François Coquet
Jean Jacod

Convergence des surmartingales. Application aux vraisemblances partielles

Séminaire de probabilités (Strasbourg), tome 24 (1990), p. 282-299
<http://www.numdam.org/item?id=SPS_1990__24__282_0>

CONVERGENCE DES SURMARTINGALES - APPLICATION AUX VRAISEMBLANCES PARTIELLES

François COQUET et Jean JACOD

1 - INTRODUCTION

Cet article est motivé par des questions de statistique: nous nous proposons de montrer des résultats de convergence pour les processus de vraisemblance partielle, analogues à ceux obtenus dans [8] pour les processus de vraisemblance usuels. Cela permet par exemple de redémontrer de manière simple divers résultats de normalité asymptotique: cf. Andersen et Gill [1], Dzhaparidze [2], Hutton et Nelson [5], Sørensen [13], et à titre d'application nous retrouverons ceux de Greenwood et Wefelmeyer [4]. Tout ceci est exposé au paragraphe 4.

Ces résultats de convergence s'obtiennent comme corollaires de résultats généraux de convergence pour les suites de surmartingales positives. Ceux-ci s'expriment de manière élégante en termes des "processus de Hellinger" qu'on peut associer à toute famille finie de surmartingales positives (paragraphe 3).

Enfin, un outil essentiel pour ceci est un théorème de convergence analogue à la condition nécessaire et suffisante de convergence des lois indéfiniment divisibles, exprimée en terme de convergence des drifts, mesures de Lévy et covariances des parties gaussiennes. Nous consacrons donc le paragraphe 2 à l'étude de la convergence dans une formule de type "Lévy-Khintchine", mais associée aux transformées de Mellin: cela nous semble intéressant en soi.

Nous utilisons les notations usuelles de théorie générale: $H \cdot X$ désigne l'intégrale stochastique du processus prévisible H par rapport à la semimartingale X, $W * \mu$ est le processus intégrale de W par rapport à la mesure aléatoire μ: cf. [6].

2 - UNE FORMULE DE LEVY-KHINTCHINE: PROPRIETES DE CONVERGENCE

La formule de Lévy-Khintchine donne la transformée de Fourier des lois indéfiniment divisibles. Ici, nous étudions une formule analogue, liée aux transformées de Mellin. Cette formule a été introduite dans [7], et nous rappelons d'abord de quoi il s'agit (avec des notations plus naturelles, sans rapport avec les modèles statistiques).

On fixe un entier d, et on note $\mathfrak{M}_d$ l'ensemble des matrices $d\times d$ symétriques semi-définies positives.

On désigne par $\mathcal{Q}$ l'ensemble des triplets (a,c,F) constitués de $a=(a^i)_{i\leq d}\in\mathbb{R}^d$, de $c=(c^{ij})_{i,j\leq d}\in\mathfrak{M}_d$, et d'une mesure positive F sur $E=[-1,\infty[^d$ avec $F(\{0\})=0$ et $\int(|x|\wedge|x|^2)F(dx)<\infty$. Soit $\mathcal{Q}_+$ l'ensemble des $(a,c,F)\in\mathcal{Q}$ avec $a^i\geq 0$ pour tout i.

Soit $\mathcal{B}$ l'ensemble des $\gamma=(\gamma^i)_{i\leq d}\in\mathbb{R}^d_+$ avec $\sum\gamma^i<1$; sa fermeture $\overline{\mathcal{B}}$ est l'ensemble des $\gamma\in\mathbb{R}^d_+$ avec $\sum\gamma^i\leq 1$. Si $\gamma\in\overline{\mathcal{B}}$ on considère la fonction ψ_γ: $E\to\mathbb{R}_+$ définie par

2.1 $$\psi_\gamma(x) = \sum_i \gamma^i x^i + 1 - \prod_i (1+x^i)^{\gamma^i}.$$

A tout triplet $(a,c,F)\in\mathcal{Q}$ on associe la fonction continue $g_{a,c,F}$: $\overline{\mathcal{B}}\to\mathbb{R}$ par

2.2 $$g_{a,c,F}(\gamma) = \sum_i \gamma^i(a^i + c^{ii}/2) - \frac{1}{2}\sum_{i,j}\gamma^i\gamma^j c^{ij} + F(\psi_\gamma),$$

et on a $g_{a,c,F}\geq 0$ si $(a,c,F)\in\mathcal{Q}_+$. Le lemme 7.9(a) de [7] donne:

2.3 LEMME. <u>La restriction de</u> $g_{a,c,F}$ <u>à une partie dense de</u> $\overline{\mathcal{B}}$ <u>détermine le triplet</u> (a,c,F).

Nous allons à présent étudier la convergence des fonctions $g_{a,c,F}$ pour des triplets appartenant à $\mathcal{Q}_+$. A cet effet, on considère une fonction de troncation <u>continue</u> $h=(h^i)_{i\leq d}$: $\mathbb{R}^d\to\mathbb{R}^d$ fixée une fois pour toutes: c'est une fonction bornée, telle que $h(x)=x$ pour x dans un voisinage de 0, et $h^i(x)=0$ si $|x^i|$ est assez grand. A tout triplet $(a,c,F)\in\mathcal{Q}_+$ on associe $b\in\mathbb{R}^d$, $\tilde{c}\in\mathfrak{M}_d$ par

2.4 $$b^i = a^i + \int F(dx)(x^i-h^i(x)),\quad \tilde{c}^{ij} = c^{ij} + \int F(dx)h^i(x)h^j(x).$$

Introduisons également deux classes de fonctions sur $E=[-1,\infty[^d$:

- $\mathcal{C}_1$ = ensemble des fonctions continues bornées nulles autour de 0;
- $\mathcal{C}_2$ = ensemble des fonctions continues, telles que $f(x)=o(|x|^2)$ si $x\to 0$ et $f(x)=o(|x|)$ si $|x|\to\infty$.

On considère enfin des triplets (a_n,c_n,F_n) et (a,c,F) dans $\mathcal{Q}_+$ et on introduit les conditions suivantes (b_n et $\tilde{c}_n$ sont associés à (a_n,c_n,F_n) par 2.4):

[B] $b_n\to b$;

[C] $\tilde{c}_n\to\tilde{c}$;

$[D_i]$ $F_n(f) \to F(f)$ pour toute $f \in \mathcal{C}_i$ (où $i=1,2$).

Si h' est une autre fonction de troncation, les fonctions $h'^i - h^i$ et $h'^i h'^j - h^i h^j$ sont dans $\mathcal{C}_1$, de sorte que sous $[D_1]$ les conditions [B] et [C] ne dépendent pas de la fonction de troncation choisie.

2.5 THEOREME. Soit $(a_n, c_n, F_n) \in \mathcal{Q}^+$ et $g_n = g_{a_n, c_n, F_n}$. Soit $\mathcal{B}'$ une partie dense de $\mathcal{B}$.

a) Si la suite g_n converge simplement sur $\mathcal{B}'$ vers une limite g, il existe un triplet $(a,c,F) \in \mathcal{Q}_+$ avec $g = g_{a,c,F}$ sur $\mathcal{B}'$.

b) Soit $(a,c,F) \in \mathcal{Q}_+$ et $g = g_{a,c,F}$. Il y a équivalence entre

(i) $g_n \to g$ simplement sur $\mathcal{B}'$,

(ii) $g_n \to g$ simplement sur $\mathcal{B}$,

(iii) On a [B], [C], $[D_1]$,

(iv) On a [B], [C], $[D_2]$.

Dans ce cas, on a aussi

$$2.6 \qquad \sup_n (|a_n| + |c_n| + \int F_n(dx)(|x| \wedge |x|^2) < \infty.$$

Preuve. Etant donné que sous $[D_1]$ les conditions [B] et [C] ne dépendent pas de la fonction de troncation, nous prenons pour h la fonction $h^i(x) = \ell(x^i)$, où $\ell(y) = y$ (resp. $2-y$, resp. 0) pour $-1 \le y \le 1$ (resp. $1 \le y \le 2$, resp. $y \ge 2$). Nous procédons en plusieurs étapes.

