

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

R. SENOUSI

Problème d'identification dans le modèle de Cox

Annales de l'I. H. P., section B, tome 26, n° 1 (1990), p. 45-64

http://www.numdam.org/item?id=AIHPB_1990__26_1_45_0

© Gauthier-Villars, 1990, tous droits réservés.

L'accès aux archives de la revue « *Annales de l'I. H. P., section B* » (<http://www.elsevier.com/locate/anihpb>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

Problème d'identification dans le modèle de Cox

par

R. SENOUSI

Bât. n° 425, Mathématiques,
Université de Paris-Sud, 91405 Orsay Cedex, France

RÉSUMÉ. — On utilise ici la méthode de la vraisemblance compensée d'Akaïké afin d'estimer presque sûrement la dimension d'un modèle de données de survie censurées. L'étude des informations de Kullback et de Fisher pour la vraisemblance partielle de Cox montre la consistance forte des estimateurs ainsi que leur convergence en loi. On établit ensuite une loi du logarithme itéré pour cette vraisemblance partielle qui conduit à la consistance de l'estimateur d'Akaïké.

Mots clés : vraisemblance partielle de Cox, identification de modèle, loi du logarithme itéré.

ABSTRACT. — We use here Akaïké's likelihood method in order to estimate almost surely the dimension of a survival's statistical model with censored data. The study of Kullback's and Fisher's information for Cox's partial likelihood shows the strong consistency of the estimators and their convergence in distribution. Then we prove an iterated logarithm law for this partial likelihood, which leads to the consistency of Akaïké's estimator.

Classification A.M.S. : 62 F 10, 62 M 07.

INTRODUCTION

Dans l'analyse des données de survie, le modèle de Cox [3] s'avère souvent bien adapté en présence de covariables explicatives Z_1, \dots, Z_q de la survie X . L'intensité y est supposée proportionnelle à $\exp \left\{ \sum_{j=1}^q \theta_j z_j \right\}$ pour un paramètre $\theta = (\theta_1, \dots, \theta_q)$ inconnu et à une intensité fixe $\lambda_0(\cdot)$. Lorsque cette fonction λ_0 est inconnue le modèle est dit semi-paramétrique. Dans la pratique, on dispose de façon générale d'un grand nombre de covariables susceptibles d'expliquer la survie observée et le problème consiste ici à rechercher les « bonnes » covariables, autrement dit à déterminer le modèle statistique où l'on se place, c'est-à-dire déterminer l'espace du paramètre θ . Une abondante littérature démontre le succès d'un tel modèle, et traite les différents problèmes de la consistance, de la convergence en loi et de tests dans l'estimation de θ et de λ_0 ([2], [6], etc.). Andersen et Gill [2], ont montré grâce à la théorie des processus ponctuels comment généraliser ces modèles à des observations de processus ponctuels censurés, tout en l'étendant aux cas des covariables dépendantes du temps. On se place dans le schéma d'étude proposé dans [9] pour répondre à ce problème d'identification de modèle. Notre cadre est un peu plus restreint que celui de [2], on peut le généraliser mais au prix d'hypothèses peu maniables. On précise l'étude asymptotique presque sûre de l'estimateur du maximum de vraisemblance ou de vraisemblance partielle. On vérifie aussi la régularité asymptotique du modèle dans un premier temps par un calcul de quantités d'informations analogues à celles de Kullback et de Fisher. On établit ensuite une loi du logarithme itéré pour l'estimateur. La méthode de la vraisemblance compensée d'Akaïké permet alors de déterminer p. s. l'espace du paramètre. Des simulations dans le cas simple de lois exponentielles illustrent l'intérêt de ces méthodes pour le problème d'identification.

1. MODÈLE STATISTIQUE

On considère ici un type d'observation de temps de survie censuré X courant : pour une filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ et si le paramètre vaut α , le processus $N = (N_t)$, $N_t = \mathbf{1}_{\{X \leq t\}}$ est $(\mathbb{F}, \mathbb{P}_\alpha)$ compensé par le processus prévisible $\tilde{N} = (\tilde{N}_t)$, $\tilde{N}_t(\alpha) = \int_0^t \lambda(\alpha, s) ds$ où l'intensité $\lambda(\alpha, s)$ est de la forme :

$$\lambda(\alpha, s) = Y(s) \lambda_0(s) \exp \{ \alpha' Z(s) \}$$

α' désigne le vecteur ligne $(\alpha^1, \dots, \alpha^q)$ et $\alpha'Z$ le produit scalaire.

Le processus des covariables $Z = (Z(s))$ est à valeurs dans \mathbb{R}^q , adapté à \mathbb{F} , continu à gauche et possédant des limites à droite. Sa loi μ indépendante du paramètre θ est le plus souvent assez bien connue, par exemple bornée ou gaussienne et l'hypothèse suivante n'est donc pas trop restrictive :

C.1. La loi μ de Z vérifie, pour tout $\rho > 0$, $\mathbb{E}_\mu \{ \exp(\rho \|Z\|_\infty) \} < \infty$ où $\|Z\|_\infty = \sup_{s \leq t} \sup_{1 \leq j \leq q} |Z^j(s)|$ et la loi de $Z(X) = (Z^1, \dots, Z^q)(X)$ n'est concentrée sur aucun hyperplan de \mathbb{R}^q .

Le processus de censure $Y = (Y(s))$ est à valeurs dans $\{0, 1\}$, décroissant, continu à gauche et adapté à \mathbb{F} . Dans les cas usuels de censure à droite des données de survie, Y s'écrit $Y(s) = \mathbf{1}_{\{X \geq s\}}$ avec $X = \inf(T, C)$: T est la variable d'intérêt (la survie) et C la variable de censure. Si $C = t_0$ fixe, la censure est dite à temps fixe et si C est indépendante de T la censure est dite à risques compétitifs. Les autres cas de censure classiques peuvent être appréhendés de la même façon (type II, type II progressive, etc.).

L'étude statistique du modèle se fait à partir de l'observation de n copies indépendantes $(X_i, Y_i, Z_i)_{i \geq 1}$ du triplet (X, Y, Z) jusqu'au temps t . La quasi-vraisemblance $\mathcal{V}_n^1(\alpha, t)$ suivante se confond souvent avec la vraisemblance du modèle, mais cela nous importe peu. L'essentiel est qu'elle en joue le rôle lorsque λ_0 est connue. Si $N_i(t) = \mathbf{1}_{\{X_i \leq t\}}$ posons

$$\mathcal{V}_n^1(\alpha, t) = \exp \left\{ \sum_{i=1}^n \left(\int_0^t \alpha' Z_i(s) dN_i(s) + \int_0^t (1 - \exp \{ \alpha' Z_i(s) \}) Y_i(s) \lambda_0(s) ds \right) \right\}.$$

Dans les cas intéressants la fonction λ_0 est inconnue et, pour cela, Cox a introduit la vraisemblance partielle qui évite cette inconnue λ_0 :

$$\mathcal{V}_n^2(\alpha, t) = \exp \left\{ \sum_{i=1}^n \left(\int_0^t \alpha' Z_i(s) dN_i(s) - n \int_0^t \text{Log} \left(\sum_{i=1}^n Y_i(s) \exp \{ \alpha' Z_i(s) \} \right) d\bar{N}^n(s) \right) \right\}$$

où

$$\bar{N}^n(s) = \left(\sum_{i=1}^n N_i(s) \right) / n$$

$\mathcal{V}_n(\alpha, t)$ désignera indifféremment l'une ou l'autre de ces quasi-vraisemblances et v_n leur logarithme. $\hat{\theta}_{n,t}$ sera l'estimateur du maximum de la quasi-vraisemblance \mathcal{V}_n .

Ainsi v_n^1 peut être étudiée sans trop de difficultés dans le cadre des échantillons des modèles réguliers (voir § 1 du [9]), mais il n'en est pas ainsi pour v_n^2 , qui nécessite une étude plus approfondie.

