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

JEAN DESHAYES
DOMINIQUE PICARD

Lois asymptotiques des tests et estimateurs de rupture dans un modèle statistique classique

Annales de l'I. H. P., section B, tome 20, nº 4 (1984), p. 309-327
<http://www.numdam.org/item?id=AIHPB_1984__20_4_309_0>

Ann. Inst. Henri Poincaré,
Vol. 20, n° 4, 1984, p. 309-327. *Probabilités et Statistiques*

Lois asymptotiques des tests et estimateurs de rupture dans un modèle statistique classique

par

Jean DESHAYES
E. R. A. CNRS 1044, E. N. S. T.
46, rue Barrault
75634 Paris Cedex 13

et

Dominique PICARD
E. R. A. CNRS 532
Université Paris-Sud

RÉSUMÉ. — Nous étudions ici le comportement asymptotique des méthodes liées à la vraisemblance sur des modèles statistiques réguliers avec rupture. L'utilisation de théorèmes d'invariance sur le processus de vraisemblance donne la clé de nombreux problèmes : comportement aux bords, renormalisation de la statistique de test, liaison des paramètres statistiques et temporel d'une rupture.

SUMMARY. — We investigate here the asymptotic behaviour of likelihood methods for regular statistical models with change-point. The use of invariance theorems for the likelihood process is the key for many problems: behaviour near the edges, renormalization of the test statistic, link of the time-parameter and the statistical ones of a change.

INTRODUCTION

Lorsqu'on observe un phénomène durant une longue période, le problème de la stabilité du modèle est fréquemment posé. La non stationnarité du modèle sera ici envisagée sous l'angle de la segmentation de l'intervalle d'observation en phases stationnaires. Nous étudions les tests d'existence de rupture et les méthodes d'identification de cette rupture. Nous nous concentrerons sur les méthodes liées à la vraisemblance et ferons une étude approfondie de leurs comportements asymptotiques.

Le problème de la loi limite de la statistique du rapport de vraisemblance pour tester l'existence d'une rupture à un instant inconnu date des années 60 : est-ce une loi de χ^2 ou non [*15*], [*6*] ? Faut-il renormaliser la statistique [*7*]? ou la « bootstrapper » [*10*] ? Des éléments de réponse ont été fournis par Hawkins [*7*], Mc Neil [*14*], Bhattacharya-Brockwell [*2*]. De même le comportement asymptotique des paramètres et en particulier de l'instant de rupture n'a été que partiellement résolu : l'asymptotique d'Hinkley [*8*] [*9*] (rupture d'amplitude fixe) obtenue par des méthodes de fluctuations donne des résultats difficilement calculables et fortement dépendants de la loi des observations ; l'alternative consiste à envisager l'asymptotique pour une rupture d'amplitude tendant lentement vers zéro (Bhattacharya [*1*]).

Nous montrons ici que ces problèmes trouvent une solution globale par l'utilisation de théorèmes d'invariance sur le processus de vraisemblance tels celui démontré dans [*4*]. Ces théorèmes, en rendant continu le paramètre temporel, permettent d'éclairer les problèmes de bords et le travail dans des espaces fonctionnels liant les paramètres statistiques classiques et temporel ; l'utilisation d'échelles de temps appropriées permet de régler à la fois les problèmes de test et d'estimation.

1. MODÈLE ET NOTATIONS

Pour un ouvert Θ non vide de R^p, considérons une famille $\{ \mathrm{P}_\theta, \theta \in \Theta \}$ de lois de probabilité sur un espace polonais $\mathscr{X}$, absolument continues par rapport à une mesure μ σ-finie et de densité notée $g(\,.\,,\theta)$.

Pour une valeur θ du paramètre (appartenant à Θ), nous définissons l'hypothèse $\mathrm{H}_{0,n}(\theta)$ correspondant à l'observation d'un n-échantillon X_1, $\mathrm{X}_2, \ldots, \mathrm{X}_n$ de loi P_θ.