Etape 1. L'implication (ii)$\Rightarrow$(i) est évidente, tandis que (iv)$\Rightarrow$(ii) découle de ce que pour chaque $\gamma \in \mathcal{B}$ la fonction

$$\hat{\psi}_\gamma(x) = \psi_\gamma(x) + \sum_i \gamma^i [h^i(x) - x^i - (h^i(x))^2/2] + \frac{1}{2} \sum_{i,j} \gamma^i \gamma^j h^i(x) h^j(x)$$

appartient à $\mathcal{C}_2$, et de ce que par un calcul simple on a

$$2.7 \qquad g_{a,c,F}(\gamma) = \sum_i \gamma^i (b^i + \tilde{c}^{ii}/2) - \frac{1}{2} \sum_{i,j} \gamma^i \gamma^j \tilde{c}^{ij} + F(\hat{\psi}_\gamma).$$

Etape 2. Montrons que la convergence simple de g_n vers une limite g, en restriction à $\mathcal{B}'$, entraine 2.6. Fixons $\gamma \in \mathcal{B}'$ avec $\gamma^i \ge 1/2d$ pour tout i. Comme $a_n^i \ge 0$ on a $|a_n| \le 2d^2 \sum \gamma^i a_n^i$ ($|.|$ désigne la norme euclidienne); on a donc $|x|^2 \wedge |x| \le K\psi_\gamma(x)$ pour une constante k; enfin d'après [7,7.19] il existe une constante K' telle que $|c_n| \le K'(\sum \gamma^i c_n^{ii} - \sum \gamma^i \gamma^j c_n^{ij})$. Par suite 2.2 implique

$$|a_n| + |c_n| + \int F_n(dx)(|x| \wedge |x|^2) \le (4d^2 + K + K')g_n(\gamma),$$

et 2.6 suit de ce que $g_n(\gamma) \to g(\gamma)$.

Etape 3. Montrons que (iii) $\Rightarrow$ 2.6. En se rappelant la forme de h, et si $g\in\mathcal{C}_1$ vérifie $1_{\{x^i\geq 1\}} \leq g(x)$, il vient

$$b_n^i \geq a_n^i + \int_{x^i\geq 2} x^i F_n(dx), \qquad \tilde{c}_n^{ii} \geq c_n^{ii} + \int_{|x^i|\leq 1} (x^i)^2 F_n(dx),$$

$$F_n(g) \geq \int_{x^i\geq 1} F_n(dx).$$

Comme $a_n^i\geq 0$ et $c_n^{ii}\geq 0$, on déduit alors 2.6 de [B], [C], [D_1].

Etape 4. Montrons que (iii) $\Rightarrow$ (iv). Soit k_α: $\mathbb{R}^d \to [0,1]$ continue avec $1_{\{\alpha\leq |x|\leq 1/\alpha\}} \leq k_\alpha(x) \leq 1_{\{\alpha/2\leq |x|\leq 2/\alpha\}}$. Si $f\in\mathcal{C}_2$ on pose $\rho(\alpha) = \sup_{x:\ |x|\leq\alpha \text{ ou } |x|\geq 1/\alpha} |f(x)|/(|x|\wedge|x|^2)$. On a

$$|F_n(f) - F_n(fk_\alpha)| \leq \rho(\alpha) \int F_n(dx)(|x|^2\wedge|x|),$$

et de même pour F. On a $fk_\alpha\in\mathcal{C}_1$, donc $F_n(fk_\alpha) \to F(fk_\alpha)$ par [D_1]. On a 2.6 par l'étape 3, et $\lim_{\alpha\downarrow 0} \rho(\alpha) = 0$, donc $F_n(f) \to F(f)$: par suite on a [D_2], d'où (iv).

Etape 5. Supposons que $g_n \to g$ simplement sur $\mathcal{B}'$. Nous allons montrer qu'il existe un triplet $(a,c,F)\in\mathcal{Q}_+$ tel que $g = g_{a,c,F}$ sur $\mathcal{B}'$ et qu'on ait (iii), ce qui achèvera de prouver le théorème. Noter qu'on peut raisonner sur des sous-suites.

On a 2.6 par l'étape 2, donc la suite des mesures (F_n), restreintes à $E\backslash\{0\}$ est vaguement relativement compacte. Quitte à prendre une sous-suite, on peut donc supposer qu'elle converge vaguement vers une mesure positive F sur $E\backslash\{0]$, qu'on prolonge à E en posant $F(\{0\})=0$. De plus, pour tout $\alpha\in]0,1[$,

$$\int_{\alpha<|x|<1/\alpha} F(dx)(|x|^2\wedge|x|) \leq \limsup_n \int_{\alpha<|x|<1/\alpha} F_n(dx)(|x|^2\wedge|x|),$$

d'où $\int F(dx)(|x|\wedge|x|^2)<\infty$.

En refaisant le raisonnement de l'étape 4, on déduit de 2.6 et de ce qui précède qu'on a [D_2] (car fk_α est à support compact dans $E\backslash\{0\}$ si $f\in\mathcal{C}_2$), donc a fortiori [D_1].

Comme $\hat{\psi}_\gamma\in\mathcal{C}_2$, il vient alors $F_n(\hat{\psi}_\gamma) \to F(\hat{\psi}_\gamma)$. Donc par 2.7,

2.8 $$\sum_i \gamma^i(b_n^i + \tilde{c}_n^{ii}/2) - \frac{1}{2}\sum_{i,j} \gamma^i\gamma^j\tilde{c}_n^{ij} \longrightarrow g(\gamma) - F(\hat{\psi}_\gamma)$$

pour tout $\gamma\in\mathcal{B}'$. $\mathcal{B}'$ étant dense dans $\mathcal{B}$, on en déduit que b_n^i et $\tilde{c}_n^{ij}$ convergent vers des limites respectives b^i et $\tilde{c}^{ij}$. On a donc [B] et [C], à condition de définir a, c par les formules 2.4.

On a aussi clairement $g = g_{a,c,F}$ sur $\mathcal{B}'$ (comparer 2.8 et 2.7). Il nous reste donc à montrer que $c\in\mathcal{M}_d$ et que $a^i\geq 0$.

Si k_α est comme dans l'étape 4, et si $k'_\alpha=1-k_\alpha$, on pose $c_n^{ij}(\alpha) = c_n^{ij} + F_n(h^ih^jk'_\alpha)$, et de même pour $c^{ij}(\alpha)$. Par $[D_1]$, $F_n(h^ih^jk_\alpha) \to F(h^ih^jk_\alpha)$, donc [C] implique $c_n^{ij}(\alpha) \to c^{ij}(\alpha)$. Comme $c_n(\alpha)\in\mathfrak{M}_d$ on a aussi $c(\alpha)\in\mathfrak{M}_d$, donc $c\in\mathfrak{M}_d$ car $\lim_{\alpha\downarrow 0} c^{ij}(\alpha)=c^{ij}$.

Enfin $b_n^i = a_n^i + F_n(k^i)$ si $k^i(x) = x^i-\ell(x^i)$. Comme ci-dessus, on voit que $F(k^i) \le \limsup_n F_n(k^i)$; comme $b_n^i \to b^i$ on en déduit $a^i \ge \liminf_n a_n^i$, d'où $a^i\ge 0$. □

Dans le théorème 2.5 on n'a pas en général la convergence $g_n(\gamma) \to g(\gamma)$ pour $\gamma\in\overline{\mathcal{B}}\backslash\mathcal{B}$, comme le montre l'exemple suivant: prendre $d=1$, $a_n=c_n=0$, $F_n = \frac{1}{n}\varepsilon_n$; on a alors les conditions de 2.5 avec $a=1$, $c=0$, $F=0$, tandis que $g_n(\gamma) = [\gamma n + 1 - (1+n)^\gamma]/n$ converge vers $g(\gamma) = \gamma$ pour $0\le\gamma<1$, mais pas pour $\gamma=1$.

Cependant, on a le résultat suivant, dans lequel on utilise les notations $\gamma_{i,\rho}\in\mathcal{B}$ et $\gamma_{ij}\in\overline{\mathcal{B}}$ ci-dessous:

2.9 $\quad \gamma_{i,\rho}^k = \begin{cases} \rho & \text{si } k=i \\ 0 & \text{si } k\neq i, \end{cases} \qquad \gamma_{ij}^k = \begin{cases} 1/2 & \text{si } k=i,j \\ 0 & \text{si } k\neq i,\ k\neq j. \end{cases}$

2.10 THEOREME. <u>Soit</u> $(a_n,c_n,F_n)\in\mathcal{Q}_+$ <u>et</u> $(a,c,F)=(0,c,0)\in\mathcal{Q}_+$. <u>Il y a équivalence entre les conditions (i)-(iv) de 4.5 et chacune des conditions suivantes</u> (où $g_n = g_{a_n,c_n,F_n}$, $g = g_{0,c,0}$):

(v) $g_n \to g$ <u>simplement sur</u> $\overline{\mathcal{B}}$,

(vi) $g_n(\gamma) \to g(\gamma)$ <u>pour</u> $\gamma=\gamma_{ij}$, $\gamma=\gamma_{i,1/2}$, $\gamma=\gamma_{i,\rho}$, $\gamma=\gamma_{i,1-\rho}$, <u>pour tous</u> $i\neq j$ <u>et pour un</u> $\rho\in]0,1/2[$,

(vii) <u>On a [C] et les deux propriétés</u>

[B'] $\quad a_n \to 0$,

[D'] $\quad \int_{|x|\ge\varepsilon} |x|F_n(dx) \to 0$ pour tout $\varepsilon>0$.

