrons aussi que cette vraisemblance partielle jouit d'agréables propriétés d'invariance : par exemple, elle ne change pas si on remplace la semi-martingale de base X par $Y=f(X)$, où f est une fonction inversible de classe C^2. Finalement, nous montrons que les résultats de normalité asymptotique de [10] restent vrais pour cette définition plus générale.

Mots clés : Processus de vraisemblance, vraisemblance partielle, semi-martingales, normalité asymptotique.

ABSTRACT. — We propose a definition of partial likelihood processes for general semimartingales, which extends the definition given in [10] and is better fit for statistical problems: we illustrate this statement with some statistical applications. We also prove that partial likelihoods enjoy nice invariance properties: for example they do not change if we replace the basic semimartingale X by $Y=f(X)$ with an invertible C^2 function f. Finally, we observe that the asymptotic normality result of [10] remains true in the present setting.

Key words : Likelihood processes, partial likelihood, semimartingales, asymptotic normality.

Classification A.M.S. : .

1. INTRODUCTION

On considère un modèle statistique « filtré » : l'espace d'état est Ω, l'ensemble des temps est $T=\mathbb{N}$ ou $T=\mathbb{R}_+$, on note $\mathcal{F}_t$ la tribu des événements observables jusqu'à l'instant t, et $\mathcal{F}=\bigvee\mathcal{F}_t$. On s'intéresse à la valeur d'un paramètre $\theta\in\Theta$.

Dans le cas « idéal » la valeur de θ détermine la probabilité P_θ, et le rôle essentiel est joué par les *processus de vraisemblance* $Z_t^{\zeta/\theta}=(dP_\zeta/dP_\theta)_{|\mathcal{F}_t}$.

Lorsque θ ne détermine pas complètement la probabilité, Cox a suggéré dans [3] et [4] (pour $T=\mathbb{N}$) de remplacer les $Z^{\zeta/\theta}$ par des *processus de vraisemblance partielle* (PVP), quand on est dans la situation suivante : on observe un processus « de base » $X=(X_t)_{t\in T}$ (disons, à valeurs réelles), plus un ou plusieurs autres processus auxiliaires, symbolisés par $Y=(Y_t)_{t\in T}$; et, de manière un peu vague, le paramètre θ ne détermine pas complètement la probabilité P_θ, mais seulement la « loi conditionnelle de X si on connaît Y ». Cela permet d'englober certains modèles avec paramètre nuisible : la probabilité $P_{\theta,\eta}$ dépend de θ et d'un *paramètre nuisible* η qui décrit le comportement du processus auxiliaire Y ; cela permet aussi de considérer la situation *semi-paramétrique* où on ne veut pas préciser la loi de Y.

Exemple 1 : *le cas discret.* — On suppose que $T=\mathbb{N}$ et, pour simplifier, que $\mathcal{F}_0$ est la tribu triviale. C'est le cas étudié par Cox ([3], [4]) et Wong [17]. L'hypothèse de base est que, pour chaque $n\geqq 1$, le paramètre θ détermine la loi conditionnelle $\mu_{\theta,n}(\omega, dx)$ de X_n-X_{n-1} par rapport à $\mathcal{F}_{n-1}$. Si on suppose en plus que $\mu_{\theta,n}(\omega,.)$ admet la densité $f_{\theta,n}(\omega,.)$ par rapport à la mesure de Lebesgue, Cox propose d'appeler PVP de P_ζ par rapport à P_θ, relativement à X, le processus

$$\bar{Z}_n^{\zeta/\theta}(\omega)=\prod_{1\leqq p\leqq n}\frac{f_{\zeta,p}(\omega, X_p(\omega)-X_{p-1}(\omega))}{f_{\theta,p}(\omega, X_p(\omega)-X_{p-1}(\omega))}. \tag{1.1}$$

Bien entendu, si $\mathcal{F}_n=\sigma(X_p:p\leqq n)$, les $(\mu_{\theta,n})_{n\geqq 1}$ déterminent entièrement la probabilité P_θ, et le processus (1.1) est le véritable processus de vraisemblance $Z^{\zeta/\theta}$ de P_ζ par rapport à P_θ.

En général les $\mu_{\theta,n}$ ne déterminent pas P_θ, mais plutôt la famille $\mathcal{S}_\theta$ de probabilités pour lesquelles les $\mu_{\theta,n}$ sont les lois conditionnelles de X_n-X_{n-1} si $\mathcal{F}_{n-1}$. Cox montre cependant que les PVP se comportent comme les vrais processus de vraisemblance : par exemple, $\bar{Z}^{\zeta/\theta}$ est une P_θ-surmartingale, pour tout choix de P_θ dans $\mathcal{S}_\theta$; et, sous des hypothèses convenables (comme la normalité asymptotique des logarithmes des PVP) les estimateurs de « maximum de vraisemblance partielle » ont les mêmes propriétés de consistance, normalité asymptotique, etc. que le maximum de vraisemblance classique. □

Exemple 2 : les processus ponctuels. — Ici $T=\mathbb{R}_+$ et le processus de base X est un processus ponctuel. L'hypothèse essentielle est que θ détermine le compensateur A^θ de X, qui se met sous la forme

$$A_t^\theta(\omega)=\int_0^t \alpha^\theta(\omega,s)\,ds. \tag{1.2}$$

Dans ce cas, Andersen et Gill [1], Arjas et Haara [2], Gill [6], Slud [14], proposent de prendre pour PVP de P_ζ par rapport à P_θ le processus suivant, où $T_1<\ldots<T_n<\ldots$ désignent les temps de saut successifs de X :

$$\bar{Z}^{\zeta/\theta}(\omega)_t=\left(\prod_{T_n\leqq t}\frac{\alpha^\zeta(\omega,T_n)}{\alpha^\theta(\omega,T_n)}\right)\exp\int_0^t[\alpha^\theta(\omega,s)-\alpha^\zeta(\omega,s)]\,ds. \tag{1.3}$$

Là encore ce processus est la véritable vraisemblance lorsque $(\mathscr{F}_t)$ est la filtration engendrée par X.

En général, $\bar{Z}^{\zeta/\theta}$ n'est pas la vraisemblance, mais c'est une P_θ-surmartingale pour toute probabilité P_θ appartenant à la classe $\mathscr{S}_\theta$ des probabilités sous lesquelles X admet A^θ pour compensateur. □

Dans cet article nous considérons la situation bien plus générale où le processus de base X est une semi-martingale. Les deux exemples précédents suggèrent que le cadre naturel où les PVP s'introduisent est celui où le paramètre θ détermine le triplet τ_θ des caractéristiques de X. Notons aussi que le problème de la définition du PVP concerne uniquement les couples θ, ζ de valeurs du paramètre.

On va donc étudier le problème suivant : soit X un processus adapté sur un espace filtré $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqq 0})$; soit également $\mathscr{T}$ et $\mathscr{T}'$ deux triplets de caractéristiques « potentielles ». On note $\mathscr{S}$ (resp. $\mathscr{S}'$) l'ensemble des probabilités sur $(\Omega, \mathscr{F})$ sous lesquelles X est une semi-martingale de caractéristiques $\mathscr{T}$ (resp. $\mathscr{T}'$) : ce sont donc les ensembles de solutions de certains problèmes de martingales (*cf.* [13], § III.2). Comment définir le PVP de $P'\in\mathscr{S}'$ par rapport à $P\in\mathscr{S}$?

Sous certaines conditions (impliquant notamment le fait que $\mathscr{T}$ et $\mathscr{T}'$ déterminent P et P' : (*cf.* [13], § III.5)), on sait exprimer la vraisemblance de P' par rapport à P en termes des triplets $\mathscr{T}$ et $\mathscr{T}'$ et du processus X lui-même, par une « formule de Girsanov ». L'idée essentielle consiste alors à définir, en toute généralité, le PVP de $P' \in \mathscr{S}'$ par rapport à $P \in \mathscr{S}$ par cette formule de Girsanov (comme dans les exemples 1 et 2).

Dans [10] nous avions proposé cette solution, en adoptant le point de vue suivant : fixons P et P′, et appelons Z le vrai processus de vraisemblance; le triplet $\mathscr{T}'$ n'est défini de manière unique qu'en dehors d'un ensemble P′-négligeable. Donc on peut considérer qu'il n'est P-p. s. bien défini que sur l'intervalle stochastique $\{ t : Z_{t-} > 0 \}$, et nous n'avions donc défini le PVP $\bar{Z}$ que sur ce même intervalle. Ce point de vue est naturel si on considère qu'on connaît P et P′; il ne l'est plus si on considère qu'on sait seulement que P et P′ appartiennent à $\mathscr{S}$ et $\mathscr{S}'$ (en effet, dans ce cas P et P′ sont souvent étrangères sur chaque tribu $\mathscr{F}_t$, donc $Z_t = 0$ P-p. s.).

Ici, nous définissons le PVP $\bar{Z}$ sur un intervalle « maximal » dépendant seulement de $\mathscr{T}$ et $\mathscr{T}'$: c'est l'objet du paragraphe 2; on peut noter que $\bar{Z}$ dépend de $\mathscr{T}$ et $\mathscr{T}'$, et aussi de P (dans la mesure où son expression fait intervenir des intégrales stochastiques), mais nullement de P′, et son expression formelle ne dépend pas de P. Nous illustrons cette notion sur un exemple statistique simple au paragraphe 3, et le paragraphe 4 est consacré à des propriétés naturelles d'invariance : on ne modifie pas le PVP si on remplace X par un processus Y de la forme $Y = \rho . X$ (intégrale stochastique de ρ par rapport à X) ou $Y = f(X)$, lorsque ces transformations sont « inversibles ».

Les paragraphes suivants sont de nature plus technique : au paragraphe 5 nous montrons que l'intervalle maximal de définition du PVP $\bar{Z}$ contient l'intervalle $\{ Z_- > 0 \}$ (Z = véritable vraisemblance), et que sur ce dernier $\bar{Z}$ coïncide avec la version donnée en [10]. Le paragraphe 6 est consacré au « calcul » de $\bar{Z}$ à partir de Z sur l'intervalle $\{ Z_- > 0 \}$: on peut dire que $\bar{Z}$ est une sorte de « projection » de Z sur l'espace des martingales engendré par le processus X. Enfin au paragraphe 7 nous étendons les processus de Hellinger partiels et le résultat de normalité asymptotique de [10].

Les notations sont celles usuelles en théorie générale ([5], [13]) : $H . Y_t = \int_0^t H_s \, dY_s$ (Y = semi-martingale) et $W * \eta_t = \int_0^t \int_E W(s, x) \, \eta(ds, dx)$ (η = mesure aléatoire sur $\mathbb{R}_+ \times E$) sont les processus intégrale stochastique, Y^c désigne la partie martingale continue de la semi-martingale Y, etc. Les vecteurs sont des matrices colonne, et la transposée de A est notée A^T.

2. DÉFINITION DU PROCESSUS DE VRAISEMBLANCE PARTIELLE

On se donne un espace filtré $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \geq 0})$, muni de :
(2.1) un processus càdlàg adapté q-dimensionnel $X = (X^i)_{i \leq q}$ avec $X_0 = 0$;
(2.2) deux triplets $\mathscr{T} = (B, C, \nu)$ et $\mathscr{T}' = (B', C', \nu')$ de caractéristiques potentielles pour X, relativement à une fonction de troncation fixée h sur $\mathbb{R}^q$.

Rappelons ce que veut dire (2.2). D'abord, h est une fonction de $\mathbb{R}^q$ dans $\mathbb{R}^q$, à support compact, bornée, égale à l'identité sur un voisinage de 0; ensuite, $B = (B^i)_{i \leq q}$ est un processus prévisible q-dimensionnel à variation finie avec $B_0 = 0$; $C = (C^{ij})_{i, j \leq q}$ est un processus adapté continu croissant dans l'ensemble des matrices symétriques non négatives, avec $C_0 = 0$; ν est une mesure aléatoire prévisible sur $\mathbb{R}_+ \times \mathbb{R}^q$ vérifiant

$$(|x|^2 \wedge 1) * \nu_t < \infty, \qquad \nu(\{t\} \times \mathbb{R}^q) \leq 1, \qquad \nu(\{0\} \times \mathbb{R}^q) = 0,$$

$$\nu(\mathbb{R}_+ \times \{0\}) = 0 \qquad \text{et} \qquad \Delta B_t = \int h(x)\, \nu(\{t\} \times dx).$$

On note $\mathscr{S}$ (resp. $\mathscr{S}'$) l'ensemble des probabilités sur $(\Omega, \mathscr{F})$ faisant de X une semi-martingale de caractéristiques $\mathscr{T}$ (resp. $\mathscr{T}'$).

On se donne $P \in \mathscr{S}$ et $P' \in \mathscr{S}'$. Afin de définir le PVP de P' par rapport à P, il nous faut une série de notations assez lourdes, semblables à celles du paragraphe IV.3.a de [13].