Nous appellerons rupture dans le modèle, l'existence d'un instant $k \, (p \leqslant k \leqslant n-p)$ et de paramètres θ_1 et θ_2 $(\theta_1 \in \Theta,\ \theta_2 \in \Theta,\ \theta_1 \neq \theta_2)$, hypothèse notée $\mathrm{H}_{1,n}(\theta_1, k, \theta_2)$, correspondant à l'observation des variables $\mathrm{X}_1, \ldots, \mathrm{X}_n$ indépendantes et la loi de X_i est P_{θ_1} pour $i \leqslant k$ et P_{θ_2} pour $i > k$. Le test du rapport de vraisemblance rejette l'hypothèse H_0 lorsque la statistique T_n dépasse un seuil c :

$$\mathrm{T}_n = \sup_k \sup_{\theta_1} \sup_{\theta_2} \inf_\theta \frac{\prod_{i=1}^{k} g(\mathrm{X}_i, \theta_1) \prod_{i=k+1}^{n} g(\mathrm{X}_i, \theta_2)}{\prod_{i=1}^{n} g(\mathrm{X}_i, \theta)}$$

Pour évaluer son comportement sous l'hypothèse $H_0(\theta_0)$, il est commode d'introduire ici le processus de vraisemblance défini sur $R^p \times [0, 1]$ comme suit : pour tout θ appartenant à R^p, $Z_n(\theta, t)$ est la ligne polygonale joignant les points

$$(0, 1) \quad \text{et} \quad \left(\frac{k}{n}, \prod_{i=1}^{k} \frac{g\left(X_i, \theta_0 + \frac{\theta}{\sqrt{n}}\right)}{g(X_i, \theta_0)} \right), \qquad k = 1, \ldots, n.$$

$\left(Z_n \right.$ n'est *a priori* défini que sur l'ensemble $\mathscr{D}_n = \left\{ \theta, \theta_0 + \frac{\theta}{\sqrt{n}} \in \Theta \right\}$ mais cela ne change rien : il suffit de le prendre nul en dehors d'un voisinage de $\mathscr{D}_n$ et de le prolonger ensuite par continuité).

Soit φ une fonction continue par morceaux sur $[0, 1]$, telle que $\frac{\varphi(t)}{\sqrt{t}}$ soit croissante et satisfaisant $\varphi(0) = 0$, $\varphi(1) = 1$; nous appellerons $\mathscr{C}_\varphi$ l'ensemble des fonctions continues sur $R^p \times [0, 1]$ vérifiant :

$$\begin{cases} z(\theta, 0) \equiv 1 \\ \lim\limits_{M \to \infty} \operatorname{Sup}\,(z, \varphi(t) \mid \theta \mid > M) = 0 \end{cases}$$

la notation $\operatorname{Sup}\,(z, D)$ désignant le suprémum de la fonction z sur D. Dans le théorème 1 de [*4*], il est montré la convergence étroite, sous $H_0(\theta_0)$ de la suite de processus Z_n vers le processus Z :

$$Z(\theta, t) = \exp \left\{ {}^t\theta I(\theta_0)^{1/2} W_p(t) - \frac{t}{2} \cdot {}^t\theta I(\theta_0)\theta \right.$$

où W_p est un mouvement brownien sur $[0, 1]$ à p coordonnées indépendantes.

$I(\theta_0)$ est l'information de Fisher.

Cette convergence a lieu dans l'espace fonctionnel $\mathscr{C}_\varphi$ dès que φ satisfait la condition :

$$\int_0^1 \left[\frac{\varphi(t)}{t} \right]^2 dt < \infty .$$

Dans la suite, nous supposerons vérifiées les hypothèses de régularité de [*4*] que nous rappelons :

A] *i*) la fonction $\theta \rightsquigarrow > g(x, \theta)$ est absolument continue sur R^p, μ-presque sûrement en x.

ii) $\mathrm{I}(\theta) = \int \frac{\partial}{\partial \theta} g^{1/2}(x, \theta) \frac{{}^t\partial}{\partial \theta} g^{1/2}(x, \theta) d\mu(x)$ est continue sur R^p et définie positive en θ_0.

iii) $\forall \delta > 0$, $\forall w$, $| w | = 1$,

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon^p} \int_{|v-\theta_0| \leq \varepsilon} d\lambda(v) \int_{\left\{x, \left|\mathrm{Log} \frac{g(x,\theta_0+\varepsilon w)}{g(x,\theta_0)}\right| > \delta\right\}} \frac{{}^t\partial}{\partial v} g^{1/2}(x, v) \frac{\partial}{\partial v} g^{1/2}(x, v) d\mu(x) = 0$$

où λ est la mesure de Lebesgue sur R^p.