<u>Preuve</u>. (v) ⇒ (i) est évident, ainsi que (v) ⇒ (vi).

Pour la suite, on prend $h^i(x)=\ell(x^i)$ (cf. preuve de 2.5), et on note $\mathcal{C}_3$ l'ensemble des fonctions f continues telles que $f(x) = O(|x|)$ si $|x| \to \infty$ et $f(x) = o(|x|^2)$ si $|x| \to 0$. En particulier, les fonctions $x \to k^i(x) = x^i-\ell(x^i)$ sont dans $\mathcal{C}_3$.

Supposons maintenant (vii). [D'] implique que $F_n(f) \to F(f) = 0$ pour toute $f\in\mathcal{C}_3$, et comme $k^i\in\mathcal{C}_3$ on déduit [B] de [B'] et 2.4. Comme de plus $\hat\psi_\gamma\in\mathcal{C}_3$ si $\gamma\in\overline{\mathcal{B}}$, on déduit alors (v) de 2.7.

Supposons ensuite (i), donc (iii). On a $b_n^i = a_n^i + F_n(k^i) \to b^i = 0$, et $a_n^i\ge 0$, $k^i\ge 0$, d'où [B'] et $F_n^i(k\ ) \to 0$, donc $\int_{|x|\ge 2d} |x|F_n(dx) \to 0$.

[D'] découle alors de ce résultat et de $[D_1]$, et on a donc (vii).

Supposons enfin (vi). Posons

$$\alpha_n^i = g_n(\gamma_{i,\rho}) + g_n(\gamma_{i,1-\rho}) - 8\rho(1-\rho)g_n(\gamma_{i,1/2}),$$
$$\varphi^i = \psi_{\gamma_{i,\rho}} + \psi_{\gamma_{i,1-\rho}} - 8\rho(1-\rho)\psi_{\gamma_{i,1/2}},$$

et on définit de même α^i à partir de g. Par un calcul simple, on a

$$\alpha_n^i = F_n(\varphi^i) + [1 - 4\rho(1-\rho)]a_n^i,$$

et en particulier $\alpha^i=0$. (vi) implique $\alpha_n^i \to \alpha^i = 0$, et comme $\varphi^i \geq 0$, $a_n^i \geq 0$, on en déduit [B'] et $F_n(\varphi^i) \to 0$. On sait aussi [6, p.554] que pour tout $\varepsilon>0$ il existe une constante $K_{\varepsilon,\rho}$ avec $|x^i| \leq K_{\varepsilon,\rho}\varphi^i(x)$ sur $\{|x^i|\geq\varepsilon\}$, de sorte que [D'] découle de ce qui précède.

Par ailleurs, on voit comme dans l'étape 2 de la preuve de 2.5 qu'on a 2.6, donc comme dans l'étape 4 on voit que $F_n(f) \to 0$ pour toute $f\in C_2$, donc toute $f\in C_3$ d'après [D']. En particulier $F_n(\hat{\psi}_\gamma)\to 0$ pour tout $\gamma\in\overline{B}$. Donc (vi) et 2.7 entrainent

2.11 $$\sum \gamma^i(b_n^i + \tilde{c}_n^{ii}/2) - \frac{1}{2}\sum \gamma^i\gamma^j\tilde{c}_n^{ij} \longrightarrow g_\infty(\gamma) = \frac{1}{2}[\sum \gamma^i c^{ii} - \sum \gamma^i\gamma^j c^{ij}]$$

si $\gamma=\gamma_{i,\rho}$ ou $\gamma=\gamma_{ij}$. De plus $b_n^i \to 0$ par [B']+[D']. En appliquant 2.11 à $\gamma=\gamma_{i,1/2}$ on voit que $\tilde{c}_n^{ii} \to c^{ii}$; en appliquant 2.11 à $\gamma=\gamma_{ij}$ pour $i\neq j$, on voit que $\tilde{c}_n^{ij} \to c^{ij}$. On a donc [C].

On a ainsi prouvé que (vi)⇒(vii), ce qui achève la démonstration.□

3 - THEOREMES LIMITE POUR LES SURMARTINGALES POSITIVES

§3-a. Quelques préliminaires. Dans ce paragraphe, nous allons associer aux familles finies de surmartingales positives des processus prévisibles du type "processus de Hellinger".

Soit $L=(L^i)_{i\leq d}$ un processus dont les composantes L^i sont des surmartingales locales vérifiant $L_0^i=0$ et $\Delta L^i\geq -1$, sur un espace filtré $(\Omega,\mathcal{F},(\mathcal{F}_t),P)$. On note (B,C,ν) les caractéristiques de L relativement à une fonction de troncation continue h donnée sur $\mathbb{R}^d$. La mesure ν ne charge que $\mathbb{R}_+\times[-1,\infty[^d$. Il existe des processus prévisibles croissants A^i nuls en 0, tels que les L^i+A^i soient des martingales locales (décomposition de Doob-Meyer), et on a

3.1 $$B^i = -A^i + (h^i(x)-x^i)*\nu.$$

Pour tout $\gamma\in\overline{B}$ on introduit le processus croissant prévisible:

3.2 $$k(\gamma) = \sum_{i\leq d}\gamma^i(A^i + C^{ii}/2) - \frac{1}{2}\sum_{i,j\leq d}\gamma^i\gamma^j C^{ij} + \psi_\gamma*\nu$$

(voir 2.1 pour ψ_γ; $\sum \gamma^i C^{ii} - \sum \gamma^i \gamma^j C^{ij}$ est croissant par [7, 7.19]).

On associe à L le processus $V=(V^i)_{i\leq d}$ défini par

3.3 $$V^i = \mathcal{E}(L^i)$$

(exponentielle de Doléans): chaque V^i vérifie $V_0^i=1$ et est une surmartingale locale positive (car $\Delta L^i \geq 0$), donc aussi une surmartingale. Les $k(\gamma)$ s'interprètent comme les "processus de Hellinger" de la surmartingale vectorielle V, via la

3.4 PROPOSITION. <u>Pour tout</u> $\gamma \in \overline{\mathcal{B}}$ <u>on a</u> $\Delta k(\gamma) \leq 1$ <u>et, si</u> $Y(\gamma) = \prod_{i\leq d} (V^i)^{\gamma^i}$, <u>on a</u>

3.5 $$M(\gamma) := Y(\gamma) + Y(\gamma)_- \bullet k(\gamma) \text{ est une martingale}$$

<u>Preuve</u>. La première assertion découle de $\Delta A_t^i = -\int x^i \nu(\{t\}\times dx)$. Une simple application de la formule d'Ito, jointe à 3.2 et 3.3, montre que $M(\gamma)$ est une martingale locale, donc $M(\gamma)-Y(\gamma)_- \bullet k(\gamma)$ est la décomposition de Doob-Meyer de $Y(\gamma)$. Comme $Y(\gamma)$ est une surmartingale (non locale) d'après Hölder, $M(\gamma)$ est aussi une martingale. □

<u>§3-b. Les théorèmes généraux</u>. Soit $L^n=(L^{n,i})_{i\leq d}$ des processus (chacun sur son espace filtré $(\Omega^n,\mathcal{F}^n,(\mathcal{F}_t^n),P^n)$), dont les composantes sont des surmartingales locales vérifiant $L_0^{n,i}=0$ et $\Delta L^{n,i}\geq -1$. On leur associe $k^n(\gamma)$ et $V^n=(V^{n,i})_{i\leq d}$ par 3.2 et 3.3.