D'abord, il existe un processus continu adapté croissant F et un processus prévisible $c = (c^{ij})_{i, j \leq q}$ à valeurs dans l'ensemble des matrices symétriques non négatives, tels que

$$(2.3) \qquad C^{ij} = c^{ij} . F, \quad \text{P-p. s.}$$

On note $\mathscr{P}$ la tribu prévisible sur $\Omega \times \mathbb{R}_+$; soit la tribu $\tilde{\mathscr{P}} = \mathscr{P} \otimes \mathscr{R}^q$ sur $\tilde{\Omega} = \Omega \times \mathbb{R}_+ \times \mathbb{R}^q$. Posons

$$(2.4) \qquad a_t = \nu(\{t\} \times \mathbb{R}^q), \qquad a'_t = \nu'(\{t\} \times \mathbb{R}^q).$$

On considère une mesure prévisible η qui domine ν et ν', par exemple $\eta = \nu + \nu'$. Il existe deux fonctions positives $\tilde{\mathscr{P}}$-mesurables y, y' sur $\tilde{\Omega}$ telles que

$$(2.5) \qquad \nu = y . \eta, \qquad \nu' = y' . \eta, \quad \text{P-p. s.}$$

Les quantités suivantes ne dépendent P-p. s. pas du choix de η, y, y' dans (2.5) :

$$
(2.6)\qquad \left\{
\begin{aligned}
&G=(\sqrt{y}-\sqrt{y'})^2 * \eta + \sum_{s\leqq .} (\sqrt{1-a_s}-\sqrt{1-a'_s})^2,\\
&\Sigma_1=\{ G<\infty \}.
\end{aligned}\right.
$$

Remarquer que G est prévisible croissant, et c.-à-d. excepté peut-être au point $S=\inf(t: G_t=\infty)$. On a

$$
\begin{aligned}
|y-y'|\,|h| &\leqq \frac{1}{2}(\sqrt{y}-\sqrt{y'})^2+\frac{1}{2}|h|^2(\sqrt{y}+\sqrt{y'})^2\\
&\leqq \frac{1}{2}(\sqrt{y}-\sqrt{y'})^2+|h|^2(y+y');
\end{aligned}
$$

comme $|h|^2 y * \eta_t=|h|^2 * \nu_t<\infty$ et de même $|h|^2 y' * \eta_t<\infty$, le processus $|(y'-y)h| * \eta$ est fini sur Σ_1, et le processus $B'-B-(y'-y)h * \eta$ est donc bien défini sur Σ_1. Il admet alors une (P-p. s.) unique décomposition en une somme de trois termes :

$$
(2.7)\qquad B'-B-(y'-y)h * \eta=(c\,\beta).F+\tilde{\beta}.F+F' \quad \text{sur } \Sigma_1,
$$

où

- $\beta=(\beta^i)_{i\leqq q}$ et $\tilde{\beta}=(\tilde{\beta}^i)_{i\leqq q}$ sont prévisibles,
- $\tilde{\beta}_t(\omega)$ est orthogonal dans $\mathbb{R}^q$ à l'image de $\mathbb{R}^q$ par l'application linéaire associée à la matrice $c_t(\omega)$,
- les composantes F'^i sont prévisibles à variation finie sur les compacts de Σ_1, et les mesures dF'^i_t et dF_t sont étrangères sur E_1.

Un calcul simple montre que F′ est continu. Posons enfin

$$
(2.8)\qquad
\begin{aligned}
&\Sigma_2=\{ t\in\Sigma_1 : (\beta^{T} c\,\beta).F_t<\infty \},\\
&\Sigma_3=\{ t\in\Sigma_2 : \tilde{\beta}.F_s=0 \text{ et } F'_s=0\ \forall s\leqq t\}\\
&\Sigma_4=\{ t\in\Sigma_2 : C'_s=C_s\ \forall s\leqq t\}, \qquad \Sigma_5=\Sigma_3\cap\Sigma_4.
\end{aligned}
$$

Les Σ_i sont des ensembles prévisibles de la forme $\bigcup_n [0, S_n]$, pour des temps d'arrêt S_n convenables. Noter que les ensembles Σ_i sont définis de manière non ambiguë (à un ensemble P-négligeable près) indépendamment du choix de c, F, y, y', β, $\tilde{\beta}$ dans (2.3), (2.5), (2.7).

(2.9) *Remarque.* — La probabilité P intervient dans (2.3), (2.5) et (2.7). Si on ne veut pas la faire intervenir à ce stade, il faut faire l'hypothèse (plutôt anodine) qu'on a (2.5) identiquement avec y, y' prévisibles (sans aucune hypothèse, on a toujours (2.5) identiquement avec y, y' optionnels, mais pas forcément prévisibles). Quant à (2.3) et (2.7), ils sont toujours vérifiés identiquement avec c, β, $\tilde{\beta}$ optionnels, ce qui est sans inconvénient pour la suite car ces processus interviennent comme intégrands par rapport à des martingales continues. □

Encore un dernier rappel : on a $C^{ij}=\langle X^{c,i}, X^{c,j}\rangle$ (le crochet et la partie martingale continue X^c sont relatifs à P) et ν est le P-compensateur de la mesure des sauts de X, définie par

$$\mu(\omega; dt, dx)=\sum_{s:(\omega,s)\in D}\varepsilon_{(s,\Delta X_s(\omega))}(dt, dx), \tag{2.10}$$

où

$$D=\{\Delta X\neq 0\}.$$

On a aussi P-p. s. :

$$1_{\{y=0\}}*\mu=1_{\{y=0\}}*\nu=0, \qquad a_t<1 \quad \text{pour tout } t\in D^c \tag{2.11}$$

(en effet, $1_{\{y=0\}}*\nu=0$ vient de (2.5); $1_{\{y=0\}}*\mu$ admet $1_{\{y=0\}}*\nu$ pour P-compensateur; $1-a$ est la P-projection prévisible du processus 1_{D^c} par la proposition II.1.7 de [13]).

(2.12) LEMME. — *Les formules*

$$\bar{A}=y'\,1_{\{y=0\}}*\eta+\sum_{s\leqq\,.}(1-a'_s)\,1_{\{a_s=1\}}, \tag{2.13}$$

$$\bar{M}=\beta^T.X^c+U*(\mu-\nu), \tag{2.14}$$

où

$$U=\frac{y'}{y}1_{\{y>0\}}-1+\frac{a'-a}{1-a}1_{\{a<1\}},$$

$$\bar{L}=\bar{M}-\bar{A}, \tag{2.15}$$

définissent un processus croissant prévisibles fini $\bar{A}$ *sur* Σ_2, *une* P-*martingale locale* $\bar{M}$ *sur* Σ_2, *et une* P-*surmartingale locale* $\bar{L}$ *sur* Σ_2 *vérifiant* $\Delta\bar{L}\geqq -1$ *et*

$$\Delta\bar{L}_t=\left(\frac{y'}{y}(t,\Delta X_t)-1\right)1_D(t)+\frac{a_t-a'_t}{1-a_t}1_{D^c}(t). \tag{2.16}$$

Étant donné (2.11), la formule (2.16) a un sens P-p. s.; comme $\Sigma_2=\bigcup_n[0,S_n]$ pour des temps d'arrêt S_n, dire que $\bar{M}$ (resp. $\bar{L}$) est une P-martingale (resp. surmartingale) locale sur Σ_2 signifie que les processus arrêtés $\bar{M}^{S_n}$ (resp. $\bar{L}^{S_n}$) sont des P-martingales (resp. surmartingales) locales, et ceci ne dépend pas de la suite (S_n) choisie.

Noter aussi que $\bar{A}$, $\bar{M}$ et $\bar{L}$ ne dépendent pas du choix de y, y' dans (2.5), ni de β dans (2.7).

(2.17) *Remarque.* — Le processus $\bar{L}$ peut aussi être caractérisé ainsi : c'est l'unique P-surmartingale locale sur Σ_2 telle que $\bar{L}^c = \beta^T . X^c$, que $\Delta\bar{L}$ soit donné par (2.16), et que le processus $\bar{A}$ de sa décomposition de Doob-Meyer $\bar{L} = \bar{M} - \bar{A}$ soit donné par (2.13). □

Preuve de (2.12). — On a $\bar{A} \leqq G$, d'où la première assertion. Comme $c^{ij} . F = \langle X^{c, i}, X^{c, j} \rangle$ et $(\beta^T c\, \beta) . F < \infty$ sur Σ_2, la première intégrale stochastique de (2.14) définit une P-martingale locale continue sur Σ_2.

Pour la seconde intégrale, introduisons d'abord la notation suivante, pour toute fonction V sur $\tilde{\Omega}$:

$$\hat{V}_t = \int_{\mathbb{R}^q} V(t, x)\, \nu(\{t\} \times dx); \tag{2.18}$$

puis, U étant défini par (2.14), on pose $\tilde{U}_t = U(t, \Delta X_t)\, 1_D(t) - \hat{U}_t$ et

$$G' = ((1 + U - \hat{U})^{1/2} - 1)^2 * \nu + \sum_{s \leqq .} (1 - a_s)((1 - \hat{U}_s)^{1/2} - 1)^2.$$

Appliquons alors le résultat II.1.36 de [13] : si $\tilde{U} \geqq -1$ et si le processus G' est fini sur Σ_2, alors l'intégrale stochastique $U * (\mu - \nu)$ est bien définie sur Σ_2 ; de plus, c'est une martingale locale sur Σ_2 dont le processus de sauts égale $\tilde{U}$.

Posons $H_t = \int y'(t, x)\, 1_{\{y(t, x) > 0\}}\, \eta(\{t\} \times dx)$, de sorte que $\hat{U} = H - a + a \dfrac{a' - a}{1 - a} 1_{\{a < 1\}}$, donc $1 - \hat{U} = a' - H + \dfrac{1 - a'}{1 - a}$ si $a < 1$, et $1 + U - \hat{U} = \dfrac{y'}{y} 1_{\{y > 0\}} - H + 1_{\{a = 1\}} + a'\, 1_{\{a < 1\}}$. Comme $0 \leqq H \leqq a' \leqq 1$, il vient $1 + \tilde{U} \geqq 0$. Comme $|\sqrt{x + z} - \sqrt{x}|^2 \leqq z$ pour x, $z \geqq 0$ et comme $\nu = y . \eta$, on a aussi

$$\begin{aligned} G' = {} & \left[\left(\frac{y'}{y} - H + 1_{\{a=1\}} + a'\, 1_{\{a<1\}} \right)^{1/2} - \left(\frac{y'}{y} \right)^{1/2} + \left(\frac{y'}{y} \right)^{1/2} - 1 \right]^2 * \nu \\ & + \sum_{s \leqq .} [\sqrt{1 - a'_s} - \sqrt{1 - a_s} - \sqrt{1 - a'_s} + \sqrt{(1 - a_s)(a'_s - H_s) + 1 - a'_s}]^2 \\ & \leqq 2 (\sqrt{y} - \sqrt{y'})^2 * \eta + 2 (1_{\{a=1\}} + a'\, 1_{\{a<1\}} - H) * \nu \\ & \quad + 2 \sum_{s \leqq .} (\sqrt{1 - a'_s} - \sqrt{1 - a_s})^2 + 2 \sum_{s \leqq .} (1 - a_s)(a'_s - H_s) \\ & = 2G + 2 \sum_{s \leqq .} (1 - H_s)\, 1_{\{a_s = 1\}} + 2 \sum_{s \leqq .} (a'_s - H_s)\, 1_{\{a_s < 1\}}. \end{aligned}$$

Mais $a'-\mathrm{H}=\int y'(t,x)\,1_{\{y(t,x)=0\}}\,\eta(\{t\}\times dx)\leqq \Delta\mathrm{G}$ sur Σ_1 par (2.5) et (2.6), donc $\mathrm{G}'\leqq 4\mathrm{G}+2\sum_{s\leqq .} 1_{\{a_s=1\}}$, qui est fini sur Σ_1.

On a ainsi montré que $\bar{\mathrm{M}}$, et donc $\bar{\mathrm{L}}$, sont bien définis sur Σ_2. De plus $\Delta\bar{\mathrm{M}}=\tilde{\mathrm{U}}$, et comme $\Delta\bar{\mathrm{A}}=a'-\mathrm{H}+(1-a')\,1_{\{a=1\}}$, un calcul simple utilisant les propriétés (2.11) montre qu'on a (2.16), donc $\Delta\bar{\mathrm{L}}\geqq -1$. □

(2.19) DÉFINITION. — On appelle *processus de vraisemblance partielle* (PVP) de P′ par rapport à P, relativement à X et aux triplets (B, C, ν) et (B′, C′, ν′), le processus défini sur Σ_2 par

$$(2.20)\qquad \bar{\mathrm{Z}}=\exp\left(\bar{\mathrm{L}}-\frac{1}{2}(\beta^{\mathrm{T}}c\,\beta).\mathrm{F}\right)\prod_{s\leqq .}[(1+\Delta\bar{\mathrm{L}}_s)\exp(-\Delta\bar{\mathrm{L}}_s)],$$

c'est-à-dire $\bar{\mathrm{Z}}=\mathscr{E}(\bar{\mathrm{L}})$ (exponentielle de Doléans de $\bar{\mathrm{L}}$). □

Comme $\bar{\mathrm{L}}$ est une P-surmartingale sur Σ_2 avec $\Delta\bar{\mathrm{L}}\geqq -1$, on voit que $\bar{\mathrm{Z}}$ est une P-*surmartingale locale positive sur* Σ_2. Si $\Sigma_2=\Omega\times\mathbb{R}_+$, $\bar{\mathrm{Z}}$ est donc une P-*surmartingale positive.*

Remarquons que la dénomination : PVP de P′ par rapport à P, est un peu impropre, car la probabilité P′ ne joue aucun rôle dans cette définition. On peut aussi noter que $\bar{\mathrm{Z}}$ et Σ_2 ne dépendent pas du choix de c, η, y, y', β dans (2.3), (2.5) et (2.7).

L'ensemble Σ_2 est l'ensemble maximal sur lequel $\bar{\mathrm{Z}}$ peut être défini (car les intégrales stochastiques de (2.14) n'ont pas de sens sur son complémentaire). Toutefois nous verrons au paragraphe 4 qu'à certains égards il serait plus naturel de ne définir $\bar{\mathrm{Z}}$ que sur Σ_3 ou Σ_5.

Exemple 1. — Reprenons l'exemple 1 de l'introduction. Le cas discret se ramène au cas continu en posant $\mathrm{X}_t=\mathrm{X}_n$ et $\mathscr{F}_t=\mathscr{F}_n$ pour $n\leqq t<n+1$. Dire qu'on connaît les lois conditionnelles $\mu_{\theta,n}$ revient à se donner le triplet $\mathscr{T}_\theta=(\mathrm{B}^\theta, \mathrm{C}^\theta, \nu^\theta)$, avec (si $[t]$ = partie entière de t) :

$$\mathrm{C}_t^\theta=0,\quad \mathrm{B}_t^\theta=\sum_{p=1}^{[t]}\mu_{\theta,p}(h),$$
$$\nu^\theta(dt,dx)=\sum_{p\geqq 1}\varepsilon_p(dt)\,\mu_{\theta,p}(dx)\,1_{\{x\neq 0\}}.$$

Sous l'hypothèse où $\mu_{\theta,p}(dx)=f_{\theta,p}(x)\,dx$, on a (2.5) avec $y^\theta(\omega,t,x)=f_{\theta,p}(\omega,x)$ pour $p\leqq t<p+1$, à condition de prendre $\eta=\sum_{p\geqq 1}\varepsilon_p(dt)\,dx$, et aussi $a_t^\theta=1$ dans (2.4).