Nous supposons en outre :

$$\text{C.2. } \mathbb{P}_0(Y(t) > 0) > 0 \text{ et } \int_0^t \lambda_0(s) ds < \infty \text{ pour tout } t > 0.$$

Nous adoptons les notations suivantes et rappelons quelques remarques simples et utiles pour la suite. \mathcal{D} indique la dérivation relative au paramètre α .

$$\begin{aligned} \varphi_n(\alpha, s) &= \left(\sum_{i=1}^n Y_i(s) \exp \{ \alpha' Z_i(s) \} \right) / n \\ \Phi_n(\alpha, s) &= \text{Log } \varphi_n(\alpha, s) \\ \psi_n(\alpha, s) &= \mathcal{D} \varphi_n(\alpha, s) = \left(\sum_{i=1}^n Y_i(s) Z_i(s) \exp \{ \alpha' Z_i(s) \} \right) / n \\ \tau_n(\alpha, s) &= \mathcal{D}^2 \varphi_n(\alpha, s) = \left(\sum_{i=1}^n Y_i(s) Z_i(s) Z_i'(s) \exp \{ \alpha' Z_i(s) \} \right) / n. \end{aligned}$$

Remarques. — 1. Si β est une fonction continue de \mathbb{R} dans \mathbb{R}

$$\begin{aligned} M(\theta, t) &= \int_0^t \beta(Z(s)) dN(s) - \int_0^t \beta(Z(s)) \lambda(\theta, s) ds \\ &= \beta(Z(X)) \cdot \mathbf{1}_{\{X \leq t\}} - \int_0^t \beta(Z(s)) \lambda(\theta, s) ds \end{aligned}$$

est une $(\mathbb{F}, \mathbb{P}_\theta)$ martingale de processus croissant

$$\langle M(\theta, t) \rangle = \int_0^t \beta^2(Z(s)) \lambda(\theta, s) ds.$$

2. L'hypothèse C.1 implique que, pour tout $s \leq t$, on a

$$\mathbb{E}_\mu(\|Z(s)\|^k \cdot \exp \{ \alpha' Z(s) \}) \leq \mathbb{E}_\mu(\|Z\|_\infty^k \cdot \exp \{ \|\alpha\| \cdot \|Z\|_\infty \}) \leq \infty.$$

La fonction $\alpha \mapsto \mathbb{E}_\mu(\|Z(s)\|^k \cdot \exp \{ \alpha' Z(s) \})$ est donc finie, continue pour tout k , et on peut dériver sous le signe intégral autant de fois que l'on désire la fonction $\alpha \mapsto \mathbb{E}_\mu(\exp \{ \alpha' Z(s) \})$.

Par ailleurs la même condition C.1 suffit pour appliquer le théorème de R. Rao [7], de la loi forte des grands nombres dans $D[0, 1]$ l'ensemble des fonctions c.à.d. et l.à.g., pour la suite de processus $(\varphi, \psi, \tau)_n$ c'est-à-dire: $\sup_{s \leq t} \|\pi_n(\alpha, s) - \tau(\theta, \alpha, s)\| \rightarrow 0$ \mathbb{P}_θ p. s. quand $n \rightarrow \infty$ pour $\pi_n = \varphi_n$ et $\pi = \varphi$, (resp. $\pi_n = \psi_n$ et $\pi = \psi$), (resp. $\pi_n = \tau_n$ et $\pi = \tau$).

3. Pour tout $s \leq t$, on a $\varphi_n(\alpha, s) \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} \varphi(\theta, \alpha, s)$ avec $\varphi(\theta, \alpha, s) = \mathbb{E}_\theta$

$(Y(s) \cdot \exp\{\alpha' Z(s)\}) = \int_{\mathbb{R}^k} S(\theta, s, z) \cdot \exp\{\alpha' z\} d\mu_s(z)$ où μ_s est la loi image de $Z(s)$ et $S(\theta, s, Z) = 1 - \mathbb{E}_\theta(\mathbf{1}_{\{X < s\}}/Z(s))$. La fonction $\alpha \mapsto \varphi(\theta, \alpha, s)$ est la transformée de Fourier de la mesure $S(\theta, s, z) d\mu_s(z)$. Sous les conditions C.1 et C.2, la fonction $\alpha \mapsto \Phi(\theta, \alpha, s) = \text{Log } \varphi(\theta, \alpha, s)$ est strictement convexe et la fonction $\alpha \mapsto \varphi(\theta, \alpha, s)$ admet une dérivée $\varphi^{(k)}(\theta, \alpha, s)$ pour tout k qui vérifie l'inégalité

$$\|\varphi^{(k)}(\theta, \alpha, s)\| \leq \mathbb{E}_\mu(\|Z\|_\infty^k \exp\{\|\alpha\| \cdot \|Z\|_\infty\}).$$

2. INFORMATION DE KULLBACK ET CONSISTENCE FORTE POUR LE MODÈLE DE COX

Sous des hypothèses légèrement plus faibles, l'étude de la consistance faible et de la normalité asymptotique des estimateurs de Cox est effectuée dans [2]. Nous précisons ici les pertes d'informations par rapport aux estimateurs du maximum de vraisemblance et les comportements presque sûrs.

THÉORÈME 1. — *Sous les hypothèses C.1 et C.2, on a*

1. *Pour tout $\alpha \in \mathbb{R}^q$.*

$$\mathcal{K}_n(\theta, \alpha, t) = (v_n(\theta, t) - v_n(\alpha, t))/n \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} K(\theta, \alpha, t) < \infty$$

où

$$K^1(\theta, \alpha, t) = \mathbb{E}_\theta \left(\int_0^t (\exp\{(\alpha - \theta)' Z(s)\} - (\alpha - \theta)' Z(s) - 1) \lambda(\theta, s) ds \right)$$

$$K^2(\theta, \alpha, t) = \mathbb{E}_\theta \left(\int_0^t \left(\text{Log} \left(\frac{\mathbb{E}_\theta(Y(s) \exp\{\alpha' Z(s)\})}{\mathbb{E}_\theta(Y(s) \exp\{\theta' Z(s)\})} \right) - (\alpha - \theta)' Z(s) \right) \lambda(\theta, s) ds \right).$$

2. *L'information $t \mapsto K(\theta, \alpha, t)$ est croissante, nulle en zéro, et la perte d'information $t \mapsto (K^1 - K^2)(\theta, \alpha, t)$ est aussi croissante. De plus, pour tout $t > 0$, $\alpha \mapsto K(\theta, \alpha, t)$ est strictement convexe sur \mathbb{R}^q et vérifie, $K(\theta, \alpha, t) = 0$ seulement pour $\theta = \alpha$.*

3. *Enfin $\hat{\theta}_{n,t} \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} \theta$.*

Preuve. — 1. Dans le cas de v_n^1 :

$$v_n^1(\theta, t) - v_n^1(\alpha, t) = \sum_{i=1}^n \left((\theta - \alpha)' Z_i(X_i) \mathbf{1}_{\{X_i \leq t\}} + \int_0^t (\exp\{(\alpha - \theta)' Z_i(s)\} - 1) \lambda_i(\theta, s) ds \right)$$

C'est une somme de v. a., i.i.d. et il suffit de montrer qu'elles sont \mathbb{P}_θ -intégrables. Des remarques précédentes, on tire

$$\begin{aligned} \mathbb{E}_\theta \left((\theta - \alpha)' Z(X) \cdot \mathbf{1}_{\{X \leq t\}} + \int_0^t (\exp\{(\alpha - \theta)' Z(s)\} - 1) \lambda(\theta, s) ds \right) \\ \leq \mathbb{E}_\theta (\exp\{\|\alpha - \theta\| \cdot \|Z\|_\infty\} + \|\alpha - \theta\| \cdot \|Z\|_\infty + 1) < \infty. \end{aligned}$$

En ce qui concerne la vraisemblance partielle, on a :

$$\begin{aligned} \frac{1}{n} (v_n^2(\theta, t) - v_n^2(\alpha, t)) \\ = \frac{1}{n} \left(\sum_{i=1}^n (\theta - \alpha)' Z_i(X_i) \mathbf{1}_{\{X_i \leq t\}} \right) - \int_0^t \text{Log} \left(\frac{\Phi_n(\alpha, s)}{\Phi_n(\theta, s)} \right) d\bar{N}(s). \end{aligned}$$

Le premier terme converge vers $\mathbb{E}_\theta((\theta - \alpha)' Z(X) \mathbf{1}_{\{X \leq t\}})$.