3.6 THEOREME. <u>Soit</u> $\mathcal{B}'$ <u>dense dans</u> $\mathcal{B}$. <u>Supposons que pour tout</u> $t\geq 0$,

3.7 $$k(\gamma)_t^n \xrightarrow{P^n} k(\gamma)_t \qquad \forall \gamma \in \mathcal{B}',$$

<u>où chaque</u> $t \to k(\gamma)_t$ <u>est une fonction</u> (déterministe) <u>continue de</u> $\mathbb{R}_+$ <u>dans</u> $\mathbb{R}_+$. <u>Alors</u>:

a) <u>Les</u> $k(\gamma)$ <u>sont associés par 3.2 à un unique (en loi) processus</u> $L=(L^i)_{i\leq d}$, <u>qui est un PAI</u> (processus à accroissements indépendants) <u>sans discontinuité fixe, et chaque</u> L^i <u>est une surmartingale vérifiant</u> $\Delta L^i \geq -1$.

b) <u>Les processus</u> L^n <u>convergent en loi vers</u> L.

c) <u>Les processus</u> V^n <u>convergent en loi vers</u> $V=(V^i)_{i\leq d}$, <u>où</u> $V^i = \mathcal{E}(L^i)$.

3.8 COMMENTAIRES. Le résultat important pour les applications est la partie (c), qui découle aussi du résultat plus général suivant:

Soit V un processus dont les composantes sont des surmartingales positives, égales à 1 en 0, et $Y(\gamma) = \prod (V^i)^{\gamma^i}$; comme en [7, 2.3] il existe un processus croissant prévisible $\overline{k}(\gamma)$ tel que $Y(\gamma) + Y(\gamma)_- \bullet \overline{k}(\gamma)$ soit une martingale, et ce processus est unique sur l'ensemble aléatoire $\{Y(\gamma)_->0\}$. Noter que si les V^i ne s'annulent pas, alors ils se mettent sous la forme 3.3 et on a $\overline{k}(\gamma)=k(\gamma)$; lorsque les V^i sont de la forme 3.3, mais peuvent s'annuler, $k(\gamma)$ est une version de $\overline{k}(\gamma)$, mais il existe d'autres versions; enfin il faut remarquer que les V^i (s'ils peuvent s'annuler) ne se mettent pas nécessairement sous la forme 3.3.

Soit aussi V^n des processus du même type, avec les processus $\overline{k}^n(\gamma)$ associés (on en prend des versions quelconques). Si alors $D\subset\mathbb{R}_+$, si $\mathcal{B}'$ est dense dans $\mathcal{B}$, et si les $k(\gamma)_t$ sont déterministes pour tous $t\in D$, $\gamma\in\mathcal{B}'$, on a les deux assertions:

(i) sous

3.9 $$\mathcal{E}(-k^n(\gamma))_t \xrightarrow{P^n} \mathcal{E}(-k(\gamma))_t, \quad \forall t\in D,\ \forall \gamma\in\mathcal{B}',$$

les V^n convergent en loi vers V, fini-dimensionnellemnt le long de D;

(ii) si en outre D est dense dans $\mathbb{R}_+$ et $t \to k(\gamma)_t$ est continu pour tout $\gamma\in\mathcal{B}'$, 3.9 équivaut à 3.7, et entraine la convergence fonctionnelle en loi de V^n vers V.

La preuve de ces deux résultats consiste en modifications mineures des démonstrations de [8]. La preuve que nous présentons ici (pour un résultat plus faible) est toutefois notablement plus simple. □

<u>Preuve de 3.6(a)</u>. On peut toujours supposer $\mathcal{B}'$ dénombrable. Quitte à prendre une sous-suite (pour chaque temps t), on peut supposer que $k^n(\gamma)_t(\omega) \to k(\gamma)_t$ pour tout $\gamma\in\mathcal{B}'$ et tout $\omega\notin N_t$, où N_t est un ensemble de mesure nulle (sur le produit $\Omega = \prod \Omega^n$, $P = \otimes P^n$).

Fixons alors $\omega\notin N_t$, et appliquons 2.5 à $a_n = A^n_t$, $c_n = C^n_t$ et $F_n = \nu^n([0,t]\times .)$: il existe un triplet $(A_t, C_t, G_t)\in\Omega_+$, unique d'après 2.3, tel que $g_{A_t,C_t,G_t}(\gamma) = k(\gamma)_t$ pour tout $\gamma\in\mathcal{B}'$.

Si on refait le même raisonnement avec $a_n = A^n_t - A^n_s$, $c_n = C^n_t - C^n_s$, $F_n(dx) = \nu^n(]s,t]\times dx)$ pour $s<t$, on obtient comme limite la fonction $k(\gamma)_t - k(\gamma)_s$, associée à $A_t - A_s$, $C_t - C_s$, $G_t - G_s$: donc les $t \to A^i_t$ sont croissants, $t \to C_t$ est croissant pour l'ordre fort de $\mathcal{M}_d$, et $G_t - G_s$ est une mesure positive.

Si $t_n \to t$, on a $k(\gamma)_{t_n} \to k(\gamma)_t$ par hypothèse pour $\gamma\in\mathcal{B}'$, de sorte qu'une nouvelle application de 2.5 montre que $t \to G_t(f)$ est continue pour toute $f\in C_1$: il existe donc une mesure positive ν sur $\mathbb{R}_+\times\mathbb{R}^d$ qui ne charge que $\mathbb{R}_+\times[-1,\infty[^d$, telle que $(|x|\wedge|x|^2)*\nu_t<\infty$ et $\nu(\{t\}\times\mathbb{R}^d)=0$ pour tous t, et $f*\nu_t = G_t(f)$ si $f\in C_1$; ainsi $k(\gamma)_t$ est associé à A,C,ν par 3.2.

Enfin si on définit B par 3.1, toujours à cause de la continuité de $t \to k(\gamma)_t$ et de 2.5, on voit que $t \to B_t$ et $t \to C_t$ sont continues, et par construction B est à variation finie sur les compacts. Ainsi (B,C,ν) est le triplet des caractéristiques (déterministes) d'un PAI L sans discontinuité fixe. Comme ν ne charge que $\mathbb{R}_+\times[-1,\infty[^d$ on a $\Delta L^i \geq -1$. Enfin comme on a 3.1 et que les A^i sont croissants, chaque composante L^i est clairement une surmartingale.

Pour terminer, 2.3 implique que les $k(\gamma)$ $(\gamma\in\mathcal{B}')$ déterminent de manière unique A, C, ν (par 3.2), donc B (par 3.1): donc le PAI L ci-dessus est la seule semimartingale (en loi) associée aux $k(\gamma)$ par 3.2. □

Pour la suite, nous rappelons quelques résultats sur la convergence des semimartingales [6]. Posons

3.10
$$\begin{cases} \hat{C}^{n,ij}_t = C^{n,ij}_t + (h^i h^j) * \nu^n_t \\ \tilde{C}^{n,ij}_t = \hat{C}^{n,ij}_t - \sum_{s\leq t} \Delta B^{n,i}_s \, \Delta B^{n,j}_s . \end{cases}$$

On associe de même $\hat{C}=\tilde{C}$ (qui est continu) au PAI L associé aux $k(\gamma)$ par 3.6(a). Considérons les conditions

[β] $\quad B^n_t \xrightarrow{P^n} B_t \quad \forall t\geq 0$,

[Sup-β] $\quad \sup_{s\leq t} |B^n_s - B_s| \xrightarrow{P^n} 0 \quad \forall t\geq 0$,

[γ] $\quad \tilde{C}^n_t \xrightarrow{P^n} \tilde{C}_t \quad \forall t\geq 0$,

[γ'] $\quad \hat{C}^n_t \xrightarrow{P^n} \tilde{C}_t \quad \forall t\geq 0$,

[δ] $\quad f*\nu^n_t \xrightarrow{P^n} f*\nu_t \quad \forall t\geq 0,\ \forall f\in C_1'$,

où C_1' désigne l'ensemble des fonctions continues bornées sur $\mathbb{R}^d$, nulles autour de 0. Rappelons que [Sup-β]+[γ]+[δ] implique la convergence en loi de L^n vers L.

3.11 LEMME. Sous les hypothèses de 3.6, il y a équivalence entre

(i) [Sup-β]+[γ]+[δ];

(ii) [β]+[γ']+[δ];

(iii) on a 3.7 pour tout $t\geq 0$.

Preuve. Sous [δ], les conditions [β] et [Sup-β] ne dépendent pas de la troncation h choisie (comme avant l'énoncé de 2.5). On choisit donc $h^i(x) = \ell(x^i)$ comme dans la preuve de 2.5. Par suite $h^i(x)-x^i\leq 0$ si $x^i\geq -1$, et on déduit de 3.1 que $B^{n,i}$ et B^i sont décroissants; comme de plus $t \to B^i_t$ est continu, il est bien connu que [β] $\Leftrightarrow$ [Sup-β]. De plus

$$\sum_{s\leq t} |\Delta B_s^{n,i}\ \Delta B_s^{n,j}| \ \leq\ (-B_t^{n,i})\ \sup_{s\leq t}|\Delta B_s^{n,j}|,$$

et [Sup-β] entraine que l'expression précédente tend vers 0 en P^n-probabilité: étant donné 3.10, on a donc l'équivalence (i) $\Leftrightarrow$ (ii).