Comparons alors $\mathcal{T}=\mathcal{T}_\theta$ et $\mathcal{T}'=\mathcal{T}_\zeta$. Il vient $\Sigma_1=\Omega\times\mathbb{R}_+$ dans (2.6), et le premier membre de (2.7) est nul, de sorte que $\Sigma_i=\Omega\times\mathbb{R}_+$ pour $2\leqq i\leqq 5$. D'après les formules du lemme (2.12), $\bar{\mathrm{L}}$ est purement discontinu et ne saute qu'aux instants entiers, tandis que les sauts sont donnés par (2.16). On a donc

$$\bar{\mathrm{L}}_t=\sum_{p=1}^{[t]}\left(\frac{f_{\zeta,p}(\mathrm{X}_p-\mathrm{X}_{p-1})}{f_{\theta,p}(\mathrm{X}_p-\mathrm{X}_{p-1})}-1\right),$$

de sorte que d'après (2.20) le PVP est bien donné par (1.1). □

Exemple 2. — Reprenons l'exemple 2 de l'introduction. Pour h choisie de sorte que $h(1)=0$, et sous l'hypothèse (1.2), le triplet $\mathcal{T}_\theta$ est

$$\mathrm{B}^\theta=0,\quad \mathrm{C}^\theta=0,$$
$$\nu^\theta(\omega\,;dt,dx)=\alpha^\theta(\omega,t)\,dt\,\varepsilon_1(dx).$$

Si alors $\mathcal{T}=\mathcal{T}_\theta$ et $\mathcal{T}'=\mathcal{T}_\zeta$, on vérifie aisément que $\Sigma_i=\Omega\times\mathbb{R}_+$ (prendre $\eta=dt\,\varepsilon_1(dx)$, $y=\alpha^\theta$ et $y'=\alpha^\zeta$ dans (2.5)) et que $a=a'=0$ dans (2.4).

Les formules (2.13) et (2.14) donnent

$$\bar{\mathrm{A}}_t=\int_0^t \alpha^\zeta\, 1_{\{\alpha^\theta=0\}}\,ds,$$

$$\bar{\mathrm{M}}_t=\sum_{\mathrm{T}_n\leqq t}(\alpha^\zeta/\alpha^\theta-1)(\mathrm{T}_n)-\int_0^t(\alpha^\zeta\, 1_{\{\alpha^\theta>0\}}-\alpha^\theta)\,ds.$$

Donc d'après (2.15) et (2.20) on voit que le PVP est donné par (1.3). □

Pour terminer ce paragraphe, et afin de motiver la définition (2.19) dans un cadre plus général que celui des deux exemples ci-dessus, supposons que $\mathrm{P}'\overset{\mathrm{loc}}{\ll}\mathrm{P}$ (*i. e.* $\mathrm{P}'\ll\mathrm{P}$ en restriction à chaque $\mathcal{F}_t$) et appelons Z le véritable processus de vraisemblance. Supposons qu'on ait la propriété de représentation prévisible pour les P-martingales, relativement à X, au sens de [13], III.4.22. D'après la version III.5.19 de [13] du théorème de Girsanov, on a alors $\bar{\mathrm{Z}}=\mathrm{Z}$ sur Σ_2, donc le *PVP coïncide avec la véritable vraisemblance* là où il est bien défini. Remarquons d'ailleurs que dans ce cas, $\mathrm{Z}=0$ sur $(\Sigma_2)^c$, et $\mathrm{Z}_->0$ sur Σ_5, et $\bar{\mathrm{A}}=0$ P-p. s. sur $\mathbb{R}_+$.

Signalons enfin que Slud [15] a donné une autre justification de (2.19) (lorsque X est un processus ponctuel ou une diffusion) en faisant une approximation « en temps discret » et en montrant que les PVP de Cox (exemple 1 ci-dessus) convergent vers le PVP de (2.19) lorsque le pas de la subdivision tend vers 0.

3. UN EXEMPLE D'APPLICATION A L'ESTIMATION

Nous allons considérer ci-dessous un modèle très simple; il s'agit d'un cas particulier des modèles proposés par Hutton et Nelson [8] (le processus de base est unidimensionnel et continu), mais étendu à la situation où on observe également des processus auxiliaires. Les résultats de ce paragraphe sont donc des extensions immédiates de ceux de Hutton et Nelson, mais ils montrent bien l'usage qu'on peut faire des PVP au sens de (2.19), alors que les PVP au sens de [10] et *a fortiori* les véritables processus de vraisemblance seraient inopérants.

Nous étudierons de manière systématique dans [12] les rapports entre les PVP et les quasi-vraisemblances au sens de Hutton et Nelson, de Godambe et Heyde [7] ou de Sorensen [16].

On prend pour Ω l'espace canonique des fonctions sur $\mathbb{R}_+$ à valeurs dans $\mathbb{R}\times[1,2]\times[1,2]$, continues, avec le processus canonique (X, U, V) et la filtration canonique $(\mathscr{F}_t)$. L'espace des paramètres est $\Theta=\mathbb{R}$, et on observe le triplet (X, U, V). Pour la valeur θ on se donne les caractéristiques suivantes pour X :

$$(3.1)\qquad B_t^\theta=\theta\int_0^t U_s\,ds,\qquad C_t^\theta=f(\theta)\int_0^t V_s\,ds,\qquad \nu^\theta=0,$$

où f est une fonction connue : $\mathbb{R}\to]0,\infty[$.

On ne veut pas préciser la loi de U et V (modèle semi-paramétrique), donc P_θ désigne un élément quelconque de l'ensemble $\mathscr{S}_\theta$ des probabilités sous lesquelles X est une semi-martingale de caractéristiques $\mathscr{T}_\theta=(B^\theta, C^\theta, \nu^\theta)$. En d'autres termes, sous P_θ le processus X vérifie

$$dX_t=\theta\,U_t\,dt+\sqrt{f(\theta)\,V_t}\,dW_t,\qquad \text{W mouvement brownien linéaire.}$$

Remarquons tout de suite qu'on ne peut pas (en général) utiliser pour ce modèle les véritables vraisemblances, car :

(1) on ne connaît pas complètement les P_θ, en général;

(2) même si on impose la « trivialité » des processus U et V (par exemple $U_t=V_t=1$ identiquement), les mesures P_θ sont deux à deux étrangères sur $\mathscr{F}_0=\mathscr{F}_{0+}$, à moins que la fonction f ne soit constante (*cf.* 3.6-(2) ci-dessous).

Pour construire le PVP $\bar{Z}^{\zeta/\theta}$ de P_ζ par rapport à P_θ, on définit $\Sigma_i^{\zeta/\theta}=\Sigma_i$ et $\beta^{\zeta/\theta}=\beta$ comme au paragraphe 2, avec $(B, C, \nu)=(B^\theta, C^\theta, \nu^\theta)$ et $(B', C', \nu')=(B^\zeta, C^\zeta, \nu^\zeta)$. Il vient $\Sigma_2^{\zeta/\theta}=\mathbb{R}_+$ et $\beta^{\zeta/\theta}=\dfrac{\zeta-\theta}{f(\theta)}\dfrac{U}{V}$. Donc si on pose

$$(3.2)\qquad Y_t=\int_0^t (U_s/V_s)\,dX_s,\qquad H_t=\int_0^t (U_s^2/V_s)\,ds,$$

il vient finalement

$$(3.3)\qquad \bar{Z}_t^{\zeta/\theta}=\exp\left\{\frac{\zeta-\theta}{f(\theta)}Y_t-\frac{\zeta^2-\theta^2}{2f(\theta)}H_t\right\}\quad \text{pour tout } t\geqq 0$$

(dans cette formule, le processus Y est une intégrale stochastique prise par rapport à la mesure P_θ).

Supposons alors qu'on calcule l'estimateur du maximum de « vraisemblance partielle » à l'instant t, c'est-à-dire le point $\hat{\theta}_t$ qui maximise la fonction $\zeta \to \bar{Z}_t^{\zeta/\theta}$. *A priori*, $\hat{\theta}_t$ devrait dépendre du point de référence θ, mais d'après (3.3) on obtient immédiatement la valeur suivante, indépendante de θ :

$$(3.4)\qquad \hat{\theta}_t=Y_t/H_t.$$

Supposons alors que la vraie valeur du paramètre soit θ. L'intégrale stochastique donnant Y est alors prise naturellement par rapport à P_θ, et d'après (3.2) on peut alors écrire

$$(3.5)\qquad \hat{\theta}_t=\theta+M_t/H_t,$$

où $M=(U/V).(X-B^\theta)$ est une P_θ-martingale de crochet $\langle M, M\rangle=f(\theta)H$. Comme $H_t\uparrow\infty$ quand $t\uparrow\infty$, la loi des grands nombres pour les martingales continues entraîne que $M_t/H_t\to 0$ P_θ-p. s. quand $t\uparrow\infty$. Par suite les estimateurs $\hat{\theta}_t$ sont *consistants :*

$$\hat{\theta}_t\to\theta,\quad P_\theta\text{-p. s.}$$

Enfin, en appliquant le théorème central limite pour les martingales continues, on obtient que $\sqrt{H_t/f(\theta)}\,(\hat{\theta}_t-\theta)$ tend en loi sous P_θ vers $\mathcal{N}(0, 1)$. Cela donne en particulier la vitesse de convergence de $\hat{\theta}_t$ vers la vraie valeur θ, en fonction de t.

(3.6) COMMENTAIRES. — (1) Si $U_t=V_t=1$ identiquement (ou plus généralement si U et V sont déterministes), et si f est constante, toutes les mesures P_θ sont localement équivalentes, et $\bar{Z}^{\zeta/\theta}=Z^{\zeta/\theta}$ est la véritable vraisemblance : on retrouve bien entendu les résultats classiques sur l'estimation d'un drift linéaire quand on observe un mouvement brownien avec drift.

(2) Si $U_t=V_t=1$, mais si f n'est pas constante, on obtient un cas particulier (très simple) des résultats de [8].

Supposons même que f soit bijective de $\mathbb{R}$ dans $]0, \infty[$. En théorie, la connaissance de la trajectoire de X permet de connaître aussi la variation quadratique $\langle X, X\rangle_t = f(\theta)\, t$, donc on peut décider sûrement de la valeur de θ au bout d'un temps d'observation arbitrairement petit (cela correspond au fait que les mesures P_θ sont deux à deux étrangères sur $\mathscr{F}_0 = \mathscr{F}_{0+}$). Dans la pratique une telle procédure est évidemment impossible, car on ne peut observer en général la trajectoire de X qu'en des instants discrets. Cependant, on a alors $Y_t = X_t$ et $H_t = t$, donc $\hat{\theta}_t = X_t/t$ est effectivement observable (aux instants t effectifs d'observation de X). Bien entendu, $\hat{\theta}_t$ n'est pas, pour un t fixé, le meilleur estimateur possible de θ.

(3) Lorsque U et V sont des processus non triviaux, mais suffisamment réguliers en temps, Y_t s'approxime correctement par des sommes discrètes (de Riemann) : on a donc un estimateur qui approche « raisonnablement » $\hat{\theta}_t$ (et, si on ne connaît pas les lois de U et V sous P_θ, on ne dispose d'aucun autre estimateur raisonnable).

(4) Dans les cas (2) et (3) ci-dessus, on ne peut pas utiliser le PVP au sens de [7], car l'ensemble $\{Z_-^{\zeta/\theta} > 0\}$ sur lequel ce dernier est défini est inconnu, ou réduit à $\{0\}$.

(5) On voit immédiatement sur (3.3) que l'égalité $\bar{Z}^{\rho/\theta} = \bar{Z}^{\rho/\zeta}\, \bar{Z}^{\zeta/\theta}$ est fausse en général : le PVP ne satisfait pas la propriété de multiplicativité des dérivées de Radon-Nikodym.

4. DEUX PROPRIÉTÉS D'INVARIANCE DE LA VRAISEMBLANCE PARTIELLE

Il est naturel de se poser la question de voir comment est transformé le PVP quant on remplace le processus de base X par un processus Y de la forme $Y = \rho \,.\, X$, ou de la forme $Y = f(X)$: *a priori*, on perd de l'information et le nouveau PVP devrait être une « partie » de l'ancien : nous étudierons ceci au paragraphe 6.

Ici, nous considérons le cas où la transformation $X \to Y$ est « inversible », c'est-à-dire qu'on a aussi $X = \rho^{-1} . Y$ ou $X = f^{-1}(Y)$. Dans ce cas on devrait trouver le même PVP pour X et Y. C'est ce que nous allons (presque) vérifier ci-dessous : l'invariance du PVP a lieu sur Σ_3 dans le premier cas, sur Σ_5 dans le second.

4-*a*. Transformation par intégrale stochastique

Considérons d'abord un processus prévisible $\rho = (\rho^{ij})_{i \leq q', j \leq q}$ tel que $Y = \rho . X$ soit bien défini (sous P), et h' une fonction de troncation sur

$\mathbb{R}^{q'}$. Soit (B^Y, C^Y, ν^Y) les caractéristiques de Y, et μ^Y la mesure des sauts de Y, et $D^Y = \{\Delta Y \neq 0\}$ [*voir* (2.10)].

Noter que $Y^c = \rho . X^c$, et $\Delta Y = \rho \Delta X$. En particulier $g * \mu^Y = g(\rho x) 1_{\{\rho x \neq 0\}} * \mu$ et cette relation passe aux compensateurs. Enfin rappelons que $X = B + X^c + h * (\mu - \nu) + (x - h(x)) * \mu$, et une relation analogue pour Y. Avec la notation (2.3), on en déduit aisément que P-*p. s.* :

$$
\begin{aligned}
&C^Y = c^Y . F, \qquad \text{avec} \quad c^Y = \rho c \rho^T,\\
&g * \nu^Y = [g(\rho x) 1_{\{\rho x \neq 0\}}] * \nu,\\
&B^Y = \rho . B + [h'(\rho x) - \rho\, h(x)] * \nu.
\end{aligned}
\tag{4.1}
$$

Passons maintenant au problème qui nous intéresse : on se donne un processus prévisible $\rho = (\rho^{ij})_{i, j \leq q}$ à valeurs matricielles *inversibles* et tel qu'il existe un processus Y vérifiant $Y = \rho . X$ *sous* P *et sous* P'. Considérons les caractéristiques (B^Y, C^Y, ν^Y) *définies par* (4.1) (avec $h' = h$), et (B'^Y, C'^Y, ν'^Y) définies de manière analogue à partir de (B', C', ν').