Le second terme nécessite plus d'attention. La loi forte [7] dans $D[0, t]$ assure la convergence uniforme sur $[0, t]$

$$\varphi_n(\alpha, s) \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} \varphi(\theta, \alpha, s) = \mathbb{E}_\theta(\exp\{\alpha' Z(s)\} \cdot \mathbf{1}_{\{X \geq s\}}).$$

D'autre part $\mathbb{C}.2$ implique que $\varphi(\theta, \alpha, s) > 0$ p. s. sur $[0, t]$, et alors

$\Phi_n(\alpha, s) \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} \Phi(\theta, \alpha, s)$ uniformément sur $[0, t]$. Le théorème de Glivenko-

Cantelli montre que $\bar{N}^n(t) \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} \mathbb{E}_\theta \left(\int_0^t \lambda(\theta, s) ds \right)$ uniformément sur \mathbb{R}_+ et

$$\mathbb{E}_\theta \left(\int_0^t \lambda(\theta, s) ds \right) = \int_0^t \mathbb{E}_\theta(\lambda(\theta, s)) ds$$

d'où

$$\begin{aligned} \int_0^t \text{Log} \left(\frac{\Phi_n(\alpha, s)}{\Phi_n(\theta, s)} \right) d\bar{N}^n(s) \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} \int_0^t \text{Log} \left(\frac{\varphi(\theta, \alpha, s)}{\varphi(\theta, \theta, s)} \right) \mathbb{E}_\theta(\lambda(\theta, s)) ds \\ 2. \mathbf{K}^1(\theta, \alpha, t) = \mathbb{E}(\{\exp\{(\alpha - \theta)' Z(X)\} - (\alpha - \theta)' Z(X) - 1\} \cdot \mathbf{1}_{\{X \geq s\}}) \end{aligned}$$

mais $e^x - x - 1 > 0$ sauf pour $x = 0$, $t \mapsto \mathbf{K}^1(\theta, \alpha, t)$ est positive, croissante, nulle en 0, et ne s'annule sous $\mathbb{C}.1$ et $\mathbb{C}.2$ que pour $(\alpha - \theta)' Z(X) = 0$ μ p.

autrement dit pour $\alpha = \theta$

$$K^2(\theta, \alpha, t) = \int_0^t ((\Phi(\theta, \alpha, s) - \Phi(\theta, \theta, s)) \mathbb{E}_\theta(\lambda(\theta, s)) - (\alpha - \theta)' \mathbb{E}_\theta(Z(s) \lambda(\theta, s))) ds.$$

D'après la remarque 3, $\alpha \mapsto \Phi(\theta, \alpha, s)$ est strictement convexe, d'où, pour $\alpha \neq \theta$,

$$\Phi(\theta, \alpha, s) - \Phi(\theta, \theta, s) > (\alpha - \theta)' \left(\frac{\mathcal{D}\Phi}{\Phi} \right) (\theta, \theta, s)$$

et

$$\frac{\mathcal{D}\Phi}{\Phi}(\theta, \theta, s) = \frac{\mathbb{E}_\theta(Y(s) \cdot Z(s) \cdot \exp\{\theta' Z(s)\}) \lambda_0(s)}{\mathbb{E}_\theta(Y(s) \cdot \exp\{\theta' Z(s)\}) \lambda_0(s)} = \frac{\mathbb{E}_\theta(Z(s) \cdot \lambda(\theta, s))}{\mathbb{E}_\theta(\lambda(\theta, s))}$$

c.à.d.

$$(\Phi(\theta, \alpha, s) - \Phi(\theta, \theta, s)) \cdot \varphi(\theta, \theta, s) > (\alpha - \theta)' \cdot \mathcal{D}\Phi(\theta, \theta, s).$$

Ceci assure le résultat désiré: $t \mapsto K^2(\theta, \alpha, t)$ est positive, croissante ne s'annulant que pour $\alpha = \theta$.

La convexité stricte de la fonction $x \mapsto e^x - 1 - x$ et la linéarité de l'opérateur \mathbb{E}_θ impliquent la convexité stricte de la fonction $\alpha \mapsto K^1(\theta, \alpha, t)$ et la convexité stricte de $\alpha \mapsto \Phi(\theta, \alpha, t)$ entraîne celle de $\alpha \mapsto K^2(\theta, \alpha, t)$.

$$(K^1 - K^2)(\theta, \alpha, t) = \mathbb{E}_\theta \left(\int_0^t \left\{ \exp\{(\alpha - \theta)' Z(s)\} - \text{Log} \left(\frac{\varphi(\theta, \alpha, s)}{\varphi(\theta, \theta, s)} \right) - 1 \right\} \lambda(\theta, s) ds \right)$$

mais

$$\begin{aligned} \mathbb{E}_\theta \left(\int_0^t (\exp\{(\alpha - \theta)' Z(s)\} - 1) \lambda(\theta, s) ds \right) \\ = \mathbb{E}_\theta \left(\int_0^t (\exp\{\alpha' Z(s)\} - \exp\{\theta' Z(s)\}) Y(s) \lambda_0(s) ds \right) \\ = \int_0^t (\varphi(\theta, \alpha, s) - \varphi(\theta, \theta, s)) ds \end{aligned}$$

c.à.d.

$$(K^1 - K^2)(\theta, \alpha, t) = \int_0^t \left(\left(\frac{\varphi(\theta, \alpha, s)}{\varphi(\theta, \theta, s)} \right) - 1 - \text{Log} \left(\frac{\varphi(\theta, \alpha, s)}{\varphi(\theta, \theta, s)} \right) \right) \varphi(\theta, \theta, s) ds$$

comme perte d'information. Elle est positive, croissante en t et nulle seulement pour le cas trivial car $x - 1 - \text{Log } x > 0$ sauf pour $x = 0$.

3. Si $\hat{\theta}_{n,t}$ réalise le minimum de la fonction convexe $\mathcal{K}_n(\theta, \alpha, t)$ pour tout n , on peut utiliser un corollaire [9] d'un théorème de Rockafellar sur la convergence uniforme d'une suite de fonctions convexes. Outre la convexité de la limite $K(\theta, \alpha, t)$, le corollaire assure la convergence p. s. de la suite $\hat{\theta}_{n,t}$ vers θ point minimum unique de la fonction $\alpha \mapsto K(\theta, \alpha, t)$.

3. ÉTUDE DES VITESSES DE CONVERGENCE

On complète dans cette partie l'étude de la convergence en loi de l'estimateur du maximum des quasi-vraisemblances en détaillant pour cela le calcul des informations de Fisher, fort utile pour établir une loi du logarithme itéré pour l'estimateur. Soient

$$\begin{aligned} \mathcal{D} v_n^1(\alpha, t) &= \sum_{i=1}^n \left(Z_i(X_i) \mathbf{1}_{\{X_i \leq t\}} - \int_0^t Z_i(s) \exp\{\alpha' Z_i(s)\} Y_i(s) \lambda_0(s) ds \right) \\ \mathcal{D}^2 v_n^1(\alpha, t) &= - \sum_{i=1}^n \left(\int_0^t Z_i(s) Z_i'(s) \cdot \lambda_i(\alpha, s) ds \right) \\ \mathcal{D} v_n^2(\alpha, t) &= \sum_{i=1}^n \left(\int_0^t \left(Z_i(s) - \frac{\Psi_n(\alpha, s)}{\Phi_n} \right) dN_i(s) \right) \\ \mathcal{D}^2 v_n^2(\alpha, t) &= - \sum_{i=1}^n \left(\int_0^t \left(\frac{\tau_n(\alpha, s)}{\Phi_n} - \left(\frac{\Psi_n}{\Phi_n} \right) \cdot \left(\frac{\Psi_n}{\Phi_n} \right)'(\alpha, s) \right) dN_i(s) \right) \end{aligned}$$

THÉORÈME 2. — *Sous les mêmes hypothèses C.1 et C.2, on a*

1. *Pour tout n et t , $\alpha \mapsto v_n(\alpha, t)$ est de classe C^2 sur \mathbb{R}^q .*
2. *Il existe une matrice inversible $\mathbb{I}(\theta, t)$ telle que, pour toute suite de v. a. $(\theta_n^*)_{n \geq 1}$ convergente vers θ , \mathbb{P}_θ p. s., on ait*

$$(\mathcal{D}^2 v_n(\theta_n^*, t))/n \xrightarrow[n \uparrow \infty]{\text{P}_\theta \text{ p. s.}} -\mathbb{I}(\theta, t)$$

où

$$\begin{aligned} \mathbb{I}^1(\theta, t) &= \int_0^t \mathbb{E}_\theta(Z(s) Z'(s) \lambda(\theta, s)) ds \\ \mathbb{I}^2(\theta, t) &= \int_0^t \left(\frac{\mathbb{E}_\theta(ZZ'(s) \lambda(\theta, s))}{\mathbb{E}_\theta(\lambda(\theta, s))} - \frac{\mathbb{E}_\theta(Z(s) \lambda(\theta, s))}{\mathbb{E}_\theta(\lambda(\theta, s))} \right. \\ &\quad \left. \times \frac{\mathbb{E}_\theta(Z'(s) \lambda(\theta, s))}{\mathbb{E}_\theta(\lambda(\theta, s))} \right) \mathbb{E}_\theta(\lambda(\theta, s)) ds. \end{aligned}$$

De plus la fonction perte d'information $t \mapsto (\mathbb{I}^1 - \mathbb{I}^2)(\theta, t)$ est croissante, positive.