Par ailleurs, si $a_n=A_t^n$, $c_n=C_t^n$, $F_n=\nu^n([0,t]\times .)$, les quantités associées par 2.4 sont $b_n=B_t^n$ et $\tilde{c}_n=\hat{C}_t^n$. Etant donné 3.2, on déduit alors immédiatement l'équivalence (ii) $\Leftrightarrow$ (iii) de 2.5. □

Preuve de 3.6(b,c). Comme [Sup-β]+[γ]+[δ] entraine la convergence en loi de L^n vers L, (b) vient de 3.11. Quant à (c), il suffit d'utiliser (b) et de combiner les théorèmes 3.2(b) et 4.2 de Jakubowski, Mémin et Pagès [10] (l'extension multidimensionnelle des résultats de [10] est très facile, selon la méthode de Mémin [12]). □

§3-c. Limite continue gaussienne. Un cas particulier intéressant du point de vue des applications est celui où le processus limite L est une martingale gaussienne continue. Ses caractéristiques (B,C,ν) vérifient alors $B=0$, $\nu=0$, et on a simplement

$$k(\gamma)\ =\ \tfrac{1}{2}[\sum \gamma^i C^{ii}\ -\ \sum \gamma^i\gamma^j C^{ij}].$$

Au vu de 2.10, pour avoir la convergence il suffit de vérifier 3.7 avec $\mathcal{B}'$ remplacé par (cf. 2.9).

3.12 $\qquad \mathcal{B}'_\rho\ =\ \{\gamma_{ij},\ \gamma_{i,1/2},\ \gamma_{i,\rho},\ \gamma_{i,1-\rho}:\ 1\leq i,j\leq d,\ i\neq j\},$

où ρ est un nombre fixé dans $]0,1/2[$. On peut aussi obtenir un résultat de convergence fini-dimensionnelle en temps (peut-être plus utile que la convergence fonctionnelle). Voici l'énoncé complet:

3.13 THEOREME. Supposons que L soit une martingale gaussienne continue; soit $D\subset\mathbb{R}_+$, $\rho\in]0,1/2[$ et $\mathcal{B}'_\rho$ défini par 3.12.

a) Les conditions suivantes sont équivalentes:

(i) on a 3.7 pour tout $t\in D$;

(ii) on a $k^n(\gamma)_t \xrightarrow{P^n} k(\gamma)_t \quad \forall t\in D,\ \forall\gamma\in\mathcal{B}'_\rho$;

(iii) on a $B_t^n \xrightarrow{P^n} 0$, $\tilde{C}_t^n \xrightarrow{P^n} C_t$, $1_{\{|x|\geq\varepsilon\}}*\nu_t^n \xrightarrow{P^n} 0 \quad \forall t\in D,\ \forall\varepsilon>0$;

(iv) on a $A_t^n \xrightarrow{P^n} 0$, $\hat{C}_t^n \xrightarrow{P^n} C_t$, $(|x|1_{\{|x|\geq\varepsilon\}})*\nu_t^n \xrightarrow{P^n} 0 \quad \forall t\in D,\forall\varepsilon>0$.

b) Les conditions ci-dessus assurent la convergence en loi fini-dimensionnelle le long de D des processus L^n et V^n, vers L et V respectivement.

c) Si de plus D est dense dans $\mathbb{R}_+$, elles entrainent aussi la convergence en loi fonctionnelle.

Preuve. a) Comme dans 3.11, on sait que $B^{n,i}$ est décroissant; donc $|\hat{C}_t^{n,ij} - \tilde{C}_t^{n,ij}| \le (-B_t^{n,i})(-B_t^{n,j})$ par 3.10, et $B_t^n \to 0$ entraine l'équivalence de $\hat{C}_t \to C_t$ et de $\tilde{C}_t \to C_t$. L'équivalence (i) ⇔ (iii) se montre alors comme la fin de 3.11, à l'aide de 2.5. Enfin les équivalences (i) ⇔ (ii) ⇔ (iv) découlent de 2.10.

b) Comme $t \to k(\gamma)_t$ est continu croissant, (i) avec D dense implique 3.7 pour tout $t \ge 0$. L'assertion découle alors de 3.6.

c) D'après [6, VIII.2.4], (iii) implique que

3.14 L^n converge vers L en loi fini-dimensionnellement le long de D.

Par ailleurs, on a

$$V_t^{n,i} = W_t^{n,i} \exp(L_t^{n,i} - \hat{C}_t^{n,ii}/2),$$

où

$$W_t^{n,i} = [\prod_{s\le t}(1 + \Delta L_s^{n,i})\exp(-\Delta L_s^{n,i})] \exp[\tfrac{1}{2}(h^i)^2 * \nu_t^n].$$

D'après 3.14, la seconde propriété dans (iv) et la formule $V^i = \exp(L^i - C^{ii}/2)$, pour obtenir la convergence fini-dimensionnelle de V^n vers V le long de D, il nous suffit de montrer que:

3.15 $$W_t^{n,i} \xrightarrow{P^n} 1 \quad \forall t \in D.$$

Dans la suite on fixe i. Si $\varepsilon \in]0,\frac{1}{2}]$, soit $T_\varepsilon^n = \inf(t: |\Delta L_s^{n,i}| \ge \varepsilon)$. D'après la troisième propriété (iii) et l'inégalité de Lenglart,

3.16 $$P^n(T_\varepsilon^n \le t) \longrightarrow 0 \quad \forall t \in D.$$

Par ailleurs $h^i(x) = \ell(x^i) = x^i$ si $|x^i| \le 1$, donc si $\alpha^n = h^i(\Delta L^n)$ on a

$$\text{Log } W_t^{n,i} = \tfrac{1}{2}(h^i)^2 * \nu_t^n - \sum_{s\le t}[\alpha_s^n - \text{Log}(1+\alpha_s^n)]$$

sur $[0,T_\varepsilon^n[$. Il existe aussi une constante K telle que pour $|y| \le \frac{1}{2}$, $|\text{Log}(1+y) - y + y^2/2| \le K|y|^3$. Donc

3.17 $$|\text{Log } W_t^{n,i,\varepsilon} - U_t^n/2| \le K\varepsilon \sum_{s\le t}(\alpha_s^n)^2 \quad \text{sur } [0,T_\varepsilon^n[,$$

où

$$U_t^n = (h^i)^2 * \nu_t^n - \sum_{s\le t}(\alpha_s^n)^2.$$

Par définition du compensateur ν^n, U^n est une martingale locale et, ses sauts étant bornées, elle est localement de carré intégrable. Par [6], $\langle U^n,U^n\rangle = (h^i)^4 * \nu^n - \sum_{s\le\cdot}[\nu^n(\{s\}\times(h^i)^2]^2$. Comme $|h^i| \le 2$,

$$\langle U^n,U^n\rangle \le (h^i)^4 * \nu^n \le \eta^2\hat{C}^{n,ii} + 2^4\, 1_{\{|x|\ge\eta\}} * \nu^n$$

pour tout $\eta > 0$. On déduit alors des deux dernières propriétés (iv) que $\langle U^n,U^n\rangle_t \xrightarrow{P^n} 0$ si $t \in D$, donc par l'inégalité de Lenglart:

3.18 $$\sup_{s\le t}|U_s^n| \xrightarrow{P^n} 0 \quad \forall t \in D.$$

Enfin, le compensateur de $\sum_{s\leq .}(\alpha_s^n)^2$ est majoré par $\hat{C}^{n,ii}$, donc une dernière application de l'inégalité de Lenglart donne

3.19 $\qquad \lim_{a\uparrow\infty} \limsup_n P^n(\sum_{s\leq t}(\alpha_s^n)^2 > a) = 0 \qquad \forall t\in D.$

Il est alors immédiat de déduire 3.15 de 3.16, 3.17, 3.18 et 3.19, et le résultat est démontré. □

4 - APPLICATIONS AUX VRAISEMBLANCES PARTIELLES

§4-a. Rappels sur la convergence des vraisemblances. Pour chaque n entier, soit un modèle statistique filtré $\mathcal{E}^n=(\Omega^n,\mathcal{F}^n,(\mathcal{F}_t^n),(P_\theta^n)_{\theta\in\Theta})$; l'espace des paramètres Θ est quelconque, mais indépendant de n. Pour simplifier, on suppose que toutes les probabilités P_θ^n $(\theta\in\Theta)$ coïncident sur $\mathcal{F}_0^n$. On note $Z^{n,\zeta/\theta}$ le processus de vraisemblance de P_ζ^n par rapport à P_θ^n: c'est une P_θ^n-surmartingale positive, $Z_0^{n,\zeta/\theta}=1$.