Notons $\bar{Z}^Y$ le PVP de P' par rapport à P, relativement à Y et aux triplets (B^Y, C^Y, ν^Y) et (B'^Y, C'^Y, ν'^Y), et aussi Σ_i^Y les ensembles aléatoires associés par (2.8).

(4.2) Théorème. — *Si* ρ *est inversible, on a* $\Sigma_3^Y = \Sigma_3$ *et les processus* $\bar{Z}$ *et* $\bar{Z}^Y$ *coïncident sur* Σ_3.

On n'a pas en général $\Sigma_2 = \Sigma_2^Y$.Vu le caractère naturel de la propriété $\bar{Z} = \bar{Z}^Y$, il peut donc sembler souhaitable de ne définir le PVP que sur l'ensemble Σ_3 (les exemples du paragraphe 3 restent valides avec cette définition restrictive).

Preuve. — On utilise les notations du paragraphe 2, et on affecte de l'exposant « Y » celles relatives à Y et aux triplets (B^Y, C^Y, ν^Y), (B'^Y, C'^Y, ν'^Y). Comme $\nu(\mathbb{R}_+ \times \{0\}) = 0$ et comme ρ est inversible, on déduit de (4.1) que $a^Y = a$ et $a'^Y = a'$.

Définissons la mesure η^Y par

$$
g * \eta^Y = g(\rho x) 1_{\{\rho x \neq 0\}} * \eta,
\tag{4.3}
$$

de sorte qu'on a (2.5) pour Y avec

$$
y^Y(x) = y(\rho^{-1} x), \qquad y'^Y(x) = y'(\rho^{-1} x)
\tag{4.4}
$$

(notation abrégée pour $y^Y(\omega, t, x) = y(\omega, t, \rho_t^{-1}(\omega) x)$). Dans (2.6), on a alors $G^Y = G$ et $\Sigma_1^Y = \Sigma_1$. En utilisant encore (4.1) et (2.5), puis (2.7), on a

$$
\begin{aligned}
B'^Y - B^Y - (y'^Y - y^Y) h' * \eta^Y &= \rho . (B' - B) - (\rho h)(y' - y) * \eta\\
&= (\rho c \beta) . F + (\rho \tilde{\beta}) . F + \rho . F'.
\end{aligned}
$$

On ne sait pas si $\rho\tilde{\theta}$ est orthogonal à l'image de $\mathbb{R}^q$ par c^Y, sauf bien sûr si $\tilde{\beta}=0$. Par suite, sur l'ensemble Σ_3 on a

$$B'^Y - B^Y - (y'^Y - y^Y)\, h' * \eta^Y = (c^Y \beta^Y).F \tag{4.5}$$

en posant $\beta^Y = \rho^{T,-1}\beta$; de plus $\beta^{Y,T} c^Y \beta^Y = \beta^T c\,\beta$, de sorte qu'on obtient $\Sigma_3 \subset \Sigma_3^Y$. En échangeant les rôles de X et Y, il vient finalement $\Sigma_3 = \Sigma_3^Y$.

Enfin d'après (2.17) il reste à montrer que sur Σ_3, on a $\bar{A}^Y = \bar{A}$, $\Delta\bar{L}^Y = \Delta\bar{L}$ et $\bar{L}^{Y,c} = \bar{L}^c$. La première propriété est évidente; la seconde vient de (4.4) et de ce que $D^Y = D$ et $\Delta Y = \rho\,\Delta X$; enfin $\bar{L}^{Y,c} = \beta^{Y,T}.Y^c = (\beta^{Y,T}\rho).X^c = \beta^T \times X^c = \bar{L}^c$, car $\beta^Y = \rho^{T,-1}\beta$ sur Σ_3. □

4-*b*. Transformation fonctionnelle

Considérons d'abord une fonction $f : \mathbb{R}^q \to \mathbb{R}^{q'}$ de classe C^2, donc $Y = f(X)$ est une P-semi-martingale de caractéristiques notées (B^Y, C^Y, ν^Y), relativement à la fonction de troncation h' sur $\mathbb{R}^{q'}$.

Si $\nabla f = (\partial f^i/\partial x^j)_{i \leq q',\, j \leq q}$, on a $Y^c = \nabla f(X_-).X^c$. On a aussi $\Delta Y = f(X_- + \Delta X) - f(X_-)$. En utilisant la formule d'Ito, on voit facilement qu'une version des caractéristiques est donnée par

$$\tag{4.6} \begin{cases} C^Y = c^Y.F, \qquad \text{avec} \quad c^Y = \nabla f(X_-)\, c\, (\nabla f(X_-))^T, \\ \qquad g * \nu^Y = g(\rho)\, 1_{\{\rho \neq 0\}} * \nu, \\ \text{avec} \\ \qquad \rho(\omega, t, x) = f(X_{t-}(\omega) + x) - f(X_{t-}(\omega)), \\ B^Y = \nabla f(X_-).B + [h'(\rho) - \nabla f(X_-)\, h] * \nu + \delta.F, \\ \text{avec} \\ \qquad \delta_t^i = \dfrac{1}{2} \displaystyle\sum_{j,k \leq q} (\partial^2 f^i/\partial x^j \partial x^k)(X_{t-})\, c_t^{jk}. \end{cases}$$

Passons maintenant au cas « inversible ». On suppose que f est une bijection de $\mathbb{R}^q$ dans lui-même, de classe C^2, et on pose $Y = f(X)$. On considère les caractéristiques (B^Y, C^Y, ν^Y) définies par (4.6) (avec $h' = h$) et (B'^Y, C'^Y, ν'^Y) définies de la même manière à partir de (B', C', ν'). On utilise les mêmes notations $\bar{Z}^Y$, Σ_i^Y qu'au paragraphe 4*a*.

(4.7) Théorème. — *Si f est inversible et de classe C^2, on a $\Sigma_5 = \Sigma_5^Y$, et $\bar{Z}$ et $\bar{Z}^Y$ coïncident sur Σ_5.*

Il semblerait donc naturel de définir le PVP $\bar{Z}$ en restriction à Σ_5 seulement, mais alors cela excluerait les applications du paragraphe 3.

Preuve. — On reprend la preuve de (4.2) avec les modifications suivantes : si ρ est défini par (4.6), on définit η^Y par

$$(4.8)\qquad g * \eta^Y = g(\rho)\, 1_{\{\rho \neq 0\}} * \eta.$$

Comme $x \to \rho(\omega, t, x)$ est inversible, on a (2.5) pour Y avec

$$(4.9)\qquad y^Y = y(\rho^{-1})\ (i.e.\ y^Y(\omega, t, x) = y(\omega, t, \rho^{-1}(\omega, t, x)),\ y'^Y = y'(\rho^{-1})).$$

On a aussi $a^Y = a$, $a'^Y = a'$ (car $\rho(\omega, t, x) = 0 \Leftrightarrow x = 0$), donc $G^Y = G$ et $\Sigma_1^Y = \Sigma_1$. On choisit F de sorte qu'on ait (2.3) et aussi $C' = c'.F$, et on définit δ' comme δ dans (4.6), avec c' au lieu de c. Il vient

$$\begin{aligned} B'^Y - B^Y - (y'^Y - y^Y)\, h' * \eta^Y & \\ &= \nabla f(X_-).(B' - B) - (y' - y)(\nabla f(X_-)\, h) * \eta + (\delta' - \delta).F \\ &= (\nabla f(X_-)\, c\, \beta).F + (\nabla f(X_-)\, \tilde{\beta}).F + \nabla f(X_-).F' + (\delta' - \delta).F. \end{aligned}$$

Sur Σ_5 on a (4.5) avec $\beta^Y = \nabla f(X_-)^{T,-1}\beta$ (noter que $\delta' = \delta$ sur Σ_4, et que ∇f est inversible d'après l'hypothèse faite sur f). On termine alors comme dans la preuve du théorème (4.2). □

5. COMPARAISON AVEC LE PROCESSUS DE VRAISEMBLANCE PARTIELLE DE [10]

Nous allons vérifier que la définition (2.19) étend la version du PVP introduite en [10].

Commençons par rappeler comment on construit le PVP dans [10]. Les hypothèses sont celles du paragraphe 2. On considère aussi une probabilité Q telle que $P \overset{\text{loc}}{\ll} Q$, $P' \overset{\text{loc}}{\ll} Q$, et qui fait de X une semi-martingale, de caractéristiques notées (B'', C'', ν'') : on peut par exemple prendre $Q = (P + P')/2$. On pose $a''_t = \nu''(\{t\} \times \mathbb{R}^q)$. On note z et z' les processus densité de P et P' par rapport à Q, avec les conventions $z_{0-} = z'_{0-} = 1$, et

$$(5.1)\qquad \Gamma = \{z_- > 0\}, \qquad \Gamma' = \{z'_- > 0\}.$$

D'après le théorème de Girsanov [13], il existe des processus prévisibles $\gamma = (\gamma^i)_{i \leq q}$, $\gamma' = (\gamma'^i)_{i \leq q}$ et des fonctions $\tilde{\mathscr{P}}$-mesurables positives $\bar{y}$, $\bar{y}'$ sur $\tilde{\Omega}$, tel qu'on ait Q-p.s. sur $\Gamma \cap \Gamma'$:

$$(5.2)\quad \left\{ \begin{array}{c} \cdot = \bar{y}.\nu'', \quad \nu' = \bar{y}'.\nu'', \qquad a'' = 1 \ \Rightarrow\ a = 1 \quad et \quad a' = 1, \\ C = C' = C'', \\ B = B'' + (c\gamma).F + (\bar{y} - 1)\, h * \nu'', \\ B' = B'' + (c\gamma').F + (\bar{y}' - 1)\, h * \nu'' \end{array} \right.$$

(les relations entre (B, C, ν), resp. (B′, C′, ν′) et (B″, C″, ν″) sont en fait vraie Q-p. s. sur Γ, resp. sur Γ′). D'après [10] on peut définir les Q-martingales locales sur $\Gamma \cap \Gamma'$ suivantes :

$$(5.3)\qquad \begin{aligned} \bar{\mathrm{N}} &= \gamma^{\mathrm{T}}.\mathrm{X}^{c,\,\mathrm{Q}} + \left(\bar{y} - 1 + \frac{a - a''}{1 - a''}\right) * (\mu - \nu''), \\ \bar{\mathrm{N}}' &= \gamma'^{\mathrm{T}}.\mathrm{X}^{c,\,\mathrm{Q}} + \left(\bar{y}' - 1 + \frac{a' - a''}{1 - a''}\right) * (\mu - \nu''); \end{aligned}$$

ensuite, on pose (toujours sur $\Gamma \cap \Gamma'$) :

$$(5.4)\qquad \bar{z} = \mathscr{E}(\bar{\mathrm{N}}), \qquad \bar{z}' = \mathscr{E}(\bar{\mathrm{N}}').$$

Ce sont des Q-martingales locales positives sur $\Gamma \cap \Gamma'$; de plus sur $\Gamma \cap \Gamma'$ on a les implications $z > 0 \Rightarrow \bar{z} > 0$ et $z' > 0 \Rightarrow \bar{z}' > 0$, Q-p. s.

Par ailleurs, notons Z le processus de vraisemblances de P′ par rapport à P : il peut être défini par $\mathrm{Z} = z'/z$ (rappelons que $\inf_t z_t > 0$ P-p. s.); c'est une P-surmartingale positive, et on a $\Gamma' = \{\mathrm{Z}_- > 0\}$ à un ensemble P-évanescent près (toujours avec $\mathrm{Z}_{0-} = 1$).

Selon [10], on appelle *processus de vraisemblance partielle* de P′ par rapport à P, relativement à X, le processus

$$(5.5)\qquad \bar{\mathrm{Z}} = \bar{z}'/\bar{z}, \quad \text{défini P-p. s. sur l'ensemble } \Gamma'.$$

C'est une P-surmartingale positive sur Γ′, et $\mathrm{Z} > 0 \Rightarrow \bar{\mathrm{Z}} > 0$ sur Γ′. De plus, $\bar{\mathrm{Z}}$ ne dépend pas de la mesure dominante Q choisie.

Étant donné (5.2), on peut dans (2.5) remplacer la mesure η par la mesure $\eta' = 1_{\Gamma \cap \Gamma'} . \nu'' + 1_{(\Gamma \cap \Gamma')^c} . \eta$, de sorte qu'on a (2.5) avec

$$(5.6)\qquad \bar{y} = y, \qquad \bar{y}' = y' \quad \text{sur } \Gamma \cap \Gamma'.$$

(5.7) LEMME. — *On a* $\Gamma \cap \Gamma' \subset \Sigma_5$ Q-*p. s., et on peut choisir* β *de sorte que*

$$(5.8)\qquad \beta = \gamma' - \gamma \quad \text{sur } \Gamma \cap \Gamma'.$$

Preuve. — Comme les intégrales stochastiques dans (5.3) sont bien définies sur $\Gamma \cap \Gamma'$, relativement à Q, les processus suivants sont Q-p. s. finis sur $\Gamma \cap \Gamma'$ (utiliser [10]-II.(1.36), (5.6) et aussi le fait que C = C″ et donc $\langle X^{c,i,Q}, X^{c,j,Q}\rangle = c^{ij}.F$ sur $\Gamma\cap\Gamma'$) :

$$G^1=(\gamma^T c\gamma).F,\qquad G^2=(\gamma'^T c\gamma').F,$$

$$G^3=(\sqrt{\bar y}-1)^2 * \nu'' + \sum_{s\leq .}(1-a_s'')[\sqrt{(1-a_s)/(1-a_s'')}-1]^2$$

$$=(\sqrt{y}-1)^2 * \eta + \sum_{s\leq .}(\sqrt{1-a_s}-\sqrt{1-a_s''})^2\, 1_{\{a_s''<1\}},$$

$$G^4=(\sqrt{y'}-1)^2 * \eta)+\sum_{s\leq .}(\sqrt{1-a_s'}-\sqrt{1-a_s''})^2\, 1_{\{a_s''<1\}}.$$

Comme $a''=1\Rightarrow a=1$, $a'=1$ sur $\Gamma\cap\Gamma'$, on voit donc que sur cet ensemble $G\leq 2G^3+2G^4$, et par suite $\Gamma\cap\Gamma'\subset\Sigma_1$.