3. $\mathcal{D} v_n(\theta, t) / \sqrt{n} \xrightarrow[n \uparrow \infty]{\mathcal{L}(\mathbb{P}_\theta)} \mathbf{N}(0, \mathbb{I}(\theta, t)).$

4. La famille $\left(\frac{\mathbb{I}^{-1/2}(\theta, t)}{\sqrt{2n \text{Log Log } n}} \mathcal{D} v_n(\theta, t) \right)_{n \geq 3}$ est \mathbb{P}_θ p. s. relativement

compacte et a pour ensemble de valeurs d'adhérence la boule unité de \mathbb{R}^q .

Démonstration. — 1. est évidente.

2. Considérons les informations de Fisher empiriques.

$$\mathbb{I}_n^1(\theta, t) = -(\mathcal{D}^2 v_n^1(\theta, t))/n \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} \mathbb{E}_\theta \left(\int_0^t \mathbf{Z}(s) \cdot \mathbf{Z}'(s) \lambda(\theta, s) ds \right)$$

et

$$\mathbb{I}_n^2(\theta, t) = -(\mathcal{D}^2 v_n^2(\theta, t))/n.$$

Les remarques sur les conditions C.1 et C.2 permettent de conclure que la convergence p. s. est uniforme sur tout intervalle $[0, t]$ pour les processus :

$$\psi_n(\alpha, s) \lambda_0(s) \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} \psi(\theta, \alpha, s) \lambda_0(s) = \mathbb{E}_\theta(\mathbf{Z}(s) \lambda(\alpha, s))$$

$$\tau_n(\alpha, s) \lambda_0(s) \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} \tau(\theta, \alpha, s) \lambda_0(s) = \mathbb{E}_\theta(\mathbf{Z}(s) \mathbf{Z}'(s) \lambda(\alpha, s)).$$

On en déduit la forme de $\mathbb{I}^2(\theta, t)$.

D'un autre côté, pour la quasi-vraisemblance $v_n^1(\alpha, t)$ si $B(\theta, \eta)$ désigne la boule centrée en θ et de rayon η , on a

$$\sup_{\alpha \in B(\theta, \eta)} |\exp\{\alpha' \mathbf{Z}(s)\} - \exp\{\theta' \mathbf{Z}(s)\}| \leq \eta \cdot \|\mathbf{Z}\|_\infty \cdot \exp\{\eta \|\mathbf{Z}\|_\infty\} \cdot \exp\{\theta' \mathbf{Z}(s)\}$$

d'où

$$\sup_{\|\alpha - \theta\| \leq \eta} \|\mathbb{I}_n^1(\alpha, t) - \mathbb{I}_n^1(\theta, t)\| \leq \frac{1}{n} \left(\sum_{i=1}^n g_\eta(\|\mathbf{Z}_i\|_\infty, Y_i, X_i) \right) \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} g(\eta)$$

avec $g_\eta(\|\mathbf{Z}_i\|_\infty, X_i, Y_i) = \eta \cdot \|\mathbf{Z}_i\|_\infty^3 \cdot \exp\{\eta \|\mathbf{Z}_i\|_\infty\} \cdot \tilde{N}_i(\theta, t)$ intégrable sous C.1 c.à.d. $\lim_{\eta \rightarrow 0} g(\eta) = 0$ et ceci implique la seconde partie de 2 pour v_n^1 . Il

est plus délicat d'obtenir le même résultat pour v_n^2 :

$$\begin{aligned} \sup_{\|\alpha - \theta\| \leq \eta} \|\mathbb{I}_n^2(\alpha, t) - \mathbb{I}_n^2(\theta, t)\| &\leq \int_0^t \sup_{\|\alpha - \theta\| \leq \eta} \left\| \frac{\tau_n(\alpha, s) - \tau_n(\theta, s)}{\Phi_n} \right\| d\bar{N}^n(s) \\ &+ \int_0^t \sup_{\|\alpha - \theta\| \leq \eta} \left\| \frac{\Psi_n}{\Phi_n} \cdot \frac{\Psi_n'(\alpha, s) - \Psi_n'(\theta, s)}{\Phi_n} \right\| d\bar{N}^n(s) \end{aligned}$$

Le premier terme se décompose en

$$\begin{aligned} & \left\| \frac{\tau_n(\alpha, s) - \tau_n(\theta, s)}{\varphi_n} \right\| \\ & \leq \frac{\|\tau_n(\alpha, s) - \tau_n(\theta, s)\|}{\varphi_n(\theta, s)} \cdot \frac{|\varphi_n(\alpha, s) - \varphi_n(\theta, s)|}{\varphi_n(\theta, s)} \cdot \frac{\varphi_n(\theta, s)}{\varphi_n(\alpha, s)} \\ & + \frac{\|\tau_n(\alpha, s) - \tau_n(\theta, s)\|}{\varphi_n(\theta, s)} + \frac{\|\tau_n(\theta, s)\|}{\varphi_n(\theta, s)} \cdot \frac{|\varphi_n(\alpha, s) - \varphi_n(\theta, s)|}{\varphi_n(\theta, s)} \cdot \frac{\varphi_n(\theta, s)}{\varphi_n(\alpha, s)} \end{aligned}$$

et le second en :

$$\begin{aligned} & \left\| \frac{\Psi_n}{\varphi_n} \cdot \frac{\Psi'_n(\alpha, s)}{\varphi_n} - \frac{\Psi_n}{\varphi_n} \cdot \frac{\Psi'_n(\theta, s)}{\varphi_n} \right\| \leq \left\| \frac{\Psi_n(\alpha, s) - \Psi_n(\theta, s)}{\varphi_n} \right\|^2 \\ & + 2 \left\| \frac{\Psi_n(\theta, s)}{\varphi_n} \right\| \cdot \left\| \frac{\Psi_n(\alpha, s) - \Psi_n(\theta, s)}{\varphi_n} \right\|. \end{aligned}$$

Il suffit de majorer ensuite les expressions suivantes :

$$\frac{\|\tau_n(\alpha, s) - \tau_n(\theta, s)\|}{\varphi_n(\theta, s)/\varphi_n(\alpha, s)} \quad \text{et} \quad \left\| (\Psi_n/\varphi_n)(\alpha, s) - (\Psi_n/\varphi_n)(\theta, s) \right\|.$$