Notons $\mathfrak{a}$ l'ensemble des applications α: $\Theta \to \mathbb{R}_+$ à support fini, avec $\sum\alpha^\theta = 1$. Si I est une partie finie de Θ avec $\theta\in I$, notons $\mathfrak{a}(\theta,I)$ l'ensemble des $\alpha\in\mathfrak{a}$ à support dans I, vérifiant $0<\alpha^\theta<1$.

Rappelons alors les résultats de convergence obtenus dans [8]. On désigne par $h^n(\alpha)$ une version du processus de Hellinger d'ordre $\alpha\in\mathfrak{a}$ pour l'expérience $\mathcal{E}^n$. On se donne aussi une expérience "limite" $\mathcal{E} = (\Omega,\mathcal{F},(\mathcal{F}_t),(P_\theta)_{\theta\in\Theta})$, avec les processus de vraisemblance $Z^{\zeta/\theta}$, et vérifiant $P_\theta=P_\zeta$ sur $\mathcal{F}_0$ pour tous θ,ζ, et:

4.1 Hypothèse H1: Il existe des versions $h(\alpha)$ des processus de Hellinger de $\mathcal{E}$ qui sont déterministes, continues en temps. □

Les théorèmes 3.1 et 4.1 de [8] peuvent alors s'énoncer ainsi, compte tenu des hypothèses faites ici (on reconnaitra les résultats du commentaire 3.8, en écrivant $I = \{\theta_0,\theta_1,\dots,\theta_d\}$ avec $\theta_0=\theta$, et $V^{n,i} = Z^{n,\theta_i/\theta}$, de sorte que $h^n(\alpha) = k^n(\gamma)$ où $\gamma\in\mathcal{B}$ est défini par $\gamma^i = \alpha^{\theta_i}$ pour $1\leq i\leq d$):

4.2 THEOREME. Supposons H1, et soit I une partie finie de Θ, et $\theta\in I$, et D une partie de $\mathbb{R}_+$.

a) Si $\mathcal{E}(-h(\alpha)^n)_t \to \mathcal{E}(-h(\alpha))_t$ en P_θ^n-probabilité pour tous $t\in D$, $\alpha\in\mathfrak{a}(\theta,I)$, les processus $(Z^{n,\zeta/\theta})_{\zeta\in I}$ (sous P_θ^n) convergent en loi, fini-dimensionnellement le long de D, vers $(Z^{\zeta/\theta})_{\zeta\in I}$ (sous P_θ).

b) Si $h(\alpha)_t^n \to h(\alpha)_t$ en P_θ^n-probabilité pour tous $t\in D$, $\alpha\in\mathfrak{a}(\theta,I)$,

<u>et si</u> D <u>est dense dans</u> $\mathbb{R}_+$, <u>les processus</u> $(Z^{n,\zeta/\theta})_{\zeta\in I}$ (sous P^n_θ) <u>convergent fonctionnellement vers</u> $(Z^{\zeta/\theta})_{\zeta\in I}$ (sous P_θ).

<u>§4-b. Les vraisemblances partielles</u>. Pour chaque n on se donne aussi un processus X^n (à valeurs dans $\mathbb{R}^{d_n}$), qui est une semimartingale pour chaque P^n_θ. Dans [9] nous avons défini pour tous θ,ζ:

1) Un intervalle "maximal" $\Sigma^{n,\zeta/\theta}$ de la forme $\bigcup_p [\![0,S_p]\!]$ pour des $(\mathcal{F}^n_t)$-temps d'arrêt S_p convenables (c'est l'intervalle Σ_2 de [9]).

2) Une surmartingale locale $\overline{L}^{n,\zeta/\theta}$ sur $\Sigma^{n,\zeta/\theta}$ (i.e., le processus arrêté en chaque S_p ci-dessus est une surmartingale locale), nulle en 0 et vérifiant $\Delta\overline{L}^{n,\zeta/\theta}\geq -1$ ($\Sigma^{n,\zeta/\theta}$ et $\overline{L}^{n,\zeta/\theta}$ dépendent de manière essentielle des caractéristiques de X^n sous P^n_θ et P^n_ζ).

3) Finalement, le processus $\overline{Z}^{n,\zeta/\theta}=\mathcal{E}(\overline{L}^{n,\zeta/\theta})$, qui est défini aussi sur $\Sigma^{n,\zeta/\theta}$ et est une surmartingale locale positive sur cet intervalle: il est appelé <u>processus de vraisemblance partielle</u> de P^n_ζ par rapport à P^n_θ, relativement à X^n.

Rappelons aussi que $\{Z^{n,\zeta/\theta}_->0\}\subset\Sigma^{n,\zeta/\theta}$, $\Sigma^{n,\theta/\theta}=\mathbb{R}_+$, $\overline{Z}^{n,\theta/\theta}=1$.

Soit I une partie finie de Θ et $\theta\in I$. Posons

4.3 $$\Sigma^{n,\theta,I} = \bigcap_{\zeta\in I}\Sigma^{n,\zeta/\theta}.$$

Si $I\neq\{\theta\}$, on peut écrire $I=\{\theta,\theta_1,\ldots,\theta_d\}$, et on note $L^n=L^n(I)$ le processus d-dimensionnel de composantes $L^{n,i}=\overline{L}^{n,\theta_i/\theta}$: c'est un processus du type de ceux étudiée au §3, sauf qu'il n'est défini que sur l'intervalle $\Sigma^{n,\theta,I}$. Pour tout $\alpha\in\mathcal{Q}$ de support dans I, on appelle <u>processus de Hellinger partiel</u> d'ordre α, relativement à (θ,I), le processus $\overline{h}^n(\alpha,\theta,I)=k^n(\gamma)$ associé à L^n par 3.2, <u>sur</u> $\Sigma^{n,\theta,I}$, avec $\gamma^i=\alpha^{\theta_i}$ pour $1\leq i\leq d$ (cf. [8], formule 7.9).

Si les véritables vraisemblances égalent les vraisemblances partielles, $\overline{h}^{n}(\alpha,\theta,I)$ est une version de $h^n(\alpha)$.

Voici maintenant le théorème limite qui constitue l'un des résultats motivant cet article; pour simplifier l'énoncé nous posons (de manière arbitraire) $\overline{Z}^{n,\zeta/\theta}_t=0$ si $t\notin\Sigma^{n,\zeta/\theta}$ et $\overline{h}^n(\alpha,\theta,I)_t=\infty$ si $t\notin\Sigma^{n,\theta,I}$.

4.4 THEOREME. <u>Supposons H1, et soit</u> I <u>une partie finie de</u> Θ, <u>et</u> $\theta\in I$. <u>Supposons qu'on ait les deux proprétés suivantes</u>:

4.5 $$P^n_\theta(t\in\Sigma^{n,\theta,I})\longrightarrow 1 \quad\text{pour tout } t\geq 0.$$

4.6 $\qquad \overline{h}^n(\alpha,\theta,I)_t \longrightarrow h(\alpha)_t$ en P^n_θ-probabilité, $\forall t\geq 0$, $\forall \alpha\in\mathfrak{A}(\theta,I)$.

Alors, les processus $(\overline{Z}^{n,\zeta/\theta})_{\zeta\in I}$ sous P^n_θ convergent en loi vers $(Z^{\zeta/\theta})_{\zeta\in I}$ sous P_θ.