Il reste à comparer (2.7) et (5.2) : on voit que, Q-p.s. sur $\Gamma\cap\Gamma'$, $\tilde\beta.F=F'=0$ et $(c\beta).F=(c(\gamma'-\gamma)).F$ et $C'=C$. On peut donc prendre $\beta=\gamma'-\gamma$ sur $\Gamma\cap\Gamma'$. Enfin $(\beta^T c\beta).F<\infty$ sur $\Gamma\cap\Gamma'$ découle de ce que G^1 et G^2 sont finis sur cet ensemble, et d'après (2.8) on a finalement $\Gamma\cap\Gamma'\subset\Sigma_5$ Q-p.s. □

(5.9) Théorème. — *Les processus* $\bar Z$ *définis par* (2.19) *et* (5.5) *coïncident* P-*p.s. sur l'ensemble* $\Gamma\cap\Gamma'$.

Preuve. — (*a*) Soit $\bar Z$ le processus défini par (5.5). On a vu que c'est une P-surmartingale positive sur Γ', avec $\bar Z_->0$ sur Γ'. On peut donc définir une nouvelle P-surmartingale sur Γ' par $\bar L=(1/\bar Z_-).\bar Z$, dont on note $\bar L=\bar M-\bar A$ la décomposition de Doob-Meyer. On a $\bar Z=\mathscr{E}(\bar L)$ sur Γ', donc d'après (2.17) il suffit de montrer que sur Γ', on a les formules (2.13) et (2.16) et $\bar L^c=\beta^T.X^c$.

(*b*) Commençons par $\bar A$. D'après la fin de la preuve de (5.8) de [10], ce processus est le Q-compensateur sur $\Gamma\cap\Gamma'$ du processus suivant :

$$\bar A'=\sum_{s\leq .}(1+\Delta\bar N_s')\,1_{\{1+\Delta\bar N_s=0\}}.$$

Par (5.3), il vient sur $\Gamma\cap\Gamma'$:

$$(5.10)\qquad \begin{aligned}\Delta\bar N_t&=((\bar y(t,\Delta X_t)-1)\,1_D(t)-\frac{a_t'-a_t''}{1-a_t''}1_{D^c}(t),\\ \Delta\bar N_t'&=((\bar y'(t,\Delta X_t)-1)\,1_D(t)-\frac{a_t'-a_t''}{1-a_t''}1_{D^c}(t).\end{aligned}$$

Donc sur $\Gamma\cap\Gamma'$:

$$\bar A'=\bar y'\,1_{\{\bar y=0\}} * \mu+\sum_{s\leq .}\frac{1-a_s'}{1-a_s''}1_{\{a_s=1\}}\,1_{D^c}(s).$$

Comme la Q-projection prévisible du processus 1_{D^c} est $1-a''$ [*cf.* [13]-II. (1.17)], et étant donné (5.6), on voit que le Q-compensateur de $\bar{A}'$ sur $\Gamma \cap \Gamma'$ est donné par (2.13).

(*c*) Passons ensuite aux parties martingales continues. D'après la formule d'Ito, on a sur $\Gamma \cap \Gamma'$:

$$\bar{Z}^{c,Q} = (1/\bar{z}_-) . \bar{z}'^{c,Q} - (\bar{z}'_-/\bar{z}_-^2) . \bar{z}^{c,Q} = \bar{Z}_- . (\bar{N}'^{c,Q} - \bar{N}^{c,Q}).$$

D'après le théorème de Girsanov, on a aussi

$$\bar{Z}^{c,P} = \bar{Z}_- . (\bar{N}'^{c,P} - \bar{N}^{c,P}).$$

D'autre part on a $\bar{N}^{c,Q} = \gamma^T . X^{c,Q}$, donc aussi (encore par le théorème de Girsanov) $\bar{N}^{c,P} = \gamma^T . X^{c,P}$ sur $\Gamma \cap \Gamma'$, et de même $\bar{N}'^{c,P} = \gamma'^T . X^{c,P}$. Vu (5.8), il vient $\bar{Z}^{c,P} = \bar{Z}_- \beta^T . X^{c,P}$ sur $\Gamma \cap \Gamma' = \Gamma'$ (P-p. s.). Par suite $\bar{L}^{c,P} = \bar{M}^{c,P} = \beta^T . X^{c,P}$ P-p. s. sur Γ'.

(*d*) On sait que, sur $\Gamma \cap \Gamma'$, $z > 0$ implique $\bar{z} > 0$ Q-p. s.; on a donc $\bar{z} > 0$ et $1 + \Delta\bar{N} > 0$ P-p. s. sur Γ' (car $z > 0$ P-p. s.). D'après (5.5), il vient alors $\Delta\bar{Z} = \bar{Z}_- (\Delta\bar{N}' - \Delta\bar{N})/(1 + \Delta\bar{N})$, donc $\Delta\bar{L} = (\Delta\bar{N}' - \Delta\bar{N})/(1 + \Delta\bar{N})$ P-p. s. sur Γ'. Étant donné (5.10) et (5.6), on obtient aisément (2.16) sur Γ', ce qui achève la démonstration. □

6. COMPARAISON ENTRE VRAISEMBLANCE ET VRAISEMBLANCE PARTIELLE

6-*a*. Dans ce paragraphe, de nature très calculatoire, nous allons voir comment « calculer » (dans un certain sens) le PVP en fonction du véritable processus de vraisemblance Z de P′ par rapport à P, sur l'ensemble $\{Z_- > 0\}$ (en dehors de cet ensemble, la vraisemblance n'apporte pas d'information sur l'évolution des processus).

Lorsque $P' \overset{\text{loc}}{\ll} P$, on obtient essentiellement que $\bar{Z}$ est une P-martingale locale (et pas seulement une surmartingale), qui est une sorte de « projection » de Z sur le sous-espace de martingales engendré (au sens des sous-espaces stables) par le processus X lui-même. C'est un résultat facile; il n'en est malheureusement pas de même du cas général, qui sera utilisé de manière essentielle dans [12].

Commençons par introduire deux notations :

(*a*) Pour tout processus H on note ${}^pH$ sa projection prévisible, si elle existe (donc pour tout temps prévisible T, on a ${}^pH_T = E(H_T \mid \mathscr{F}_{T-})$ sur $\{T < \infty\}$).

(*b*) Si R est une probabilité et ζ une mesure aléatoire sur $\mathbb{R}_+ \times \mathbb{R}^q$, on définit la mesure M_ζ^R sur $(\tilde{\Omega}, \mathscr{F} \otimes \mathscr{R}_+ \otimes \mathscr{R}^q)$ par $M_\zeta^R(W) = E_R(W * \zeta_\infty)$, où E_R désigne l'espérance relativement à R. On note $M_\zeta^R(.\,|\tilde{\mathscr{P}})$ l'espérance conditionnelle par rapport à $\tilde{\mathscr{P}}$ pour cette mesure.

Comme au paragraphe 5, on note Z le processus densité de P′ par rapport à P; l'ensemble Γ'' ci-dessous est P-indistinguable de Γ' :

(6.1) $$\Gamma'' = \{Z_- > 0\}.$$

En outre, on définit la P-surmartingale locale L sur Γ'' par :

(6.2) $L = (1/Z_-).Z$, de décomposition canonique $L = M - A$

(M est une P-martingale locale sur Γ'', A est croissant prévisible).

(6.3) PROPOSITION. — (*a*) *Sur* Γ'' *on a*

(6.4) $$\langle L^c, X^{i,c} \rangle = (c\,\beta)^i . F.$$

(*b*) *Définissons le processus prévisible* K′ *et la fonction* $\tilde{\mathscr{P}}$*-mesurable* U′ *sur* Γ'' *par*

(6.5) $$K' = \{a - a' - {}^p(\Delta L\, 1_{D^c})\}\, 1_{\{a<1\}}$$

(6.6) $$U' = \left\{\frac{y'}{y} - 1 - M_\mu^P(\Delta L\,|\,\tilde{\mathscr{P}})\right\} 1_{\{y>0\}}.$$

On a alors $K' \geqq 0$, $U' \geqq 0$, *et le processus* $A - \bar{A} - U' * \nu - \sum_{s \leqq .} K'_s$ *est croissant sur* Γ''.

Un cas particulier important est celui où $P' \overset{\text{loc}}{\ll} P$: dans ce cas Z est une martingale locale, donc L également et A = O sur Γ'' : ci-dessus on a donc $\bar{A} = 0$, $U' = 0\ M_\mu^P$-p. s. et $K' = 0$ sur Γ''. De plus on sait d'après le théorème de Girsanov que, sur Γ'', $a = 1 \Rightarrow a' = 1$ et que $\nu' = Y.\nu$ avec $Y = (y'/y)\, 1_{\{y>0\}}$, et (avec la notation (2.18)) $\hat{Y} = a'$. Ainsi, (2.15) devient

(6.7) $$\bar{L} = \beta . X^c + \left(Y - 1 + \frac{\hat{Y} - a}{1-a} 1_{\{a<1\}}\right) * (\mu - \nu) \quad \text{sur } \Gamma'',$$

tandis qu'on a (6.4) et

(6.8) $$Y - 1 = M_\mu^P(\Delta L\,|\,\tilde{\mathscr{P}}) \quad \text{sur } \Gamma''.$$

D'après les théorèmes III.(4.11) et III.(4.20) de [13], on $L = \bar{L} + L'$, où L′ est une martingale locale sur Γ'' vérifiant $\langle L^c, X^{i,c} \rangle = 0$ et $M_\mu^P(\Delta L'\,|\,\tilde{\mathscr{P}}) = 0$: on dit que $\bar{L}$ est la *projection* de L sur X. Ainsi :

(6.9) COROLLAIRE. — *Si* $P' \overset{\text{loc}}{\ll} P$, *et sur l'ensemble aléatoire* Γ'', $\bar{L}$ *est la martingale locale projection de* L *sur* X.

Preuve de (6.3). — (i) Reprenons les notations Q, z, z', Γ, Γ' du paragraphe 5, et rappelons qu'on a (5.6) et (5.8). Le théorème de Girsanov (*voir* [13]-III.3.28) permet un calcul « explicite » de $\bar{y}$, $\bar{y}'$, γ, γ' en termes de la probabilité Q (les crochets ci-dessous sont relatifs à Q) :

$$\begin{aligned} &\langle z^{c,\,\mathrm{Q}}, \mathrm{X}^{i,\,c,\,\mathrm{Q}} \rangle = z_- (c\,\gamma)^i . \mathrm{F}, \\ (6.10) \qquad &\langle z'^{c,\,\mathrm{Q}}, \mathrm{X}^{i,\,c,\,\mathrm{Q}} \rangle = z'_- (c\,\gamma')^i . \mathrm{F}, \\ &\mathrm{M}^{\mathrm{Q}}_{\mu} (z \,|\, \tilde{\mathscr{P}}) = z_- \bar{y}, \qquad \mathrm{M}^{\mathrm{Q}}_{\mu} (z' \,|\, \tilde{\mathscr{P}}) = z'_- \bar{y}'. \end{aligned}$$

Étant donné (5.6), on choisira $\eta = \nu''$ et donc $y = \bar{y}$, $y' = \bar{y}'$ sur $\Gamma \cap \Gamma'$.

(ii) Montrons la partie (*a*). Les processus $\mathrm{N} = (1/z_-).z$ et $\mathrm{N}' = (1/z'_-).z'$ sont des Q-martingales locales sur $\Gamma \cap \Gamma'$. Comme en (*c*) de la preuve de (5.9), on a $\mathrm{Z}^{c,\,\mathrm{P}} = \mathrm{Z}_- . (\mathrm{N}'^{c,\,\mathrm{P}} - \mathrm{N}^{c,\,\mathrm{P}})$. Par ailleurs (6.10) implique $\langle \mathrm{N}^{c,\,\mathrm{Q}}, \mathrm{X}^{i,\,c,\,\mathrm{Q}} \rangle = (c\,\gamma)^i . \mathrm{F}$, tandis que $\langle \mathrm{N}^{c,\,\mathrm{Q}}, \mathrm{X}^{i,\,c,\,\mathrm{Q}} \rangle = \langle \mathrm{N}^{c,\,\mathrm{P}}, \mathrm{X}^{i,\,c,\,\mathrm{P}} \rangle$ P-p. s., et de même pour N′, d'après le théorème de Girsanov. Comme $\beta = \gamma' - \gamma$ par (5.8), on en déduit (6.4).

(iii) Comme $\Delta \mathrm{L} = \Delta \mathrm{Z}/\mathrm{Z}_-$ sur $\Gamma \cap \Gamma''$, on a sur cet ensemble :

$$(6.11) \qquad z\,\Delta \mathrm{L} = \frac{z_-}{z'_-} z' - z - \frac{z_-}{z'_-} z' \, 1_{\{z=0\}}.$$

Notons ${}^{p,\,\mathrm{Q}}\mathrm{H}$ la Q-projection prévisible du processus H. La formule de Bayes implique $z_-{}^p\mathrm{H} = {}^{p,\,\mathrm{Q}}(z\,\mathrm{H})$ (car ${}^{p,\,\mathrm{Q}}z = z_-$). D'après (6.10) et [13]-III (4.19) on a $z_-\,a = {}^{p,\,\mathrm{Q}}(z\,1_{\mathrm{D}})$, donc $z_-(1-a) = {}^{p,\,\mathrm{Q}}(z\,1_{\mathrm{D}^c})$, et de même $z'_-(1-a') = {}^{p,\,\mathrm{Q}}(z'\,1_{\mathrm{D}^c})$. Il vient alors sur $\Gamma \cap \Gamma'$, par (6.11) :

$$z_-\,{}^p(\Delta \mathrm{L}\, 1_{\mathrm{D}^c}) = {}^{p,\,\mathrm{Q}}(z\,\Delta\,\mathrm{L} 1_{\mathrm{D}^c}) = z_- \left(a - a' - \frac{1}{z'_-} {}^{p,\,\mathrm{Q}}(z'\,1_{\{z=0\}}\,1_{\mathrm{D}^c}) \right),$$

de sorte que le processus K′ défini par (6.5) vaut sur $\Gamma \cap \Gamma''$:

$$(6.12) \qquad \mathrm{K}' = 1_{\{a<1\}} \frac{1}{z'_-} {}^{p,\,\mathrm{Q}}(z'\,1_{\{z=0\}}\,1_{\mathrm{D}^c}),$$

et en particulier $\mathrm{K}' \geq 0$.