Pour un sous-ensemble u non vide de $\{1, 2, \ldots, p\}$, on notera ∂_u ou $\frac{\partial^{l(u)}}{\partial \theta^u}$ l'opérateur de dérivation par rapport aux coordonnées d'indices appartenant à u : ($l(u)$ désigne le cardinal de u).

$$\text{Si} \qquad u = \{ j_1, j_2, \ldots, j_{l(u)} \}, \qquad \partial_u = \frac{\partial^{l(u)}}{\partial \theta^{(j_1)} \ldots \partial \theta^{(j_{l(u)})}}$$

B] Il existe $a > 0$, un entier $m \geq 1$ et un entier pair q tels que $m \cdot q > p$

$$\text{et} \qquad \sup_{\theta \in \mathbb{R}^p} \frac{1}{1 + |\theta|^a} \int | \partial_u g^{\frac{1}{q}}(x, \theta_0 + \theta) |^q d\mu(x) < \infty$$

pour tout sous-ensemble u vérifiant $l(u) \leq m$.

C] Il existe $\eta > 0$ tels que $\overline{\lim}_{|\theta| \to \infty} | \theta |^\eta \cdot \rho(\theta_0, \theta_0 + \theta) < \infty$ où ρ désigne l'affinité de Hellinger : $\rho(\theta_0, \theta_0 + \theta) = \int_{\mathscr{X}} [g(x, \theta_0) g(x, \theta_0 + \theta)]^{\frac{1}{2}} d\mu(x)$.

D] Pour tout ε positif

$$\mathrm{P}_{\theta_0} \left\{ \sup_{\substack{\theta \\ |\theta| > \mathrm{M}}} \frac{g(\mathrm{X}, \theta_0 + \theta)}{g(\mathrm{X}, \theta_0)} > \varepsilon \right\}$$

tend vers 0 lorsque M tend vers l'infini.

E] La matrice Hessienne en θ de $\mathrm{L}(\mathrm{X}, \theta) = \mathrm{Log}\, g(\mathrm{X}, \theta_0 + \theta)$ est p. s. continue en $\theta = 0$ et il existe $\delta > 0$ tel que $\mathrm{E} \sup (| \ddot{\mathrm{L}}(\mathrm{X}, .) |, | \theta | \leq \delta) < + \infty$.

Remarque sur les hypothèses

i) nous renvoyons à [*4*] pour des premiers commentaires techniques et compléments.

ii) nous pouvons ajouter que ces conditions sont remplies pour les modèles réguliers classiques : lois gaussiennes évidemment, loi de Bernoulli, loi de Poisson, lois Gamma,...

Les hypothèses A et B sont classiques ([*11*]) et avec la condition E, elles remplacent avantageusement les traditionnelles conditions de Cramer. Les hypothèses C et D précisent la séparation des lois pour des paramètres s'éloignant indéfiniment ; elles sont trivialement satisfaites lorsque l'ensemble des paramètres est compact.

A titre d'exemple, nous pouvons vérifier que ces conditions A, B, C, D, E sont satisfaites pour un modèle d'homothétie-translation de paramètre ${}^t\theta = (\eta, \sigma)$ où la densité s'écrit : $\frac{1}{\sigma} g\left(\frac{x-\eta}{\sigma}\right)$ pour des fonctions g vérifiant :

. $\int \left|\frac{\dot{g}(x)}{g(x)}\right|^{2+\varepsilon} g(x)dx < \infty$ avec $\varepsilon > 0$, ce qui assure les conditions A et B.