Preuve. Il suffit de montrer les résultats sur tout intervalle fini, donc on fixe $t\geq 0$. On peut écrire $\Sigma^{n,\theta,I} = \bigcup_{p\geq 1}[\![0,S(n,p)]\!]$ pour des temps d'arrêt $S(n,p)$ croissant en p, donc par 4.5 il existe $p_n\in\mathbb{N}$ avec $P^n_\theta(S(n,p_n)<t)\to 0$. Il suffit donc de regarder la limite des V^n où $V^{n,i} = \mathcal{E}(L^{n,i}_{\cdot\wedge T(n,p_n)})$, et le résultat découle de 3.6. $\square$

4.7 REMARQUE. Nous avons énoncé ce théorème ainsi par souci de simplicité. En utilisant toute la force de 3.6, si on a 4.5 et 4.6 avec des fonctions déterministes continues $h(\alpha)$ on montre qu'il existe nécessairement une expérience $\mathcal{E}$ vérifiant H1, avec les $h(\alpha)$ pour processus de Hellinger. $\square$

4.8 REMARQUE. Il existe aussi une version de 4.4 dans laquelle les processus limite $Z^{\zeta/\theta}$ sont remplacés par des vraisemblances partielles $\overline{Z}^{\zeta/\theta}$, à condition de supposer que les ensembles Σ^θ associés par 4.3 sont déterministes, et de rempacer 4.4 par $P^n_\theta(t\in\Sigma^{n,\theta}) \to 1_{\Sigma^\theta}(t)$. Cela ne semble pas d'un grand intérêt ... $\square$

4.9 REMARQUE. On peut définir des processus de Hellinger partiels "non canoniques" par la propriété 3.5 sur $\Sigma^{n,\theta,I}$, pour V^n remplacé par $(\overline{Z}^{n,\zeta/\theta})_{\zeta\in I}$: grâce au résultat énoncé (mais non démontré) dans 3.8, le théorème 4.4 reste alors valide (ainsi qu'un résultat de convergence fini-dimensionnelle en temps). Cela dit, dans les applications ce sont les processus de Hellinger partiels "canoniques" qui s'expriment simplement en fonction des caractéristiques des X^n (cf. [9], 7.7), donc seul l'énoncé ci-dessus de 4.4 semble avoir un réel intérêt. $\square$

§4-c. Normalité asymptotique. Un cas particulier important est celui où le modèle limite $\mathcal{E}$ est gaussien: on entend par là que les $Z^{\zeta/\theta}$ sont continus (en temps), ne s'annulent pas, et que leurs logarithmes sont des processus gaussiens. D'après [7], cela revient à dire qu'on a H1 avec des fonctions $h(\alpha)$ à valeurs finies et vérifiant

4.10 $$h(\alpha) = 2\sum_{\theta\neq\zeta} \alpha^\theta\alpha^\zeta\, h(\alpha_{\theta\zeta}),$$

où $\alpha_{\theta\zeta}\in\mathfrak{A}$ est caractérisé par ses coordonnées $\alpha^\theta_{\theta\zeta} = \alpha^\zeta_{\theta\zeta} = 1/2$.

Toujours d'après [7], il existe alors une fonction $\mathbb{R}_+\times\Theta^2\ni(t,\theta,\zeta)$

$\to \sigma_t(\theta,\zeta)$ qui est continue en t, nulle pour $t=0$, telle que $\sigma_t-\sigma_s$ soit de type positif sur Θ^2 si $s\leq t$, et telle que

4.11 $$h(\alpha_{\theta\zeta})_t = \frac{1}{8}[\sigma_t(\theta,\theta) + \sigma_t(\zeta,\zeta) - 2\sigma_t(\theta,\zeta)].$$

De plus, $\sigma_t(\theta,\zeta)$ égale la covariance de $\text{Log } Z_t^{\theta/\eta}$ et $\text{Log } Z_t^{\zeta/\eta}$ sous n'importe quelle mesure P_ρ, et pour n'importe quel η.

Inversement, si on se donne σ comme ci-dessus, il existe un modèle filtré $\mathcal{E}$ admettant les processus de Hellinger $h(\alpha)$ donnés par 4.10 et 4.11, et ce modèle est gaussien au sens ci-dessus.

Pour simplifier, on notera $H^{\zeta\eta}(\rho)$ et $\overline{H}^{\zeta\eta}(\rho,\theta,I)$, pour $\rho\in]0,1[$, les processus $h(\alpha)$ et $\overline{h}^n(\alpha,\theta,I)$, pour $\alpha\in\mathfrak{A}$ donné par $\alpha^\zeta=\rho$ et $\alpha^\eta=1-\rho$: ainsi $h(\alpha_{\theta\zeta}) = H^{\theta\zeta}(1/2)$ dans 4.10.

On déduit alors immédiatement de 3.13 le

4.12 THEOREME. Soit les hypothèses ci-dessus concernant les $h(\alpha)$. Soit I une partie finie de Θ, $\theta\in\Theta$ et $D\subset\mathbb{R}_+$.

a) Si on a

4.13 $$P_\theta^n(t\in\Sigma^{n,\theta,I}) \longrightarrow 1 \quad \forall t\in D,$$

4.14 $$\overline{H}^{n,\zeta\eta}(\tfrac{1}{2},\theta,I)_t \longrightarrow H^{\zeta\eta}(\tfrac{1}{2})_t \text{ en } P_\theta^n\text{-probabilité}, \quad \forall t\in D,\ \forall\zeta,\eta\in I,$$

4.15 $$\overline{H}^{n,\theta\zeta}(\rho,\theta,I)_t \longrightarrow H^{\theta\zeta}(\rho)_t \text{ en } P_\theta^n\text{-probabilité}, \quad \forall t\in D,\ \forall\zeta\in I$$

pour un $\rho\in]0,1/2[$, les processus $(\overline{Z}^{n,\zeta/\theta})_{\zeta\in I}$ sous P_θ^n convergent en loi, fini-dimensionnellement le long de D, vers $(Z^{\zeta/\theta})_{\zeta\in I}$.

b) Si de plus D est dense dans $\mathbb{R}_+$, il y a même convergence fonctionnelle, en loi.

4.16 REMARQUE. Si D n'est pas dense, il n'y a pas de raison de supposer les $h(\alpha)_t$ définis a priori pour tout $t\in\mathbb{R}_+$. Si par exemple $D=\{1\}$, il suffit de considérer les $h(\alpha)_1$, définis par 4.10 à l'aide des $h(\alpha_{\theta\zeta})_1$: on peut alors se donner σ_1, et considérer $\sigma_t = t\sigma_1$ pour avoir un modèle $\mathcal{E}$ filtré. □

Le résultat 4.12(a) est apparenté à la normalité asymptotique locale (modèles "LAN") de LeCam et Hajek. Plus précisément, supposons que $D=\{1\}$, et donc $\sigma_t=t\sigma_1$ dans 4.11, d'après 4.16. Supposons également que Θ soit un espace vectoriel, et que l'application σ_1 soit bilinéaire; elle définit alors un produit scalaire sur Θ, et la semi-norme correspondante est notée $\|\theta\|$. 4.10 et 4.11 s'écrivent alors pour $t=1$:

4.17 $$h(\alpha)_1 = \frac{1}{4}\sum_{\theta,\zeta} \|\theta-\zeta\|^2.$$

Si de plus Θ est <u>séparable</u> pour $\|.\|$, un résultat de LeCam [11, p. 176] permet de déduire de 4.12(a) le

4.18 THEOREME. <u>Sous les hypothèses précédentes, si on a 4.13, 4.14 et 4.15 pour</u> $D=\{1\}$, $\theta=0$ <u>et toute partie finie</u> I <u>de</u> Θ <u>contenant</u> 0, <u>il existe des processus linéaires</u> $(U^n_\zeta)_{\zeta\in\Theta}$ <u>sur</u> $(\Omega^n,\mathcal{F}^n_1,P^n_0)$ <u>avec</u>:

(i) U^n_ζ <u>est asymptotiquement normale centrée de variance</u> $\|\zeta\|^2$, <u>sous</u> P^n_0;

(ii) <u>pour chaque</u> $\zeta\in\Theta$, $\mathrm{Log}\ \overline{Z}^{n,\zeta/0}_1 - U^n_\zeta + \frac{1}{2}\|\zeta\|^2$ <u>tend en</u> P^n_0<u>-probabilité vers</u> 0.

§4-d. Un exemple. Ici nous allons retrouver des résultats de normalité asymptotiques, dûs à Greenwood et Wefelmeyer [3], [4], à l'aide du théorème 4.12 (ou 4.18). Le cadre est celui des processus ponctuels, mais des résultats analogues s'obtiendraient facilement dans le cas où on observe des semimartingales quelconques.