De même, la version de la formule de Bayes pour les mesures $\mathrm{M}^{\mathrm{P}}_{\mu}$ donnée en (7.36) de [9] s'écrit $\mathrm{M}^{\mathrm{Q}}_{\mu}(\mathrm{W}\,z \,|\, \tilde{\mathscr{P}}) = \mathrm{M}^{\mathrm{P}}_{\mu}(\mathrm{W} \,|\, \tilde{\mathscr{P}})\, \mathrm{M}^{\mathrm{Q}}_{\mu}(z \,|\, \tilde{\mathscr{P}})$. D'après (6.10) et (6.11), on a alors sur $\Gamma \cap \Gamma''$:

$$z_-\,y\,\mathrm{M}^{\mathrm{P}}_{\mu}(\Delta \mathrm{L} \,|\, \tilde{\mathscr{P}}) = \mathrm{M}^{\mathrm{Q}}_{\mu}(z\,\Delta \mathrm{L} \,|\, \tilde{\mathscr{P}}) = z_- \left(y' - y - \frac{1}{z'_-} \mathrm{M}^{\mathrm{Q}}_{\mu}(z'\,1_{\{z=0\}} \,|\, \tilde{\mathscr{P}}) \right),$$

de sorte que la fonction U′ définie par (6.6) vaut sur $\Gamma \cap \Gamma'$:

$$(6.13) \qquad \mathrm{U}' = \frac{1}{z'_-} \mathrm{M}^{\mathrm{Q}}_{\mu}(z'\,1_{\{z=0\}} \,|\, \tilde{\mathscr{P}}),$$

et en particulier $U' \geqq 0$.

(iv) Posons $Y = \bar{A} + U' * \nu + \sum_{s \leqq .} K'_s$, qui est un processus croissant prévisible. Il reste à montrer que $A - Y$ est P-p. s. croissant sur Γ''. Comme $\Gamma \cap \Gamma' = \Gamma''$ P-p. s., il suffit de montrer que pour tout processus prévisible H vérifiant $0 \leqq H \leqq 1_{\Gamma \cap \Gamma'}$, on a

$$E_Q(H . Y_\infty) \leqq E_Q(H . A_\infty). \tag{6.14}$$

Soit H comme ci-dessus. Rappelons que $\eta = \nu''$ sur $\Gamma \cap \Gamma'$ et que les mesures M^Q_μ et $M^Q_{\nu''}$ coïncident sur $\tilde{\mathscr{P}}$; d'après (2.13) on a donc :

$$E_Q(H . Y_\infty) = M^Q_\mu[H(y' 1_{\{y=0\}} + y U')] + E_Q\left[\sum_{s>0} H_s((1-a'_s) 1_{\{a_s=1\}} + K_s)\right]. \tag{6.15}$$

En utilisant (6.10) (et $y' = \bar{y}'$ sur $\Gamma \cap \Gamma'$) et (6.13), on voit que

$$\begin{aligned} M^Q_\mu[H(y' 1_{\{y=0\}} + y U')] &= M^Q_\mu\left[\frac{H}{z'_-} z'(1_{\{y=0\}} + 1_{\{y>0,\ z=0\}})\right] \\ &= M^Q_\mu\left(\frac{H}{z'_-} z' 1_{\{z=0\}}\right) = E_Q\left[\sum_{s \in D} H_s \frac{z'_s}{z'_{s-}} 1_{\{z_s=0\}}\right], \end{aligned} \tag{6.16}$$

grâce à la propriété $y=0 \Rightarrow z=0$ M^Q_μ-p. s. sur $\Gamma \cap \Gamma'$ [qui vient de (6.10)].

Par ailleurs, l'ensemble prévisible $J = \{K \neq 0\} \cup \{a=1\}$ est à coupes dénombrables. On sait que pour tout processus optionnel $H' \geqq 0$, on a $E_Q(\sum_{s \in J} H'_s) = E_Q(\sum_{s \in J} {}^{p,\,Q}H'_s)$. On a vu plus haut que $z'_-(1-a') = {}^{p,\,Q}(z 1_{D^c})$ et $z_-(1-a) = {}^{p,\,Q}(z 1_{D^c})$, donc $a=1 \Rightarrow z=0$ sur D^c; en utilisant ces propriétés et (6.12), on voit que le dernier terme de (6.15) vaut

$$E_Q\left[\sum_{s \in J} H_s\left(1_{D^c}(s) 1_{\{a_s=1\}} \frac{z'_s}{z'_{s-}} + \frac{1}{z'_{s-}} 1_{\{a_s<1\}} z'_s 1_{\{z_s=0\}} 1_{D^c}(s)\right)\right] = E_Q\left[\sum_{s \in J \cap D^c} H_s \frac{z'_s}{z'_{s-}} 1_{\{z_s=0\}}\right].$$

par suite d'après (6.15) et (6.16) il vient

$$E_Q(H . Y_\infty) \leqq E_Q(H . A'_\infty),$$

où $A' = \sum_{s \leqq .} (z'_s/z'_{s-}) 1_{\{z_s=0\}} 1_{\Gamma \cap \Gamma'}(s)$ (rappelons que $0 \leqq H \leqq 1_{\Gamma \cap \Gamma'}$).

Mais d'après le lemme (3.15) de [11] on sait que le processus A est le Q-compensateur sur $\Gamma \cap \Gamma'$ du processus A' ci-dessus. On en déduit donc (6.14), et la preuve est achevée. □

6-*b*. Vraisemblances partielles et intégrales stochastiques

On va ici résoudre la question posée au paragraphe 4, pour la transformation $X \to \rho . X$.

Soit $\rho = (\rho^{ij})_{i \leqq q',\, j \leqq q}$ un processus prévisible, et Y un processus tel que $Y = \rho . X$ (P − p. s, et P′ − p. s.). Notons (B^Y, C^Y, ν^Y) les caractéristiques définies par (4.1), et de même (B'^Y, C'^Y, ν'^Y) associées à (B', C', ν'). Nous affectons de l'exposant « Y » toutes les quantités du paragraphe 2, relativement à $Y : \mu^Y, D^Y, a^Y, \Sigma_i^Y, \bar{A}^Y, \bar{L}^Y$, etc. On pose aussi $\tilde{\Omega}' = \Omega \times \mathbb{R}_+ \times \mathbb{R}^{q'}$, $\tilde{\mathscr{P}}' = \mathscr{P} \otimes \mathscr{R}^{q'}$.

Évidemment le théorème (4.2) n'est plus valide, mais la propriété $\Sigma_3 = \Sigma_3^Y$ devient :

(6.17) Proposition. — *On a* $\Sigma_3 \subset \Sigma_3^Y$, *et sur* Σ_3 *on a* $\beta^T c \beta = c^Y \beta^Y$ *et* $\beta^{Y,\,T} c^Y \beta^Y \leqq \beta^T c \beta$.

Preuve. — (*a*) Introduisons les notations :

$$\alpha(U)_t(\omega) = \int_{\mathbb{R}^q} U(\omega, t, x)\, \eta(\omega, \{t] \times dx), \tag{6.18}$$

$$W(\omega, t, x) = 1_{\{\rho_t(\omega)\, x \neq 0\}}, \qquad W' = 1 - W. \tag{6.19}$$

Soit $\mathscr{G}$ la sous-tribu de $\tilde{\mathscr{P}}$ engendrée par les fonctions de la forme $(\omega, t, x) \to U(\omega, t, \rho_t(\omega) x)$ [en abrégé, $U(\rho x)$], où U est $\tilde{\mathscr{P}}'$-mesurable. Toute fonction $\mathscr{G}$-mesurable est évidemment de cette forme, et W est $\mathscr{G}$-mesurable. Si U est $\tilde{\mathscr{P}}$-mesurable, $M_\eta^P(U \mid \mathscr{G})$ se met sous la forme $U'(\rho x)$ avec U' $\tilde{\mathscr{P}}'$-mesurable, et pour tout processus prévisible H :

$$E(H . (U * \eta)_\infty) = M_\eta^P(HU) = M_\eta^P(HU'(\rho x)) = E(H . (U'(\rho x) * \eta)_\infty),$$

donc les processus prévisibles $U * \eta$ et $U'(\rho x) * \eta$ sont égaux. On a ainsi montré la propriété :

$$U'(\rho x) = M_\eta^P(U \mid \mathscr{G}) \quad \Rightarrow \quad U'(\rho x) * \eta = U * \eta. \tag{6.20}$$

(*b*) Définissons η^Y par (4.3) [*i.e.* $g * \eta^Y = g(\rho x)\, W * \eta$]; il existe deux fonctions $\tilde{\mathscr{P}}'$-mesurables positives y^Y, y'^Y telles que

$$y^Y(\rho x) = M_\eta^P(y W \mid \mathscr{G}), \qquad y'^Y(\rho x) = M_\eta^P(y' W \mid \mathscr{G}). \tag{6.21}$$

Vu (4.1), (6.20) et (2.5), il est alors clair que $\nu^Y = y^Y . \eta^Y$ et $v'^Y = y'^Y . \eta^Y$. D'après (2.5) et (6.18), on a $a = \alpha(y)$ et $a' = \alpha(y')$. On a aussi $a_t^Y = \int y^Y(t, x)\, \eta^Y(\{t\} \times dx)$, qui par (4.3) égale $\alpha(y^Y(\rho x) W)_t$, qui par (6.20) vaut encore $\alpha(y W)_t$. Finalement, il vient

$$1 - a^Y = 1 - a + \alpha(y W'), \qquad 1 - a'^Y = 1 - a' + \alpha(y' W'). \tag{6.22}$$

(*c*) Après ces préliminaires, passons à la preuve proprement dite. Grâce à la convexité de $(x, x') \to (\sqrt{x} - \sqrt{x'})^2$, on déduit de (6.21) que

$$(\sqrt{y^Y(\rho x)} - \sqrt{y'^Y(\rho x)})^2 \leqq M_\eta^P((\sqrt{y} - \sqrt{y'})^2\, W \mid \mathscr{G}),$$

donc par (6.20) :

$$(\sqrt{y^{\mathrm{Y}}}-\sqrt{y'^{\mathrm{Y}}})^2 * \eta^{\mathrm{Y}} = (\sqrt{y^{\mathrm{Y}}(\rho x)}-\sqrt{y'^{\mathrm{Y}}(\rho x)})^2\ \mathrm{W} * \eta \leqq (\sqrt{y}-\sqrt{y'})^2\, \mathrm{W} * \eta.$$

On a $(\sqrt{x+z}-\sqrt{x'+z'})^2 \leqq (\sqrt{x}-\sqrt{x'})^2+(\sqrt{z}-\sqrt{z'})^2$ pour tous x, x', z, $z' \geqq 0$. Donc, par la convexité de $(x, x') \to (\sqrt{x}-\sqrt{x'})^2$,

$$\begin{aligned}(\sqrt{1-a^{\mathrm{Y}}}-\sqrt{1-a'^{\mathrm{Y}}})^2 & \\ &\leqq (\sqrt{1-a}-\sqrt{1-a'})^2+(\sqrt{\alpha(y\mathrm{W}')}-\sqrt{\alpha(y'\mathrm{W}')})^2 \\ &\leqq (\sqrt{1-a}-\sqrt{1-a'})^2+\alpha[(\sqrt{y}-\sqrt{y'})^2\,\mathrm{W}'].\end{aligned}$$

On arrive finalement à

$$\mathrm{G}^{\mathrm{Y}} \leqq (\sqrt{y}-\sqrt{y'})^2\, \mathrm{W} * \eta + \sum_{s \leqq .} (\sqrt{1-a_s}-\sqrt{1-a'_s})^2+(\sqrt{y}-\sqrt{y'})^2\, \mathrm{W}' * \eta = \mathrm{G},$$

(rappelons que $\mathrm{W}+\mathrm{W}'=1$), et donc $\Sigma_1 \subset \Sigma_1^{\mathrm{Y}}$.

(*d*) En utilisant (4.1), le même calcul que dans (4.2) montre que sur Σ_1

(6.23) $$\mathrm{B}'^{\mathrm{Y}}-\mathrm{B}^{\mathrm{Y}}-(y'^{\mathrm{Y}}-y^{\mathrm{Y}})\, h' * \eta^{\mathrm{Y}} = (\rho\, c\, \beta) . \mathrm{F}+(\rho\, \tilde{\beta}) . \mathrm{F}+\rho . \mathrm{F}'.$$

Soit $\sqrt{c}$ un processus prévisible matriciel $q \times q$ tel que $c=\sqrt{c}\sqrt{c}^{\mathrm{T}}$. Notons E $(=\mathrm{E}(\omega, t))$ l'image de $\mathbb{R}^{q'}$ dans $\mathbb{R}^q$ par l'application linéaire associée à $\sqrt{c}^{\mathrm{T}} \rho^{\mathrm{T}}$. On a donc $\sqrt{c}^{\mathrm{T}} \beta = \gamma_1+\gamma_2$, avec γ_1 et γ_2 prévisibles, $\gamma_1 \in \mathrm{E}$ et γ_2 orthogonal à E. Donc il existe un processus prévisible $\beta^{\mathrm{Y}}=(\beta^{\mathrm{Y}, i})_{i \leqq q'}$ tel que $\gamma_1=\sqrt{c}^{\mathrm{T}} \rho^{\mathrm{T}} \beta^{\mathrm{Y}}$, et il vient (rappelons que $c^{\mathrm{Y}}=\rho\, c\, \rho^{\mathrm{T}}$) :

(6.24) $$\rho\, c\, \beta=\rho \sqrt{c}(\gamma_1+\gamma_2)=c^{\mathrm{Y}} \beta^{\mathrm{Y}}+\rho \sqrt{c}\, \gamma_2=c^{\mathrm{Y}} \beta^{\mathrm{Y}}$$

car pour tout $x' \in \mathbb{R}^{q'}$ on a $x'^{\mathrm{T}} \rho \sqrt{c}\, \gamma_2=(\sqrt{c}^{\mathrm{T}} \rho^{\mathrm{T}} x')^{\mathrm{T}} \gamma_2=0$. On a aussi

(6.25) $$\beta^{\mathrm{T}} c\, \beta=(\gamma_1+\gamma_2)^{\mathrm{T}}(\gamma_1+\gamma_2)=\gamma_1^{\mathrm{T}} \gamma_1+\gamma_2^{\mathrm{T}} \gamma_2 \geqq \gamma_1^{\mathrm{T}} \gamma_1=\beta^{\mathrm{Y}, \mathrm{T}} c^{\mathrm{Y}} \beta^{\mathrm{Y}},$$

d'où la dernière assertion de la proposition.