. $g(x)$ décroissante à l'infini et $\overline{\lim_{x \to \pm\infty}} \, |x|^{1+\delta} g(x) < \infty$ avec $\delta > 0$, ce qui assure les conditions C et D.

. $\int \sup_{[x-\delta, x+\delta]} |\dot{\log} g| \, . \, g(x) \, . \, dx < \infty$, ce qui assure la condition E ; cette dernière contrainte est en particulier remplie lorsque $\dot{\log}\, g$ est uniformément continue ou que $\ddot{\log}\, g$ est bornée.

2. TESTS D'EXISTENCE D'UNE RUPTURE

2.1. Renormalisation du test du rapport de vraisemblance.

Le comportement asymptotique de la statistique T_n et d'autres tests se déduit de l'étude, sous $H_{0,n}(\theta_0)$, de la suite $\Lambda_n\left(\frac{k}{n}, \theta_1, \theta_2 ; \theta\right)$ de rapports de vraisemblance de l'hypothèse alternative $H_{1,n}\left(\theta_0 + \frac{\theta_1}{\sqrt{n}}, k, \theta_0 + \frac{\theta_2}{\sqrt{n}}\right)$ à l'hypothèse nulle $H_{0,n}\left(\theta_0 + \frac{\theta}{\sqrt{n}}\right)$. Posons

$$\mathscr{L}_n\left(\varphi, \frac{k}{n}\right) = \varphi\left(\frac{k}{n}\right) \cdot \varphi\left(1 - \frac{k}{n}\right) \operatorname*{Sup}_{\theta_1} \operatorname*{Sup}_{\theta_2} \operatorname*{Inf}_{\theta} \operatorname{Log} \Lambda_n\left(\frac{k}{n}, \theta_1, \theta_2 ; \theta\right)$$

THÉORÈME 2. — Sous la suite d'hypothèses $H_{0,n}(\theta_0)$,

i) si la fonction φ vérifie :

$$\lim_{t \to 0} \varphi(t) \log \log \frac{1}{t} = +\infty \tag{2.1}$$

alors $$\lim_{n \to \infty} \operatorname*{Sup}_{k} \mathscr{L}_n\left(\varphi, \frac{k}{n}\right) = +\infty$$

ii) si la fonction φ vérifie :

$$\int_0^1 \frac{\varphi(t)}{t}\,dt < \infty \tag{2.2}$$

alors $\sup\limits_k \mathscr{L}_n\left(\varphi, \frac{k}{n}\right)$ est borné en probabilité et plus précisément, la suite de processus définis par interpolation linéaire sur les points $\left(\frac{k}{n}, \mathscr{L}_n\left(\varphi, \frac{k}{n}\right)\right)$ converge étroitement dans $\mathscr{C}([0, 1])$ vers le processus $\mathscr{L}(\varphi, .)$ défini par :

$$\mathscr{L}(\varphi, t) = \varphi(t)\varphi(1-t)\frac{{}^t[W_p(t)-tW_p(1)][W_p(t) - tW_p(1)]}{2t(1-t)}.$$

Conséquences :

. La fonction φ identiquement égale à 1 satisfait (2.1) et par suite le théorème 2 implique qu'un niveau asymptotique α nécessite pour le test du rapport de vraisemblance un seuil $c_n(\alpha)$ tendant vers l'infini lorsque n tend vers l'infini.

. La fonction $\varphi(t) = 1_{[\varepsilon,1]}(t)$, avec $\varepsilon \in]0, 1[$ satisfait (2.2) et par conséquent, le test de vraisemblance excluant les bords de l'intervalle d'observation ($n\varepsilon \leqslant k \leqslant n - n\varepsilon$) nécessite un seuil $c_n(\alpha)$ convergeant vers $c(\alpha)$ borné. Le test du rapport de vraisemblance original n'a donc aucune puissance pour les ruptures intervenant au milieu de l'intervalle d'observation, à cause de son comportement aux bords.