Pour chaque n, on observe un processus ponctuel multivarié μ^n à valeurs dans un espace E_n (dans [3,4] il s'agit de la superposition de n processus ponctuels simples sans sauts communs), dont le compensateur $\overline{\nu}^n$ s'écrit sous P^n_θ:

4.19 $$\overline{\nu}^n(\omega,ds,dx) = y^n_\theta(\omega,s,x)\ ds\ F_n(dx)$$

pour une mesure positive finie F_n sur E_n, et une fonction prévisible positive y^n_θ. On suppose que Θ est un espace vectoriel, et qu'il existe des fonctions prévisibles $D^n_\theta(\omega,t,x)$ telles que les $\theta\to D^n_t$ soient linéaires, et qu'on ait les convergences suivantes en P^n_0-probabilité, si $n\uparrow\infty$:

4.20 $$\int_{[0,1]\times E_n} [\sqrt{y^n_\theta/y^n_0} - 1 - D^n_\theta/2c_n]^2\ y^n_0\ ds\ F_n(dx) \longrightarrow 0,$$

4.21 $$\int_{[0,1]\times E_n} y^n_\theta\ 1_{\{y^n_0=0\}}\ ds\ F_n(dx) \longrightarrow 0,$$

4.22 $$\int_{[0,1]\times E_n} (D^n_\theta/c_n)^2\ 1_{\{|D^n_\theta/c_n|\geq\varepsilon\}}\ y^n_0\ ds\ F_n(dx) \longrightarrow 0 \quad \forall\varepsilon>0,$$

4.23 $$\int_{[0,1]\times E_n} (D^n_\theta/c_n)^2\ y^n_0\ ds\ F_n(dx) \longrightarrow u(\theta),$$

où c_n est une suite de réels positifs (qui dans les applications tendent vers $+\infty$), et u une fonction: $\Theta\to\mathbb{R}_+$.

Comme $\theta\to D^n_\theta$ est linéaire, on déduit aisément que $\|\theta\| = u(\theta)$ est une semi-norme sur Θ, associée à un produit scalaire que nous notons $\sigma_1(\theta,\zeta)$. On a alors le résultat suivant, dû à Greenwood et Wefelmeyer [4], et qui contient aussi le résultat de convergence de

[3]: ci-dessous, $\bar{Z}^{n,\theta/0}$ désigne le processus de vraisemblance partielle associé à μ^n, et le modèle limite est celui associé aux $h(\alpha)$ définis par 4.10 et 4.11 avec $\sigma_t = t\sigma_1$ (cf. 4.16).

4.24 THEOREME. <u>Sous les hypothèses précédentes, on a les conclusions du théorème 4.18</u>.

<u>Preuve</u>. Soit la fonction $\varphi_\rho(x,y) = \rho x + (1-\rho)y - x^\rho y^{1-\rho}$ sur $\mathbb{R}_+^2$, pour $\rho\in]0,1[$. D'après [9, 7.8] et 4.19, les processus de Hellinger partiels s'écrivent

4.25 $$\bar{H}^{n,\theta\zeta}(\rho)_1 = \int_{[0,1]\times E_n} \varphi_\rho(y_\theta^n, y_\zeta^n)\ ds\ F_n(dx),$$

tandis que $\Sigma^{n,0,I} = \mathbb{R}_+$ pour toute partie finie I; on a donc 4.13. On note $\hat{H}^{n,\theta\zeta}(\rho)_1$ et $\tilde{H}^{n,\theta\zeta}(\rho)_1$, repectivement, les intégrales sur $\{y_0^n>0\}$ et sur $\{y_0^n=0\}$ dans 4.25. Comme $\varphi_\rho(x,y) \leq x+y$, on déduit de 4.21 que $\tilde{H}^{n,\theta\zeta}(\rho)_1 \to 0$ en P_0^n-probabilité.

Il reste donc à montrer 4.14 et 4.15, avec $\hat{H}^{n,\theta\zeta}(\rho)_1$ au lieu de $\bar{H}^{n,\theta\zeta}(\rho)_1$. Pour 4.14, on remarque que

$$\hat{H}^{n,\theta\zeta}(\tfrac{1}{2})_1 = \frac{1}{2}\int_{[0,1]\times E_n} [\sqrt{y_\theta^n/y_0^n} - \sqrt{y_\zeta^n/y_0^n}]^2\ y_0^n\ ds\ F_n(dx).$$

Etant donnés 4.20, 4.23 et la linéarité de $\theta \to D_\theta^n$, on vérifie aisément que la quantité ci-dessus tend en P_0^n-probabilité vers $u(\theta-\zeta)/8 = \|\theta-\zeta\|^2/8$, qui vaut $H^{\theta\zeta}(\frac{1}{2})_1$ par 4.17.

Pour 4.15, on remarque que

$$\hat{H}^{n,0\theta}(\rho)_1 = \int_{[0,1]\times E_n} \varphi_\rho(1, y_\theta^n/y_0^n)\ y_0^n\ ds\ F_n(dx),$$

et que $|\varphi_\rho(1,x) - 4\rho(1-\rho)\varphi_{1/2}(1,x)| \leq K(\sqrt{x} - 1)^2[\varepsilon + 1_{\{|\sqrt{x}-1|\geq\varepsilon\}}]$ pour tout $\varepsilon>0$, et pour une constante $K = K_\rho$. On a alors

4.26 $$|\hat{H}^{n,0\theta}(\rho)_1 - 4\rho(1-\rho)\hat{H}^{n,0\theta}(\tfrac{1}{2})_1|$$
$$\leq K\int_{[0,1]\times E_n} (\sqrt{y_\theta^n/y_0^n} - 1)^2[\varepsilon + 1_{\{|\sqrt{y_\theta^n/y_0^n} - 1|\geq\varepsilon\}}]y_0^n\ ds\ F_n(dx).$$

Mais on a $u^2 1_{\{|u|\geq\varepsilon\}} \leq 4(u-v)^2 + 2v^2 1_{\{|v|\geq\varepsilon/2\}}$ pour tous u,v,ε. En appliquant ceci à $u = \sqrt{y_\theta^n/y_0^n} - 1$ et à $v = D_\theta^n/2c_n$, on voit que le premier membre de 4.26 est majoré par

$$K(2\varepsilon+4)\int_{[0,1]\times E_n} (\sqrt{y_\theta^n/y_0^n} - 1 - D_\theta^n/2c_n)^2\ y_0^n\ ds\ F_n(dx)$$
$$+ \frac{K\varepsilon}{2}\int_{[0,1]\times E_n} (D_\theta^n/c_n)^2\ y_0^n\ ds\ F_n(dx)$$
$$+ \frac{K}{2}\int_{[0,1]\times E_n} (D_\theta^n/c_n)^2\ 1_{\{|D_\theta^n/c_n|>\varepsilon\}}\ y_0^n\ ds\ F_n(dx)$$

pour tout $\varepsilon>0$: on déduit alors de 4.20, 4.22 et 4.23 que le premier membre de 4.26 tend vers 0 en P_0^n-probabilité; étant donné 4.14, cela entraine 4.15, et la preuve est terminée. □

REFERENCES

[1] P.K. ANDERSEN, R.D. GILL: Cox's regression model for counting processes: a large sample study. Ann. Statist. 10, 1100-1120, 1982.

[2] K. DZHAPARIDZE: On asymptotic efficiency of the Cox estimator. In Proc. 1st World Congress Bernoulli Soc. (Yu Prokhorov, W. Sazonov, eds), Vol. 2, 59-61, Science Press: Utrecht, 1987.

[3] P.E. GREENWOOD, W. WEFELMEYER: Efficiency bounds for estimating functionals of stochastic processes. Preprint in Statistics 117, Dept. of Math., Univ. of Cologne, 1989.

[4] P.E. GREENWOOD, W.WEFELMEYER: Efficiency of estimators for partially specified filtered models. Preprint in Statistics 119, Dept. of Math., Univ. of Cologne, 1989.

[5] J.E. HUTTON, P.I. NELSON: Quasi-likelihood estimation for semimartingales. Stoch. Proc. Appl. 22, 245-257, 1986.

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

[7] J. JACOD: Filtered statistical models and Hellinger processes. Stochastic Process. Appl. 32, 3-45, 1989.

[8] J. JACOD: Convergence of filtered statistical models and Hellinger processes. Stochastic Process. Appl. 32, 47-68, 1989.

[9] J. JACOD: Sur le processus de vraisemblance partielle. Prépubl. 11, Laboratoire de Probabilités, Univ. Paris 6, 1989.

[10] A. JAKUBOWSKI, J. MEMIN, G. PAGES: Convergence en loi des suites d'intégrales stochastiques sur l'espace D de Skorokhod. Probab. Theo. Rel. Fields, 81, 111-137, 1989.

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

[12] J. MEMIN: Théorèmes limite fonctionnels pour les processus de vraisemblance (cadre asymptotiquement a-normal). Séminaire de Proba. Rennes 1985. Univ. de Rennes, 1986.

[13] M. SORENSEN: Some asymptotic properties of quasi-likelihood estimators for semimartingales. Res. Rep. 178, Aarhus, 1988.

F. Coquet: Institut de Mathématiques, Université de Rennes, Campus de Beaulieu, 35042-RENNES-Cedex.

J. Jacod: Laboratoire de Probabilités, Université P. et M. Curie, Tour 56, 4 Place Jussieu, 75252-PARIS-Cedex 05.