Par exemple

$$\begin{aligned} & \sup_{\|\alpha - \theta\| \leq \eta} \|\tau_n(\alpha, s) - \tau_n(\theta, s)\| \\ & \leq \frac{\eta}{n} \cdot \sum_{i=1}^n \|Z_i\|_\infty^3 \exp\{\eta \|Z_i\|_\infty\} \cdot Y_i(s) \cdot \exp\{\theta' Z_i(s)\} \\ & \sup_{\|\alpha - \theta\| \leq \eta} \|\Psi_n(\alpha, s) - \Psi_n(\theta, s)\| \\ & \leq \frac{\eta}{n} \cdot \sum_{i=1}^n \|Z_i\|_\infty^2 \exp\{\eta \|Z_i\|_\infty\} \cdot Y_i(s) \cdot \exp\{\theta' Z_i(s)\} \end{aligned}$$

il en est de même pour $\sup_{\|\alpha - \theta\| \leq \eta} \|\varphi_n(\alpha, s) - \varphi_n(\theta, s)\|$ et on peut majorer

$$\frac{\varphi_n(\theta, s)}{\varphi_n(\alpha, s)} \leq \sum_{i=1}^n \exp\{\eta \|Z_i\|_\infty\} \cdot \frac{Y_i(s) \exp\{\theta' Z_i(s)\}}{\varphi_n(\theta, s)}$$

par convexité de la fonction $x \mapsto 1/x$. Traitons pour l'exemple le cas suivant :

$$\begin{aligned} A &= \int_0^t \left(\frac{\|\tau_n(\alpha, s) - \tau_n(\theta, s)\|}{\varphi_n(\theta, s)} \cdot \frac{|\varphi_n(\alpha, s) - \varphi_n(\theta, s)|}{\varphi_n(\theta, s)} \cdot \frac{\varphi_n(\theta, s)}{\varphi_n(\alpha, s)} \right) d\bar{N}(s) \\ &\leq \int_0^t \pi_n(s) d\bar{N}(s) \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} \eta^2 \int_0^t \prod_{j=0,1,3} \left(\frac{\mathbb{E}_\theta(h_j(Z) \lambda(\theta, s))}{\mathbb{E}_\theta(\lambda(\theta, s))} \right) \mathbb{E}_\theta(\lambda(\theta, s)) ds \\ &\leq \eta^2 \cdot \prod_{j=0,1,3} \left(\int_0^t \left(\frac{\mathbb{E}_\theta(h_j(Z) \lambda(\theta, s))}{\mathbb{E}_\theta(\lambda(\theta, s))} \right)^3 \mathbb{E}_\theta(\lambda(\theta, s)) ds \right)^{1/3} \end{aligned}$$

obtenue par inégalité de Hölder avec $h_j(Z) = \|Z\|_j \cdot \exp\{\eta \|Z\|_\infty\}$.

Si on prend pour tout s la loi de probabilité

$$d\mathbb{Q}_{s,\theta} = \frac{\lambda(\theta, s)}{\mathbb{E}_\theta(\lambda(\theta, s))} \cdot d\mathbb{P}_\theta$$

on a

$$\left(\frac{\mathbb{E}_\theta(h_j(Z) \lambda(\theta, s))}{\mathbb{E}_\theta(\lambda(\theta, s))} \right)^3 = (\mathbb{E}_{\mathbb{Q}_{s,\theta}}(h_j(Z)))^3 \leq \mathbb{E}_{\mathbb{Q}_{s,\theta}}(h_j(Z))^3$$

alors

$$A^3 \leq \prod_{j=0,1,3} \left(\int_0^t \mathbb{E}_\theta(h_j^3(Z) \lambda(\theta, s)) ds \right) \leq \prod_{j=0,1,3} (\mathbb{E}_\mu(\|Z\|_\infty^j \exp\{\eta \|Z\|_\infty\}))$$

et il en est ainsi des autres termes. Cela permet enfin d'affirmer que pour

toute suite θ_n^* , si $\theta_n^* \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} \theta$ on a $\mathbb{I}_n(\theta_n^*, t) \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} \mathbb{I}(\theta, t)$. La positivité de $\mathbb{I}(\theta, t)$

résulte de :

$$\begin{aligned} \mathbb{I}^1(\theta, t) &= \int_0^t \mathbb{E}_\theta(Z(s) Z'(s) \lambda(\theta, s)) ds = \int_0^t \mathbb{E}_{\mathbb{Q}_{s,\theta}}(Z(s) Z'(s)) \cdot \mathbb{E}_\theta(\lambda(\theta, s)) ds \\ \mathbb{I}^2(\theta, t) &= \int_0^t \mathbb{E}_{\mathbb{Q}_{s,\theta}}(\tilde{Z}(s) \cdot \tilde{Z}'(s)) \mathbb{E}_\theta(\lambda(\theta, s)) ds \end{aligned}$$

où

$$\tilde{Z} = Z - \mathbb{E}_{\mathbb{Q}_{s,\theta}}(Z(s))$$

De plus

$$(\mathbb{I}^1 - \mathbb{I}^2)(\theta, t) = \int_0^t \mathbb{E}_{\mathbb{Q}_{s,\theta}}(Z(s)) \cdot \mathbb{E}_{\mathbb{Q}_{s,\theta}}(Z'(s)) ds$$

qui est positive et croissante en t .

3. et 4. : Pour la vraisemblance, le résultat est immédiat puisque $\mathcal{D}^1 v_n(\theta, t)$ est une somme de v.a. i.i.d. de carré intégrable. Et pour la

vraisemblance partielle de Cox, on peut décomposer $\mathcal{D} v_n^2$:

$$(\star) \quad \mathcal{D} v_n^2(\theta, t) = \sum_{i=1}^n \int_0^t (Z_i(s) - \frac{\Psi}{\Phi}(\theta, s)) dM_i(\theta, s) \\ + \sum_{i=1}^n \int_0^t \left(\frac{\Psi}{\Phi} - \frac{\Psi_n}{\Phi_n} \right) (\theta, s) dM_i(\theta, s)$$

où $M_i(\theta, s) = N_i(s) - \tilde{N}_i(\theta, s)$, $i=1, \dots, n$ sont des $(\mathbb{F}, \mathbb{P}_\theta)$ martingales orthogonales deux à deux, de processus croissant

$$\langle M_i, M_j \rangle(\theta, s) = \delta_{ij} \tilde{N}_i(\theta, s).$$

Le premier terme est une somme de v. a. i.i.d. de covariance $\mathbb{1}^2(\theta, t)$ et 3 en découle si on montre que :

$$W_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^t \left(\frac{\Psi}{\Phi} - \frac{\Psi_n}{\Phi_n} \right) (\theta, s) dM_i(\theta, s) \xrightarrow[n \uparrow \infty]{} 0 \text{ en } \mathbb{P}_\theta \text{ probabilité.}$$

Or $W_n(t)$ est une martingale localement de carré intégrable, de processus croissant

$$\langle W_n \rangle(t) \leq \sup_{s \leq t} \left\| \frac{\Psi}{\Phi} - \frac{\Psi_n}{\Phi_n} \right\|^2 (\theta, s) \left(\sum_{i=1}^n \tilde{N}_i(\theta, t) / n \right) \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} 0$$

ce qui entraîne la convergence en probabilité de W_n vers 0. Si on reprend la décomposition (\star) de $\mathcal{D}^2 v_n$, on a d'après l'appendice :

$$\sup_{s \leq t} \left\| \frac{\Psi}{\Phi} - \frac{\Psi_n}{\Phi_n} \right\|^2 (\theta, s) \sum_{i=1}^n (M_i(\theta, t) / \sqrt{2n \text{Log Log } n}) \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} 0.$$

COROLLAIRE. — La suite $\left(\frac{\sqrt{n} \mathbb{1}^{1/2}(\theta, t)}{\sqrt{2 \text{Log Log } n}} (\theta - \hat{\theta}_{n,i}) \right)_{n \geq 3}$ est p. s. relativement compacte dans \mathbb{R}^q avec pour ensemble de valeurs d'adhérence la boule unité $B_q(0, 1)$.

Preuve. — C'est une conséquence directe du théorème 1.1 de [9].

4. IDENTIFICATION DU MODÈLE STATISTIQUE

4.1. Tests des rapports de quasi-vraisemblance

Des théorèmes précédents et du théorème II.1 de [9], on tire les résultats qui suivent. Si le premier résultat est classique, le second semble récent, il

peut servir à construire des tests d'hypothèses de type « $\theta \in \varphi(\Lambda)$ » contre « $\theta \notin \varphi(\Lambda)$ » asymptotiquement de niveau 0 et de puissance 1.

THÉORÈME 3. — Soit $r \leq q$, $\Lambda \subset \mathbb{R}^r$ et $\Theta \subset \mathbb{R}^q$ des voisinages respectifs de λ et θ et soit $\varphi: \Lambda \rightarrow \Theta$ de classe C^2 et de rang r au voisinage de λ avec $\varphi(\lambda) = \theta$. Alors sous C. 1 et C. 2, on a pour $w_n(\lambda, t) = v_n(\varphi(\lambda), t)$

1. $2(v_n(\hat{\theta}_{n,i}; t) - w_n(\hat{\lambda}_{n,i}; t)) \xrightarrow[n \uparrow \infty]{\mathcal{L}(\mathbb{P}_\theta)} \chi^2(q-r)$.
2. $\limsup_{n \uparrow \infty} \frac{v_n(\hat{\theta}_{n,i}; t) - w_n(\hat{\lambda}_{n,i}; t)}{\text{Log Log } n} = 1 \mathbb{P}_\theta p. s.$

4.2. Quasi-vraisemblance compensée d'Akaïké

On reprend les notations de [9] et on suppose que la dimension de l'espace du paramètre θ est majorée par une valeur maximale q que le statisticien s'autorise. Soit par ailleurs la fonction

$$\delta: \mathbb{R}^q \rightarrow \Delta = \{0, 1\}^q$$

telle que si $\alpha = (\alpha_1, \dots, \alpha_q) \in \mathbb{R}^q$ on ait $\delta(\alpha) = (\delta_1, \dots, \delta_q)$ avec $\delta_j = \mathbf{1}_{\{\alpha_j = 0\}}$ auquel on associe la relation d'ordre partielle $\delta < \delta'$ si et ssi $\delta_j > 0 \Rightarrow \delta'_j > 0$.