. Il est donc raisonnable de pondérer aux bords la statistique de vraisemblance à l'aide de fonctions φ vérifiant (2.2) ; nous prenons donc comme région de rejet :

$$\left\{ \operatorname{Sup}_k \mathscr{L}_n\left(\varphi, \frac{k}{n}\right) > c \right\}$$

avec c déterminé par :

$$P\{ \operatorname*{Sup}_{t\in[0,1]} \mathscr{L}(\varphi, t) > c \} = \alpha$$

Cela revient en fait à mettre un seuil variant avec k : presque constant au milieu de l'intervalle d'observation, mais tendant vers l'infini aux bords, là où les paramètres sont moins bien estimés. En particulier, la fonction $\varphi(t) = \sqrt{2t}$ satisfait (2.2) et donne une loi limite simple.

Démonstration du théorème 2. — Comme le processus Λ_n s'écrit sous la forme :

$$\Lambda_n\left(\frac{k}{n}, \theta_1, \theta_2 ; \theta\right) = Z_n\left(\theta_1, \frac{k}{n}\right) . \frac{Z_n(\theta_2, 1)}{Z_n\left(\theta_2, \frac{k}{n}\right)} . \frac{1}{Z_n(\theta, 1)}$$

il suffit d'étudier le comportement du processus défini sur $[0, 1]$ par :

$$\mathrm{L}_n(\varphi, t) = \varphi(t) \operatorname{Log} \operatorname*{Sup}_{\theta} \mathrm{Z}_n(\theta, t)$$

La première partie du théorème 2 (φ vérifiant (2.1)) est une conséquence du théorème 1 de [*4*] et de la loi du logarithme itéré.

La deuxième partie (φ vérifiant (2.2)) consiste essentiellement en la tension en 0 de la suite $\mathrm{L}_n(\varphi, .)$ car la convergence étroite de $\mathrm{L}_n(\varphi, .)$ dans $\mathscr{C}([\varepsilon, 1])$ pour ε strictement positif est une conséquence du théorème 1 de [4]. Il nous reste donc à montrer :

$$\forall \varepsilon' > 0, \quad \exists \varepsilon > 0, \qquad \mathrm{P}\{ \operatorname*{Sup}_{t \in [0,\varepsilon]} \mathrm{L}_n(\varphi, t) > \varepsilon' \} \leqslant \varepsilon'$$

Pour des valeurs de M et δ à fixer ultérieurement (δ inférieur à la valeur donnée dans l'hypothèse E), décomposons le premier membre en trois parties : $\pi_1 + \pi_2 + \pi_3$.

$$\pi_1 = \mathrm{P}\{ \operatorname*{Sup}_{t \varepsilon [0,\varepsilon]} \operatorname*{Sup}_{|\theta| \leqslant \mathrm{M}} \varphi(t) \log \mathrm{Z}_n(\theta, t) > \varepsilon' \}$$

est directement contrôlé, pour ε suffisamment petit, par la tension de la suite Z_n dans $\mathscr{C}(\mathrm{R}^p \times [0, 1])$ grâce au théorème 1 de [*4*].

$$\pi_2 = \mathrm{P}\{ \operatorname*{Sup}_{t \in [0,\varepsilon]} \sup_{\mathrm{M} < |\theta| \leqslant \sqrt{n\delta}} \varphi(t) \log \mathrm{Z}_n(\theta, t) > \varepsilon' \}$$

se majore au moyen d'un développement de Taylor autour de $\theta = 0$ et de la loi des grands nombres appliquée à : $\operatorname{Sup}(\ddot{\mathrm{L}}(\mathrm{X}, .), |\theta| \leqslant \delta)$ comme dans le lemme 4 de [*4*] :