Enfin on voit d'après (6.23) et (6.24) que, sur Σ_3, on a (4.5). En utilisant (6.25), on en déduit que $\Sigma_3 \subset \Sigma_3^{\mathrm{Y}}$. □

Nous pouvons maintenant donner l'analogue de (6.3) :

(6.26) PROPOSITION. — (*a*) *Sur* Σ_3 *on a*

(6.27) $$\langle \bar{\mathrm{L}}^c, \mathrm{Y}^{i, c} \rangle = (c^{\mathrm{Y}} \beta^{\mathrm{Y}})^i . \mathrm{F}.$$

(*b*) *Définissons sur* Σ_3 *le processus prévisible* K'^{Y} *et la fonction* $\tilde{\mathscr{P}}'$*-mesurable* U'^{Y} *par*

(6.28) $$\mathrm{K}'^{\mathrm{Y}}=\{a^{\mathrm{Y}}-a'^{\mathrm{Y}}-{}^{\mathrm{P}}(\Delta \bar{\mathrm{L}}\, 1_{(\mathrm{D}^{\mathrm{Y}})^c})\}\, 1_{\{a^{\mathrm{Y}}<1\}}$$

(6.29) $$\mathrm{U}'^{\mathrm{Y}}=\left\{\frac{y'^{\mathrm{Y}}}{y^{\mathrm{Y}}}-1-\mathrm{M}_{\mu^{\mathrm{Y}}}^{\mathrm{P}}(\Delta \bar{\mathrm{L}} \mid \tilde{\mathscr{P}}')\right\} 1_{\{y^{\mathrm{Y}}>0\}}.$$

On a alors $K'^Y \geqq 0$, $U'^Y \geqq 0$, *et le processus* $\bar{A} - \bar{A}^Y - U'^Y * \nu^Y - \sum_{s \leqq .} K_s'^Y$ *est croissant sur* Σ_3.

Là encore, comme pour (6.3) on en déduit que si $\bar{L}$ est une martingale locale sur Σ_3 (donc $\bar{A}=0$ sur cet ensemble), alors $\bar{L}^Y$ est aussi une martingale locale sur Σ_3, et c'est la « projection » de $\bar{L}$ sur Y.

Preuve. — (i) Pour obtenir (6.27) il suffit de remarquer que $Y^{i,\,c}=(\rho . X^c)^i$ et $\bar{L}^c=\beta^T . X^c$, donc $\langle \bar{L}^c, X^{i,\,c} \rangle = (\rho c \beta)^i . F$, et d'appliquer (6.24).

(ii) Comme $D^Y \subset D$, on déduit de (2.16) et de (6.19) que

$$\Delta \bar{L}_t \, 1_{(D^Y)^c}(t) = \left[\frac{y'(t, \Delta X_t)}{y(t, \Delta X_t)} - 1 \right] W'(t, \Delta X_t) \, 1_D(t) + \frac{a_t - a_t'}{1 - a_t} \, 1_{D^c}(t).$$

D'après [13]-III.(4.19), et avec la notation (2.18), on a ${}^p H = \hat{U}$ si U est $\tilde{\mathscr{P}}$-mesurable et $H_t = U(t, \Delta X_t) \, 1_D(t)$. On a aussi ${}^p(1_{D^c}) = 1 - a$, donc par (2.5) et (6.18) on obtient

$$\begin{aligned} {}^p(\Delta \bar{L} \, 1_{(D^Y)^c})_t &= \int \nu(\{t\} \times dx) \left(\frac{y'(t, x)}{y(t, x)} - 1 \right) W'(t, x) + \frac{a_t - a_t'}{1 - a_t} (1 - a_t) \\ &= \alpha (y' \, 1_{\{y>0\}} W' - y W')_t + (a_t - a_t') \, 1_{\{a_t < 1\}}. \end{aligned}$$

Étant donné (6.22), on voit donc que le processus défini par (6.28) vaut

$$K'^Y = \{ \alpha (y' \, 1_{\{y=0\}} W') + (1 - a') \, 1_{\{a=1\}} \} \, 1_{\{a^Y < 1\}} \tag{6.30}$$

et en particulier $K'^Y \geqq 0$.

(iii) Comme $D^Y \subset D$, (2.16) implique $\Delta \bar{L}_t = y'(t, \Delta X_t)/y(t, \Delta X_t) - 1$ si $t \in D^Y$. Soit alors $V(\rho x) = M_\eta^P(y' \, 1_{\{y=0\}} | \mathscr{G})$. D'après le fait que $f * \mu^Y = f(\rho x) * \mu$, puis (2.5), puis (6.21), puis (4.3), et enfin $\nu^Y = y^Y . \eta^Y$, on a pour toute fonction $\tilde{\mathscr{P}}'$-mesurable V' telle que $V'=0$ sur $\{y^Y = 0\}$:

$$\begin{aligned} M_{\mu^Y}^P(\Delta \bar{L} \, V') &= M_\mu^P \left[\left(\frac{y'}{y} - 1 \right) V'(\rho x) \, W \right] \\ &= M_\eta^P [(y' - y - y' \, 1_{\{y=0\}}) \, V'(\rho x) \, W] \\ &= M_\eta^P [(y'^Y(\rho x) - y^Y(\rho x) - V(\rho x)) \, V'(\rho x) \, W] \\ &= M_{\mu^Y}^P [(y'^Y/y^Y - 1 - V/y^Y) \, V'], \end{aligned}$$

de sorte que $M_{\mu^Y}^P(\Delta \bar{L} | \tilde{\mathscr{P}}') = y'^Y/y^Y - 1 - V/y^Y$ sur $\{y^Y > 0\}$. Donc la fonction U'^Y de (6.29) vérifie

$$U'^Y = \frac{1}{y^Y} \, 1_{\{y^Y > 0\}} \, V, \tag{6.31}$$

et en particulier $U'^Y \geqq 0$.

(iv) Si $Y' = \bar{A}^Y + U'^Y * \nu^Y + \sum_{s \leqq .} K_s'^Y$, on voit d'après (2.13), (2.30) et (2.31) (et la définition de V), puis (6.20), (6.21), (6.22), que sur Σ_3 :

$$\begin{aligned} Y' = {} & y'^Y(\rho x)\, 1_{\{y^Y(\rho x)=0\}} W * \eta \\ & + \sum_{s \leqq .} (1-a_s'^Y)\, 1_{\{a_s^Y=1\}} + V(\rho x)\, W\, 1_{\{y^Y(\rho x)>0\}} * \eta \\ & + \sum_{s \leqq .} [(1-a_s')\, 1_{\{a_s=1\}} + \alpha(y'\, 1_{\{y=0\}} W')_s]\, 1_{\{a_s^Y<1\}} \\ = {} & [y'\, 1_{\{y^Y(\rho x)=0\}} + y'\, 1_{\{y^Y(\rho x)>0,\ y=0\}}]\, W * \eta \\ & + \sum_{s \leqq .} \{(1-a_s')\, [1_{\{a_s^Y=1\}} + 1_{\{a_s^Y<1,\ a_s=1\}}] \\ & \qquad + \alpha(y' W')_s\, 1_{\{a_s^Y=1\}} + \alpha(y'\, 1_{\{y=0\}} W')_s\, 1_{\{a_s^Y<1\}}\}. \end{aligned}$$

Mais (6.21) implique $y W = 0$ M_η^P-p. s. sur $\{y^Y(\rho x)=0\}$, et (6.22) implique que si $a^Y=1$ on a $a=1$ et $\alpha(y W')=0$, donc $\alpha(y' W') = \alpha(y'\, 1_{\{y=0\}} W')$. Par suite

$$Y' = y'\, 1_{\{y=0\}} W * \eta + \sum_{s \leqq .} \alpha(y'\, 1_{\{y=0\}} W')_s + \sum_{s \leqq .} (1-a_s')\, 1_{\{a_s=1\}}.$$

Mais d'après (6.18), le processus $y'\, 1_{\{y=0\}} W' * \eta - \sum_{s \leqq .} \alpha(y'\, 1_{\{y=0\}} W')$ est croissant. Étant donné la définition (2.13) de $\bar{A}$, on en déduit que $\bar{A} - Y'$ est croissant, ce qui achève la démonstration. □

6-*c*. Vraisemblance partielle et transformation fonctionnelle

Ici nous étudions la transformation $X \to Y = f(X)$, où f désigne une fonction de classe $C^2 : \mathbb{R}^q \to \mathbb{R}^{q'}$. On utilise exactement les mêmes notations qu'au paragraphe 6-*c*, sauf qu'on remplace (4.1) par (4.6). On peut alors montrer, exactement de la même manière [avec des modifications analogues à celles de la preuve de (4.7) par rapport à (4.2)], que :

(6.32) THÉORÈME. – (*a*) *On a* $\Sigma_5 \subset \Sigma_5^Y$, *et sur* Σ_5 *on a* $\nabla f(X_-) c \beta = c^Y \beta^Y$ *et* $\beta^{Y,T} c^Y \beta^Y \leqq \beta^T c \beta$.

(*b*) *On a les assertions de* (6.26), *à condition de remplacer* Σ_3 *par* Σ_5.

7. PROCESSUS DE HELLINGER PARTIELS, NORMALITÉ ASYMPTOTIQUE

Nous avions introduit les processus de Hellinger partiels pour deux probabilités dans [10], en montrant qu'ils s'exprimaient en fonction des caractéristiques de la semi-martingale de base. Nous allons faire de même ici, assez schématiquement, pour une famille finie de probabilités. Ces processus peuvent être utilisés pour obtenir des théorèmes limite pour les PVP : on verra une extension du théorème principal de [10] à la fin du paragraphe.

On se place sous les hypothèses du paragraphe 2, sauf qu'on remplace les deux triplets $\mathscr{T}$ et $\mathscr{T}'$ de (2.2) par une famille $\mathscr{T}_\theta=(B^\theta, C^\theta, \nu^\theta)$ de triplets indicés par une famille *finie* Θ; pour chaque θ on se donne une probabilité P_θ dans l'ensemble $\mathscr{S}_\theta$ des mesures faisant de X une semi-martingale de caractéristiques $\mathscr{T}_\theta$.

On peut trouver un processus croissant continu adapté F et une mesure prévisible η, tels qu'on ait les factorisations

$$(7.1)\qquad C^\theta=c^\theta . F \quad \text{et} \quad \nu^\theta=y^\theta . \eta, \quad P_\zeta\text{-p. s.}, \quad \forall \theta, \zeta\in\Theta,$$

avec $c^\theta=(c^{\theta,\, ij})_{i,\, j\leqq q}$ prévisibles et y^θ positifs $\tilde{\mathscr{P}}$-mesurables.

Pour tout couple θ, ζ on fait la construction du paragraphe 2, avec $P=P_\theta$, $(B, C, \nu)=(B^\theta, C^\theta, \nu^\theta)$, $(B', C', \nu')=(B^\zeta, C^\zeta, \nu^\zeta)$: on obtient les ensembles $\Sigma_i^{\zeta/\theta}$ [*cf.* (2.6), (2.8)], les processus $\beta^{\zeta/\theta}$, $\tilde\beta^{\zeta/\theta}$, $F'^{\,\zeta/\theta}$ [*cf.* (2.7)], $\bar A^{\zeta/\theta}$, $\bar M^{\zeta/\theta}$, $\bar L^{\zeta/\theta}$ [*cf.* (2.11)], et finalement le PVP $\bar Z^{\zeta/\theta}$ de P_ζ par rapport à P_θ [*cf.* (2.18), il est bien défini sur $\Sigma_2^{\zeta/\theta}$]. Soit enfin

$$(7.2)\qquad \Sigma_i^\theta=\bigcap_{\zeta\in\Theta}\Sigma_i^{\zeta/\theta}$$

(noter que, bien-sûr, $\Sigma_i^{\theta/\theta}=\Omega\times\mathbb{R}_+$, $\beta^{\theta/\theta}=0$, $\bar Z^{\theta/\theta}=1$).

Contrairement à [10], il n'y a pas ici de « mesure de référence ». On va donc rappeler d'abord comment construire les processus de Hellinger « usuels » sans mesure de référence. A cet effet, on note $Z^{\zeta/\theta}$ le processus de vraisemblance de P_ζ par rapport à P_θ. On fixe $\theta\in\Theta$. Pour toute famille $\alpha=(\alpha(\zeta))_{\zeta\in\Theta}$ de nombres positifs avec $\sum_{\zeta\in\Theta}\alpha(\zeta)=1$ et $\alpha(\theta)>0$, on pose $Y(\alpha, \theta)=\prod_{\zeta\in\Theta}(Z^{\zeta/\theta})^{\alpha(\zeta)}$: c'est une P_θ-surmartingale positive. Le *processus de Hellinger* $h(\alpha)$ d'ordre α est alors un processus croissant prévisible tel que

$$(7.3)\qquad Y(\alpha, \theta)+Y(\alpha, \theta)_- . h(\alpha) \quad \text{est une } P_\theta\text{-martingale locale.}$$

(Son existence provient de la décomposition de Doob-Meyer de $Y(\alpha, \theta)$; $h(\alpha)$ est défini de manière P_θ-p. s. unique sur $\bigcap_{\zeta\in\Theta}\{Z_-^{\zeta/\theta}>0\}$). Comme la notation le suggère, $h(\alpha)$ ne dépend pas du point θ choisi.