On note $|\delta(\alpha)| = \sum_{j=1}^q \delta_j$ la longueur du paramètre et on considère une suite

de mesures $(c_n)_{n \geq 1}$ de la longueur du paramètre de la forme $c_n: \mathbb{N} \rightarrow \mathbb{R}^+$ où $c_n(r) = h(r)$. $c^*(n)$ vérifie la condition :

$$\text{C. 3. } \lim_{n \uparrow \infty} c^*(n)/n = 0, \liminf_{n \uparrow \infty} c^*(n)/\text{Log Log } n > 1, h(r) - h(r-1) \geq 1.$$

Soit Θ_δ une partie convexe du sous-espace vectoriel \mathbb{R}^δ de \mathbb{R}^q $\mathbb{R}^\delta = \{\alpha \in \mathbb{R}^q: \delta(\alpha) = \delta\}$; on suppose que $\theta \in \Theta_{\delta_0}$ avec $|\delta_0| = r_0 \leq q$.

Le principe d'Akaïké de parcimonie du modèle consiste à maximiser la quasi-vraisemblance v_n tout en minimisant la longueur $|\delta|$ du paramètre. Un estimateur du maximum de la quasi-vraisemblance compensée, sera, si on omet ici l'indice t pour alléger l'écriture, une paire $(\hat{\delta}_n, \hat{\theta}_n)$ qui vérifie

$$W_n(\hat{\theta}_n) = \sup_{\delta \in \Delta} \sup_{\alpha \in \Theta_\delta} W_n(\alpha)$$

avec

$$W_n(\alpha) = v_n(\alpha) - c_n(|\delta(\alpha)|).$$

Le théorème qui suit, montre que la quasi-vraisemblance compensée permet de sélectionner ou d'identifier p.s. le « bon » modèle dès que la suite $c_n(1)$ évolue strictement entre $\text{Log Log } n$ et n . Il découle directement du théorème II.2 de [9], dont le corollaire non rapporté ici peut aussi s'avérer utile.

TABLEAU DE w_n^δ

n = 100					
c_n^*	δ_A	δ_B	δ_C	δ_0	δ_E
$n \dots \dots \dots$	(-176,28)*	-268,88	-346,93	-429,11	-529,02
$\sqrt{n} \dots \dots \dots$	-	-178,88	-166,93	(-159,11)*	-169,02
$\text{Log } n \dots \dots \dots$	-	-173,48	-156,14	(-142,93)*	-147,44
$\text{Log Log } n \dots \dots \dots$	-	-170,40	-149,98	(-133,70)*	-135,13
n = 500					
c_n^*	δ_A	δ_B	δ_C	δ_0	δ_E
$n \dots \dots \dots$	(-1350,9)*	-1794,2	-2214,9	-2632,3	-3132,3
$\sqrt{n} \dots \dots \dots$	-	-1316,5	-1259,6	(-1199,4)*	-1221,7
$\text{Log } n \dots \dots \dots$	-	-1300,4	-1227,3	(-1150,9)*	-1157,1
$\text{Log Log } n \dots \dots \dots$	-	-1296,0	-1218,5	(-1137,8)*	-1139,6
n = 1000					
c_n^*	δ_A	δ_B	δ_C	δ_0	δ_E
$n \dots \dots \dots$	(-3120,3)*	-4019,5	-4891,3	-5706,2	-6707,0
$\sqrt{n} \dots \dots \dots$	-	-3051,1	-2954,5	(-2801,0)*	-2832,5
$\text{Log } n \dots \dots \dots$	-	-3026,4	-2905,1	(-2726,9)*	-2733,6
$\text{Log Log } n \dots \dots \dots$	-	-3021,4	-2895,1	(-2711,9)*	-2713,9

() * indique la valeur maximale de w_n^δ .

THÉORÈME 4. — Sous C. 1, C. 2 et C. 3

$$(\hat{\delta}_n, \hat{\theta}_n, \hat{\delta}_n) \xrightarrow[n \uparrow \infty]{\mathbb{P}_\theta \text{ p. s.}} (\delta_0, \theta).$$

Preuve. — Il suffit de montrer que pour tout $\delta \in \Delta$, $K(\theta, \alpha, t)$ possède un minimum unique $\alpha_\delta \in \Theta_\delta$. Or $K(\theta, \alpha, t)$ est strictement convexe, positive, possède un élément minimum unique sur \mathbb{R}^q et vérifie par conséquence $\lim_{\|\alpha\| \rightarrow \infty} K(\theta, \alpha, t) = +\infty$. Elle ne peut donc avoir que des minima stricts sur tout sous-espace. Ce qui signifie que l'hypothèse H. 4 de [9] est vérifiée, C.Q.F.D.

4.3. Exemple de simulation pour la détermination du modèle

On simule dans le cas d'une censure à risques compétitifs un 1000-échantillon $(c_i, Z_i = (Z_i^1, \dots, Z_i^4), T_i); i = 1, \dots, 1000$ où $C \sim \mathcal{E}(1)$ loi exponentielle et $Z \sim N(0, \mathbb{I}_4)$, \mathbb{I}_q étant la matrice identité de \mathbb{R}^q , et telle que la loi de T conditionnellement à Z soit $\mathcal{E}(\theta' Z)$ avec $\theta = (1, 1, 1, 0)$ c.à.d. $\delta_0 = (1, 1, 1, 0)$.

Identification du modèle pour différentes vitesses c_n^* :

Soient

$$w_n^\delta = \sup_{\alpha \in \Theta_\delta} v_n(\alpha) - |\delta| c_n^*$$

$$\delta_A = (0, 0, 0, 0), \quad \delta_B = (0, 1, 0, 0), \quad \delta_C = (0, 1, 0, 1), \quad \delta_0 = (1, 1, 1, 0) \quad \text{et} \\ \delta_E = (1, 1, 1, 1), \quad c_n^* = n, \quad c_n^* = \sqrt{n}, \quad c_n^* = \text{Log } n \quad \text{et} \quad c_n^* = \text{Log Log } n.$$

La pénalisation avec la vitesse $c_n^* = n$ conduit toujours à une sous-paramétrisation du modèle, alors que les autres vitesses arrivent à détecter vite le bon modèle. Un problème important consiste à rechercher la vitesse « optimale » c_n^* , celle qui détecte le plus rapidement le modèle.

APPENDICE

LOI DU LOGARITHME ITÉRÉ POUR UNE SUITE TRIANGULAIRE DE PROCESSUS PONCTUELS

Soit X une v. a. réelle positive définie sur une base stochastique $(\Omega, \mathcal{A}, \mathbb{F}, \mathbb{P})$ de loi F continue sur \mathbb{R}_+ et \mathbb{F} une filtration $\mathbb{F} = (\mathcal{F}_s)_{s \geq 0}$ telle que $N(s) = \mathbf{1}_{\{X \leq s\}}$ soit (\mathbb{F}, \mathbb{P}) compensé par $\tilde{N}(s)$ continu. On considère une suite $(X_i, \mathbb{F}_i, \mathbb{P}_i)_{i \geq 1}$ de copies indépendantes du modèle précédent. Posant $\mathcal{F}_s = \otimes_{i \geq 1} \mathcal{F}_{i,s}$ et $\mathbb{P} = \mathbb{P}^{\otimes \mathbb{N}}$, alors $M^n(s) = \sum_{i=1}^n (N_i - \tilde{N}_i)(s)$ est une (\mathbb{F}, \mathbb{P}) martingale localement de carré intégrable, nulle en 0 et de processus croissant $\langle M^n \rangle(s) = \tilde{N}^n(s)$ où $\tilde{N}^n = \sum_{i=1}^n \tilde{N}_i$ et $N^n = \sum_{i=1}^n N_i$. M^n, N^n et \tilde{N}^n sont à valeurs dans $D[0, \infty]$, l'espace des fonctions continues à droite et possédant des limites à gauche à valeurs dans \mathbb{R} . On munit $D[0, \infty]$ de la topologie de Skorohod qui coïncide sur l'espace $\mathcal{C}[0, \infty]$ des fonctions continues avec celle de la convergence uniforme. On démontre une loi du logarithme itéré analogue à celle de la répartition empirique [5].