$$\begin{aligned}\pi_2 < \mathrm{P}\Bigg\{ & \exists k \in \{ n_0, \ldots, [n\varepsilon] \}, \sup_{\mathrm{M} < |\theta| \leqslant \sqrt{n\delta}} \frac{t_\theta}{n} \sum_{i=1}^{k} \dot{\mathrm{L}}(\mathrm{X}_i, \theta_0) \\ & - \frac{1 - \varepsilon'}{2} \cdot \frac{\lambda_{\min}}{\lambda_{\max}} \cdot {}^t\theta \mathrm{I}(\theta_0)\theta > \frac{\varepsilon'}{\varphi\left(\frac{k}{n}\right)} \Bigg\} \\ & + \mathrm{P}\left\{ \exists k < n_0, \sup_{\mathrm{M} < |\theta| \leqslant \sqrt{n\delta}} \varphi\left(\frac{k}{n}\right) \log \mathrm{Z}_n\left(\theta, \frac{k}{n}\right) > \varepsilon' \right\} + \frac{\varepsilon'}{9}\end{aligned}$$

où $\lambda_{\min}$ et $\lambda_{\max}$ désignent respectivement la plus petite et la plus grande valeur propre de $\mathrm{I}(\theta_0)$. La majoration termine en utilisant l'inégalité de Birnbaum-Marshall [*3*] pour le premier terme et l'hypothèse D pour le second.

La troisième partie :

$$\pi_3 = \mathrm{P}\left\{ \sup_{t\in[0,\varepsilon]} \sup_{|\theta|>(\mathrm{M},\sqrt{n}\delta)} \varphi\left(\frac{k}{n}\right) \log \mathrm{Z}_n\left(\theta, \frac{k}{n}\right) > \varepsilon' \right\}$$

se traite de façon parallèle au lemme 3 de [*4*].

2.2. **Commentaires sur le théorème 2 et autres applications.**

. La pondération de la vraisemblance avant de prendre le suprémum sur les valeurs possibles de l'instant de rupture peut aussi être envisagée du point de vue bayésien : la fonction $\varphi(t)\varphi(1-t)$ apparaissant comme une densité *a priori* de l'instant de rupture.

. D'autre part, ce point de vue bayésien peut aussi être adopté avec le remplacement du suprémum en t par une intégrale sur $[0, 1]$; la loi limite de la statistique découle immédiatement du théorème 1 de [*4*]; on remarque que la renormalisation n'est plus nécessaire aux bords dans ce cas :

COROLLAIRE (2.1). — La suite de statistiques de test

$$\frac{1}{n}\sum_{k=1}^{n} \operatorname*{Sup}_{\theta_1,\theta_2} \operatorname*{Inf}_{\theta} \log \Lambda_n\left(\frac{k}{n}, \theta_1, \theta_2 ; \theta\right)$$

converge, sous $\mathrm{H}_{0,n}(\theta_0)$, vers $\displaystyle\int_0^1 \frac{|\mathrm{W}_p(t) - t\mathrm{W}_p(1)|^2}{2t(1-t)}\,dt.$

. Les informations *a priori* sur les paramètres θ_1 et θ_2 peuvent aussi être prises en compte en remplaçant les supréma par une intégrale; un cas particulier intéressant est celui où le paramètre avant rupture est parfaitement connu : $\theta_1 = \theta_0$ et la renormalisation n'a lieu qu'en fin d'observation :

COROLLAIRE (2.2). — Sous la suite d'hypothèses $\mathrm{H}_{0,n}(\theta_0)$ avec θ_0 connu, la suite de statistiques de test : $\displaystyle\operatorname*{Sup}_{k} \operatorname*{Sup}_{\theta_2} \varphi\left(1 - \frac{k}{n}\right) \log \Lambda_n\left(\frac{k}{n}, 0, \theta_2 ; 0\right)$ converge vers : $\displaystyle\operatorname*{Sup}_{t\in[0,1]} \varphi(t)\cdot\frac{|\mathrm{W}_p(t)|^2}{2t}$ dès que la fonction φ vérifie (2.2).