De même ici, le processus

$$(7.4)\qquad \bar{Y}(\alpha,\theta)=\prod_{\zeta\in\Theta}(\bar{Z}^{\zeta/\theta})^{\alpha(\zeta)},\quad \text{défini sur } \Sigma_2^\theta,$$

est une P_θ-surmartingale positive sur Σ_2^θ. Exactement comme en (7.3), on peut trouver un processus croissant prévisible $\bar{h}(\alpha,\theta)$ tel que

$$(7.5)\qquad \bar{Y}(\alpha,\theta)+\bar{Y}(\alpha,\theta)_{-}.\bar{h}(\alpha,\theta)\quad \text{est une } P_\theta\text{-martingale locale sur } \Sigma_2^\theta.$$

Ce processus $\bar{h}(\alpha,\theta)$ est défini P_θ-p. s. de manière unique sur l'ensemble $\Sigma_2^\theta\cap(\bigcap_{\zeta\in\Theta}\{\bar{Z}_{-}^{\zeta/\theta}>0\})$.

On appellera *processus de Hellinger partiel d'ordre* α, associé à θ, tout processus prévisible croissant $\bar{h}(\alpha,\theta)$ vérifiant (7.5). On va voir qu'on peut en calculer une version en fonction des $\mathscr{T}_\zeta$. Introduisons la fonction suivante $\varphi_\alpha:(\mathbb{R}_+)^\Theta\to\mathbb{R}_+$:

$$(7.6)\qquad \varphi_\alpha(x)=\sum_{\zeta\in\Theta}\alpha(\zeta)\,x^\zeta-\prod_{\zeta\in\Theta}(x^\zeta)^{\alpha(\zeta)},\qquad \text{où}\quad x=(x^\zeta)_{\zeta\in\Theta}.$$

(7.7) THÉORÈME. — *Si* $\alpha(\theta)>0$, *on a la version suivante sur* Σ_2^θ :

$$(7.8)\qquad \bar{h}(\alpha,\theta)=\frac{1}{2}\Big[\sum_{\zeta}\alpha(\zeta)\,(\beta^{\zeta/\theta})^T c^\theta \beta^{\zeta/\theta}-\sum_{\zeta,\,\rho}\alpha(\zeta)\,\alpha(\rho)\,(\beta^{\zeta/\theta})^T c^\theta\beta^{\rho/\theta}\Big].F+\varphi_\alpha((y^\zeta)_{\zeta\in\Theta})\star\eta+\sum_{s\leq .}\varphi_\alpha((1-a_s^\zeta)_{\zeta\in\Theta}).$$

Preuve. — Pour simplifier on écrit $c=c^\theta$, $y=y^\theta$, $\nu=\nu^\theta$, $a=a^\theta$. Soit $\mu^{\bar{L}}$ la mesure des sauts du processus $\bar{L}=(\bar{L}^{\zeta/\theta})_{\zeta\in\Theta}$ (défini sur Σ_2^θ), et $\nu^{\bar{L}}$ son P_θ-compensateur.

Nous sommes ici exactement dans la situation du paragraphe 3 de [11], avec Y et $A^{Y,\,\zeta}$ remplacés par $\bar{L}$ et $\bar{A}^{\zeta/\theta}$. Les calculs qui permettent d'arriver à la formule (3.20) de [11] conduisent ici à la version suivante de $\bar{h}(\alpha,\theta)$ (sur Σ_2^θ bien entendu, avec des crochets pour P_θ) :

$$(7.9)\qquad \bar{h}(\alpha,\theta)=\frac{1}{2}\Big[\sum_{\zeta\in\Theta}\alpha(\zeta)\langle\bar{L}^{\zeta/\theta,\,c},\bar{L}^{\zeta/\theta,\,c}\rangle-\sum_{\zeta,\,\rho\in\Theta}\alpha(\zeta)\,\alpha(\rho)\langle\bar{L}^{\zeta/\theta,\,c},\bar{L}^{\rho/\theta,\,c}\rangle\Big]+\sum_{\zeta\in\Theta}\alpha(\zeta)\,\bar{A}^{\alpha/\theta}+\varphi_\alpha((1+x^\zeta)_{\zeta\in\Theta})\star\nu^{\bar{L}}.$$

D'abord, (2.14) montre que $\langle\bar{L}^{\zeta/\theta,\,c},\bar{L}^{\rho/\theta,\,c}\rangle=((\beta^{\zeta/\theta})^T c\,\beta^{\rho/\theta}).F$, de sorte que les premiers termes (entre crochets) des seconds membres de (7.8) et (7.9) coïncident.

Ensuite, (2.16) montre que pour toute fonction f sur $\mathbb{R}^\Theta$,

$$f * \mu^{\bar{L}} = f\left(\left(\frac{y^\zeta}{y} - 1\right)_{\zeta \in \Theta}\right) * \mu + \sum_{s \leqq .} f\left(\left(\frac{a_s - a_s^\zeta}{1 - a_s}\right)_{\zeta \in \Theta}\right) 1_{D^c}(s)$$

sur Σ_2^θ. En prenant les P_θ-compensateurs on arrive à

$$f * \nu^{\bar{L}} = f\left(\left(\frac{y^\zeta}{y} - 1\right)_{\zeta \in \Theta}\right) * \nu + \sum_{s \leqq .} (1 - a_s) f\left(\left(\frac{a_s - a_s^\zeta}{1 - a_s}\right)_{\zeta \in \Theta}\right).$$

Comme $\nu = y . \eta$, en utilisant (2.13) on voit que la somme des deux derniers termes de (7.9) égale (sur Σ_2^θ) :

$$\sum_{z \in \Theta} \alpha(\zeta)\, y^\zeta\, 1_{\{y=0\}} * \eta$$
$$+ \sum_{s \leqq .} \Big[\sum_{\zeta \in \Theta} \alpha(\zeta)\,(1 - a_s^\zeta)\Big]\, 1_{\{a_s = 1\}} + \varphi_\alpha((y^\zeta/y)_{\zeta \in \Theta})\, y * \eta$$
$$+ \sum_{s \leqq .} (1 - a_s)\, \varphi_\alpha\left(\left(\frac{1 - a_s^\zeta}{1 - a_s}\right)_{\zeta \in \Theta}\right).$$

Comme $u\, \varphi_\alpha(x/u) = \varphi_\alpha(x)$ si $u \neq 0$, on vérifie immédiatement que cette expression égale la somme des deux derniers termes de (7.8). □

En particulier, si $\Theta = \{\theta, \zeta\}$ et si $\gamma \in]0, 1[$, on écrira

(7.10) $\qquad \bar{h}^{\zeta/\theta}(\gamma) = \bar{h}(\alpha, \theta), \qquad$ avec $\quad \alpha(\theta) = \gamma, \quad \alpha(\zeta) = 1 - \gamma.$

Il faut bien noter que $\bar{h}(\alpha, \theta)$ *dépend en général du point de référence* θ choisi, ce qui en limite l'intérêt. Cependant, on a la

(7.11) PROPOSITION. – (*a*) *Les ensembles* Σ_1^θ, Σ_4^θ, Σ_5^θ *ne dépendent pas du point* θ *choisi dans* Θ.

(*b*) *Si* $\alpha(\theta) > 0$ *et* $\alpha(\theta') > 0$, *les versions* (7.8) *des processus de Hellinger partiels vérifient* $\bar{h}(\alpha, \theta) = \bar{h}(\alpha, \theta')$ *sur* $\Sigma_4^\theta = \Sigma_4^{\theta'}$.

Preuve. – (*a*) Soit $G^{\theta\zeta}$ le processus défini par (2.6), avec y, y', a, a' remplacés par y^θ, y^ζ, a^θ, a^ζ. On a $G^{\theta\zeta} = G^{\zeta\theta}$ et $G^{\theta'\zeta} \leqq 2\, G^{\theta\theta'} + 2\, G^{\theta\zeta}$, donc $\Sigma_1^{\theta'/\theta} \cap \Sigma_1^{\zeta/\theta} \subset \Sigma_1^{\zeta/\theta'}$, donc $\Sigma_1^\theta \subset \Sigma_1^{\theta'}$. On a aussi l'inclusion inverse, donc $\Sigma_1^\theta = \Sigma_1^{\theta'}$.

Soit $\tau = \inf(t : C_t^\theta \neq C_t^\zeta$ pour un couple $\theta, \zeta)$. Sur $\Sigma_1^\theta \cap [0, \tau]$ on a $c^\theta = c^{\theta'}$, donc on vérifie aisément sur (2.7) qu'on peut prendre $\beta^{\zeta/\theta'} = \beta^{\zeta/\theta} - \beta^{\theta'/\theta}$, $\tilde{\beta}^{\zeta/\theta'} = \tilde{\beta}^{\zeta/\theta} - \tilde{\beta}^{\theta'/\theta}$, $\quad F'^{\zeta/\theta'} = F'^{\zeta/\theta} - F'^{\theta'/\theta}$. Étant donnés (2.8) et $\Sigma_5^\theta \subset \Sigma_4^\theta \subset [\![0, \tau]\!]$, on en déduit que $\Sigma_4^\theta \subset \Sigma_4^{\theta'}$ et $\Sigma_5^\theta \subset \Sigma_5^{\theta'}$. On a aussi les inclusions inverses, d'où le résultat.

(*b*) Comme sur Σ_3^θ on a $\beta^{\zeta/\theta'} = \beta^{\zeta/\theta} - \beta^{\theta'/\theta}$, le résultat découle d'un calcul facile (basé sur $\Sigma\, \alpha(\zeta) = 1$). □

A titre d'application, nous allons reprendre le théorème (6.3) de [10] dans notre cadre. Pour chaque entier n on se donne un espace $(\Omega^n, \mathscr{F}^n, (\mathscr{F}^n_t)_{t \geq 0})$ muni d'un processus X^n de dimension q_n, et de deux probabilités P^n, P'^n faisant de X^n une semi-martingale. On note $\bar{Z}^n$ le PVP de P'^n par rapport à P^n, relativement à X^n, au sens de (2.19) et Σ^n_i les ensembles associés par (2.8).

Avec les notations de ce paragraphe, on a $\Theta = \{1, 2\}$, avec $P^n_1 = P^n$ et $P^n_2 = P'^n$. On note $\bar{h}^n(\gamma) = \bar{h}^{n,\, 2/1}(\gamma)$ les processus de Hellinger partiels associés par (7.10), pour $\gamma \in]0, 1[$.

Par ailleurs on se donne une martingale gaussienne continue M avec $M_0 = 0$ sur un espace $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \geq 0}, P)$, et on pose $C_t = \langle M, M \rangle_t = E(M_t^2)$.

On peut alors démontrer, exactement comme en [10] (en prenant $v^n = 1$ et $v'^n = \bar{Z}^n$ dans la preuve de [10], ce qui simplifie un peu), le

(7.12) THÉORÈME. — *Soit* R *une partie de* $\mathbb{R}_+$. *Supposons que*

$$1_{\Sigma^n_2}(t) \xrightarrow{P^n} 1 \quad \text{pour tout } t \in R, \tag{7.13}$$

$$\bar{h}^n(\gamma)_t \xrightarrow{P^n} \frac{\gamma(1-\gamma)}{2} C_t \quad \text{pour tout } t \in R, \text{ et pour } \gamma = 1/2,\ \gamma = a \text{ et} \tag{7.14}$$

$\gamma = 1 - a$, où a est un point de $\left]0, \frac{1}{2}\right[$.

Alors :

(*a*) *Les processus* $\bar{Z}^n$ *convergent en loi sous* P^n, *fini-dimensionnellement le long de* R, *vers le processus* $Z = \exp(M - C/2)$.

(*b*) *Si de plus* R *est dense dans* $\mathbb{R}_+$ *les processus* $\bar{Z}^n$ *convergent fonctionnellement en loi, sous* P^n, *vers le processus* Z.

Par rapport à [10], on a remplacé les conditions : « $P^n \overset{\text{loc}}{\ll} P'^n$ pour tout n », ou : « la suite (P^n) est localement contiguë par rapport à la suite (P'^n) », par (7.13). C'est beaucoup plus satisfaisant (*cf.* la remarque (6.4) de [10]), car cette condition s'exprime, comme (7.14), en terme des caractéristiques des X^n.

RÉFÉRENCES

[1] P. ANDERSEN et R. GILL, Cox's regression model for counting processes: a large sample study, *Ann. Stat.* vol. **10**, 1982, p. 1100-1120.

[2] E. ARJAS et P. HAARA, A marked point process approach to censored failure data with complicated covariates, *Scand. J. Stat.*, vol. **11**, 1984, p. 193-210.
[3] D. R. COX, Regression models and life tables, *J. R. Stat. Soc.*, série B, vol. **34**, 1972, p. 187-202.
[4] D. R. COX, Partial likelihood, *Biometrika*, vol. **62**, 1975, p. 69-76.
[5] C. DELLACHERIE et P. A. MEYER, *Probabilités et potentiel*, I, 1976, II, 1982, Hermann, Paris.
[6] R. GILL, *Note on product integration, likelihood and partial likelihood for counting processes*, Techn. Report, CWI, Amsterdam, 1985.
[7] V. P. GODAMBE et C. C. HEYDE, Quasi-likelihood and optimal estimation, *Intern. Stat. Rev.*, vol. **55**, 1987, p. 231-244.
[8] J. E. HUTTON et P. I. NELSON, Quasi-likelihood estimation for semimartingales, *Stoch. Proc. Appl.*, vol. **22**, 1986, p. 245-257.
[9] J. JACOD, Calcul stochastique et problèmes de martingales, *Lect. Notes Math.*, vol. **714**, 1979, Springer-Verlag, Berlin.
[10] J. JACOD, Partial likelihood process and asymptotic normality, *Stoch. Proc. Appl.*, vol. **26**, 1987, p. 47-71.
[11] J. JACOD, Filtered statistical models and Hellinger processes, *Stoch. Proc. Appl.*, vol. **32**, 1989, p. 3-45.
[12] J. JACOD, Regularity, partial regularity, partial information process, for a filtered statistical model, *Probab., Theor. Rel. Fields*, 1990 (à paraître).
[13] J. JACOD et A. N. SHIRYAEV, *Limit theorems for stochastic processes*, Springer-Verlag, Berlin, 1987.
[14] E. SLUD, *Martingale methods in Statistics*, 1986 (à paraître).
[15] E. SLUD, *Partial likelihood for continuous-time stochastic processes*, 1989, preprint.
[16] M. SORENSEN, *On quasi likelihood for semimartingales*, Research report, n° 171, Univ. of Aarhu, 1988.
[17] W. WONG, Theory of partial likelihoods, *Ann. Stat.*, vol. **14**, 1986, p. 88-123.

(*Manuscrit reçu le* 30 *mai* 1989.)