THÉORÈME. — Pour tout $t > 0$ (éventuellement infini) et $h(n) = \sqrt{2n \operatorname{Log} \operatorname{Log} n}$ la suite $\left(\frac{M^n(s \wedge t)}{h(n)} \right)_{n \geq 3}$ est p.s. relativement compacte dans $D[0, t]$ avec pour ensemble de valeurs d'adhérence le sous-ensemble \mathbb{K}_F de $\mathcal{C}[0, t]$

$$\mathbb{K}_F = \left\{ G \text{ absolument continue par rapport à } F \right. \\ \left. \text{telle que } G(0) = 0 \text{ et } \int_0^t (dG/dF)^2(s) \cdot dF(s) \leq 1 \right\}.$$

Preuve. — 1. \mathbb{K}_F est relativement compact, en effet $G(0) = 0$ et pour tout $0 \leq s \leq s' \leq t$

$$|G(s') - G(s)| = \left| \int_{s'}^s (dG/dF) dF \right| \\ \leq \left(\int_{s'}^s dF \right)^{1/2} \cdot \left(\int_{s'}^s (dG/dF)^2 dF \right)^{1/2} \leq \sqrt{F(s) - F(s')}.$$

C.Q.F.D.

Soit $D = \{t_1, \dots, t_m\}$ une partition de $[0, t]$. A D fixé on notera pour une fonction $\gamma(\cdot)$ sur \mathbb{R}_+ , $\Delta_j \gamma = \gamma(t_j) - \gamma(t_{j-1})$; $j = 1 \dots m$ où $t_0 = 0$.

Toutes les partitions utilisées supposent $\Delta_j F > 0$.

Alors un lemme de F. Riesz [5] établit l'équivalence suivante: $G \in \mathbb{K}_F$, si et seulement si, pour toute subdivision D de $[0, t]$

$$G(0) = 0 \quad \text{et} \quad \sum_{j=1}^m \frac{(\Delta_j G)^2}{\Delta_j F} \leq 1.$$

On définit l'interpolation «F-linéaire de G relativement à D » G^D par :

$$G^D(s) = G(t_j) + (F(s) - F(t_j)) \frac{\Delta_j G}{\Delta_j F}, \quad s \in [t_j, t_{j+1}]$$

et \mathbb{K}_F^D l'ensemble des fonctions γ «F-linéaires relativement à D » qui vérifient $\gamma(0) = 0$ et $\sum_{j=1}^m \frac{(\Delta_j \gamma)^2}{\Delta_j F} \leq 1$.

D'une part il est bien connu que \mathbb{K}_F^D est contenu dans \mathbb{K}_F et d'autre part il est clair que si $G \in \mathbb{K}_F$ alors $G^D \in \mathbb{K}_F^D$.

Par ailleurs, la réunion sur toutes les subdivisions D des \mathbb{K}_F^D est dense dans \mathbb{K}_F , plus précisément: pour tout $G \in \mathbb{K}_F$ et $\varepsilon > 0$, on peut choisir D tel que $\sup_j \Delta_j F \leq \varepsilon$ et $\gamma \in \mathbb{K}_F^D$ vérifiant $\gamma(t_j) = G(t_j)$ $j = 1 \dots m$. γ étant

monotone sur $[t_j, t_{j+1}]$ on aura :

$$|\gamma(s) - G(s)| \leq |\gamma(s) - G(t_j)| + |G(t_j) - G(s)| \\ \leq |G(t_{j+1}) - G(t_j)| + |G(t_j) - G(s)| \\ \leq 2|\Delta_j G| \leq 2\sqrt{\Delta_j F} \leq 2\sqrt{\varepsilon}.$$

2. Maintenant pour démontrer que la suite $M^n/h(n)$ est relativement compacte et a pour adhérence \mathbb{K}_F , il suffit de prouver que pour toute partition D il existe $\Omega^D, \mathbb{P}(\Omega^D) = 1$ et tel que l'interpolation « F-linéaire de M^n relativement à D » $(M^n)^D$ vérifie pour tout $\omega \in \Omega^D$:

- (a) L'ensemble des valeurs d'adhérence de $((M^n)^D/h(n))_{n \geq 3}$ est \mathbb{K}_F^D .
- (b) $\limsup_n \sup_{s \leq t} |M^n(s) - (M^n)^D(s)| \leq 2 \sup_{1 \leq j \leq m} \sqrt{\Delta_j F}$.

On peut alors se restreindre à une famille dénombrable de partitions, par exemple celles pour lesquels les t_j sont rationnels et bien sûr $\Delta_j F > 0$.

(a) Pour $D = \{t_1, \dots, t_m\}$, on considère le vecteur aléatoire de \mathbb{R}^m $Y_m^n = (\Delta_1 M^n, \Delta_2 M^n, \dots, \Delta_m M^n)$ formé par les accroissements de M^n où $M^n(t_0) = 0$ et

$$\Delta_j M^n = M^n(t_j) - M^n(t_{j-1}) = \sum_{i=1}^n (\mathbf{1}_{\{t_j < x_i \leq t_{j-1}\}} - (\tilde{N}_i(t_j) - \tilde{N}_i(t_{j-1}))).$$

C'est une somme de v. a. indépendantes, de même loi sur \mathbb{R}^m , centrées et de matrice de covariance $\Gamma = (\gamma_{ij})$ avec $\gamma_{ij} = \mathbf{1}_{\{i=j\}} \cdot \Delta_j F, 1 \leq i, j \leq m$. La loi du logarithme itéré [9] pour une telle somme nous assure que la suite $(Y_m^n/h(n))_{n \geq 3}$ est p. s. relativement compacte avec pour ensemble de valeurs d'adhérence l'ellipsoïde :

$$\mathcal{E}_m(0, \Gamma^{-1}) = \left\{ u \in \mathbb{R}^m; \sum_{j=1}^m ((u_j)^2 / \Delta_j F) \leq 1 \right\}.$$

Ceci assure que la suite $((M^n)^D/h(n))_{n \geq 3}$ est p. s. relativement compacte et a pour ensemble de valeurs d'adhérence l'ensemble \mathbb{K}_F^D .

(b) On démontre que, pour toute subdivision D , il existe Ω^D avec $\mathbb{P}(\Omega^D) = 1$ pour lequel : $\forall \omega \in \Omega^D, \exists N(\omega)$ tel que pour $n \geq N(\omega)$ on a

$$\sup_{t_j < s \leq t_{j+1}} \left| \frac{(M^n)^D(s) - M^n(s)}{h(n)} \right| \leq \sqrt{F(t_{j+1}) - F(t_j)}, \quad j = 1 \dots m.$$

Soit

$$(\delta_j^D)^n = \sup (|M^n(s) - (M^n)^D(s)|, s \in [t_j, t_{j+1}]) \\ \leq 2 \sup (|M^n(s) - M^n(t_j)|, s \in [t_j, t_{j+1}]).$$

Fixons j et posons $\mathbb{F}^j = (\mathcal{F}_u^j)_{u \geq 0}$ avec $\mathcal{F}_u^j = \mathcal{F}_{t_j+u}$. Le processus $M^{n,j}(u) = M^n(t_j+u) - M^n(t_j)$ est une $(\mathbb{F}^j, \mathbb{P})$ martingale localement de carré intégrable, nulle en 0 et de processus croissant $\tilde{N}^n(t_j+u) - \tilde{N}^n(t_j)$ et le

lemme suivant a pour conséquence :

$$\overline{\lim}_n \sup \left(\frac{|M^{n,j}(u)|}{h(n)}, u \in [0, t_{j+1} - t_j] \right) \leq \sqrt{F(t_{j+1}) - F(t_j)} \quad \text{p. s.}$$

LEMME. — Pour tout $t > 0$ éventuellement infini, on a :

$$\overline{\lim}_n \sup_{s \leq t} (|M^n(s)|/h(n)) \leq \sqrt{F(t)} \quad \text{p. s.}$$

Preuve. — Les sauts de M^n sont bornés par 1, et on sait [4] que, pour tout $\lambda > 0$: $Z^n(s) = \exp \{ \lambda M^n(s) - \varphi(\lambda) N^n(s) \}$ est une surmartingale positive avec $Z^n(0) = 1$ et $\varphi(\lambda) = e^\lambda (e^\lambda - 1 - \lambda)$. Remarquons aussi que N^n est en même temps le « crochet droit » de M^n : $[M^n, M^n] = N^n$.

Si $F(t) = 0$ on a $N^n(t) = \tilde{N}^n(t) = 0$ p. s. et alors $\sup_{s \leq t} |M^n(s)| = 0$. Sinon soit t fixé, $t \leq \infty$ tel que $F(t) > 0$ ($F(\infty) = 1$) et soient $\theta > 1$, $\delta > 0$. Pour tout n vérifiant $\theta^k \leq n$, $F(t) < \theta^{k+1}$ posons $\alpha_n = \frac{1+\delta}{2} h(\theta^k)$, $\lambda_n = \frac{h(\theta^k)}{\theta^k}$

$$A_n^{\theta, \delta} = \left\{ \sup_{s \leq t} (\lambda_n M^n(s) - \varphi(\lambda_n) N^n(s) - \alpha_n \lambda_n) < 0 \right\}$$

et $\bar{A}_n^{\theta, \delta}$ son complémentaire. Du lemme maximal pour les surmartingales positives on déduit :

$$\begin{aligned} \mathbb{P}(\bar{A}_n^{\theta, \delta}) &= \mathbb{P}(\sup_{s \leq t} \lambda_n M^n(s) - \varphi(\lambda_n) N^n(s) - \alpha_n \lambda_n \geq 0) \\ &\leq \exp(-\alpha_n \lambda_n) = (k \text{ Log } \theta)^{-1-\delta} = \rho_k. \end{aligned}$$

Puisque $\sum_{k=1}^{\infty} \rho_k < \infty$, on conclut par le lemme de Borel-Cantelli que pour toute suite $(n_k)_{k \geq 1}$ qui satisfait à :

$$(\star) \quad n_k \in \left[\frac{\theta^k}{F(t)}, \frac{\theta^{k+1}}{F(t)} \right] \quad \text{on a } \mathbb{P}(\limsup \bar{A}_{n_k}^{\theta, \delta}) = 0.$$

Posons $V_n^* = \sup_{s \leq t} \frac{M^n(s)}{h(n)}$, fixons une suite (n_k) satisfaisant à la propriété

(\star), notons \mathcal{N} l'ensemble de ses indices et montrons que $\overline{\lim}_k V_{n_k}^* = \overline{\lim}_{n \in \mathcal{N}} V_n^* \leq \sqrt{F(t)}$. Pour tout $\omega \notin \limsup_k \bar{A}_{n_k}^{\theta, \delta}$ on a,

$$\overline{\lim}_k V_{n_k}^* \leq \overline{\lim}_{n \in \mathcal{N}} \frac{\varphi(\lambda_n) N^n(t)}{\lambda_n \cdot h(n)} + \overline{\lim}_{n \in \mathcal{N}} \frac{\alpha_n}{h(n)}$$

or pour $n \in \mathcal{N}$,

$$\frac{\alpha_n}{h(n)} \leq \frac{(1+\delta) h(\theta^k)}{2 h(n)}$$

donc

$$\overline{\lim}_{n \in \mathcal{N}} \frac{\alpha_n}{h(n)} \leq \left(\frac{1 + \delta}{2} \right) \cdot \sqrt{F(t)}$$

et

$$\overline{\lim}_{n \in \mathcal{N}} \frac{\varphi(\lambda_n) \cdot \lambda_n \cdot N^n(t)}{(\lambda_n)^2 \cdot h(n)} \leq \overline{\lim}_{n \in \mathcal{N}} \frac{\varphi(\lambda_n)}{(\lambda_n)^2} \overline{\lim}_{n \in \mathcal{N}} \frac{N^n(t)}{n \cdot F(t)} \cdot \overline{\lim}_{n \in \mathcal{N}} \frac{n \cdot F(t) \cdot \lambda_n}{h(n)}.$$

Soit $G = \left\{ \omega; \lim_{n \uparrow \infty} \frac{N^n(t)}{n F(t)} = 1 \right\}$ et \bar{G} son complémentaire.

D'après la loi forte $\mathbb{P}(G) = 1$.

On a, $\lambda_n \rightarrow 0$ d'où $\lim_{n \in \mathcal{N}} \varphi(\lambda_n)/(\lambda_n)^2 = 1/2$ et puisque $\theta^k \leq n \cdot F(t) \leq \theta^{k+1}$ on

a aussi $\overline{\lim}_{n \in \mathcal{N}} n \cdot F(t) \cdot \lambda_n/h(n) \leq \theta$.

Donc pour tout $\omega \notin \bar{G} \cup \{ \limsup_k \bar{A}_{n_k}^{\theta, \delta} \}$,

$$\overline{\lim}_k V_{n_k}^* \leq \frac{1 + \delta + \theta}{2} \sqrt{F(t)} \quad \text{p. s.}$$

Pour θ et δ donnés, il y a un nombre dénombrable de suites (n_k) convenables satisfaisant à (\star) , alors

$$\bar{\Gamma}_{\theta, \delta} = \bigcup_{(n_k)} \left\{ \limsup_k \bar{A}_{n_k}^{\theta, \delta} \right\}$$

est négligeable, comme l'est d'ailleurs

$$\bar{\Gamma} = \bigcup \{ \bar{\Gamma}_{\theta, \delta}; \theta, \delta \text{ rationnels, } \theta > 1 \text{ et } \delta > 0 \}.$$

Montrons enfin que pour tout $\omega \notin \bar{\Gamma} \cup \bar{G}$, $\overline{\lim}_n V_n^* \leq \sqrt{F(t)}$. Supposons que,

$\overline{\lim}_n V_n^* = l > \sqrt{F(t)}$, alors on peut choisir $\theta > 1$, $\delta > 0$ tels que,

$l > (1 + \theta + \delta) \sqrt{F(t)}/2$ et extraire une sous-suite (n_k) qui vérifie (\star) pour θ et δ , mais alors $l = \lim_k V_{n_k}^* \leq (1 + \theta + \delta) \sqrt{F(t)}/2$ et c'est la contradiction

recherchée. Une démonstration similaire s'opère sur $-M^n$.

RÉFÉRENCES

- [1] A. AKAIËRÉ, A new look at the statistical model identification, *I.E.E.E. Trans. Auto. Control*, vol. **AC-19**, n° 6, 1974.
- [2] P. K. ANDERSEN et R. D. GILL, Cox's regression model for counting process: A large sample study, *Ann. Stat.*, vol. **10**, n° 4, 1982.
- [3] D. R. COX, Partial likelihood, *Biometrika*, vol. **62**, 1975.

- [4] D. DACUNHA-CASTELLE et M. DUFLO, Probabilités et Statistiques, tome 2: Problèmes à temps mobile, Masson, Paris, 1983.
- [5] H. FINKELSTEIN, The law of the iterated logarithm for empirical distribution, *Ann. Math. Stat.*, vol. **42**, n° 2, 1971.
- [6] T. NAES, The asymptotic distribution of the estimator for the regression parameter in Cox's regression model, *Scand. J. Stat.*, vol. **9**, 1982.
- [7] R. R. RAO, The law of large number for $D[0, 1]$ valued random variables, *Theor. Probab. Appl.*, vol. **8**, 1963.
- [8] T. ROCKAFELLAR, Convex analysis, Princeton University Press, 1970.
- [9] R. SENOUSI, Statistique asymptotique p. s. de modèles statistiques convexes, *Ann. Inst. Henri-Poincaré: Probabilités et Statistiques*, vol. **26**, n° 1, 1989.
- [10] A. A. TSIATIS, A large sample study of Cox's regression model, *Ann. Stat.*, vol. **9**, 1981.

(Manuscrit reçu le 14 février 1989)

(version révisée le 26 juin 1989.)