. Dans certaines applications pratiques, une rupture peut se produire avec retour à la loi initiale avant la fin de la période d'observation (par exemple, apparition d'un signal constant dans un bruit stationnaire), dans ce cas le test non renormalisé reposerait sur la statistique :

$$\operatorname*{Sup}_{1<k_1<k_2<n} \operatorname*{Sup}_{\theta_2} \frac{Z_n\left(\frac{k_2}{n}, \theta_2\right)}{Z_n\left(\frac{k_1}{n}, \theta_2\right)}$$

et il est nécessaire de pondérer ici par $\varphi\left(\frac{k_2}{n} - \frac{k_1}{n}\right)$ avec φ satisfaisant (2.2).

COROLLAIRE (2.3). — Sous la suite d'hypothèses $H_{0,n}(\theta_0)$ avec θ_0 connu, la suite

$$\operatorname*{Sup}_{k_1<k_2} \operatorname*{Sup}_{\theta_2} \varphi\left(\frac{k_2}{n} - \frac{k_1}{n}\right) \log \frac{Z_n\left(\frac{k_2}{n}, \theta_2\right)}{Z_n\left(\frac{k_1}{n}, \theta_2\right)}$$

converge vers :

$$\operatorname*{Sup}_{0<t_1<t_2<1} \varphi(t_2 - t_1) \frac{|W_p(t_2) - W_p(t_1)|^2}{2(t_2 - t_1)}.$$

. Enfin, il va de soi que les théorèmes d'invariance de processus de vraisemblance permettent aussi de déduire les lois limites des statistiques lorsqu'on veut tester l'hypothèse H_0 d'absence de rupture contre des sous-hypothèses de H_1 correspondant au changement de certaines coordonnées seulement du paramètre θ. Par exemple, dans un modèle d'homothétie-translation, la densité des observations par rapport à la mesure de Lebesque s'écrit : $\frac{1}{\sigma} g\left(\frac{x - \eta}{\sigma}\right)$ et on peut tester un saut du paramètre de position η seul.

3. ESTIMATION DES TROIS PARAMÈTRES D'UNE RUPTURE

Le problème ici est, en présence d'une rupture, d'estimer les trois paramètres k, θ_1 *et* θ_2. La méthode classique (Ibragimov-Khasminskii [*11*]) d'obtention de la loi limite des principaux estimateurs (maximum de vraisemblance, bayésiens) dans un problème standard consiste à paramétrer convenablement un voisinage (au sens de la convergence des expériences : Lecam [*13*]) de l'hypothèse $H_{1,n}$ de sorte ensuite à prouver la convergence de la suite des processus de vraisemblance dans un espace fonctionnel qui permet de déduire sans efforts la convergence de la suite des estimateurs (par exemple $\mathscr{C}(K)$ si l'espace des paramètres K est compact, $\mathscr{C}_0$ s'il ne l'est pas). Nous allons adapter ici cette méthodologie.

Pour nous situer dans un cadre asymptotique, il nous faudra supposer que l'instant de rupture k varie avec n et satisfait :

$$k(n) \to +\infty \qquad \text{et} \qquad n - k(n) \to +\infty .$$

On vérifiera alors que les estimateurs classiques de θ_1 et θ_2 sont asymptotiquement gaussiens et efficaces : leur comportement limite est le même que si l'instant de rupture $k(n)$ était connu. Le problème spécifique ici est l'estimation de $k(n)$ et plusieurs cadres asymptotiques peuvent être proposés.

Le premier point de vue consiste à envisager un « saut » $d = \theta_2 - \theta_1$ d'amplitude fixe. Dans ce cas, ce n'est pas le passage en temps continu qui donne des résultats mais des méthodes de fluctuations. Hinkley a calculé la loi limite de l'estimateur du maximum de vraisemblance pour des modèles gaussiens [*8*] et de Bernoulli [*9*]. Si l'intérêt de ce point de vue est de donner une loi limite directement discrétisée, en revanche les lois limites obtenues dépendent lourdement de la loi initiale des variables et de l'amplitude de la rupture, ce qui rend nécessaires des calculs d'approximations pour chaque cas particulier.

Une manière de court-circuiter ce problème consiste à faire tendre $d^{(n)} = \theta_2^{(n)} - \theta_1^{(n)}$ vers 0. Intuitivement, cela diminuera la concentration de la loi limite de l'estimateur $\hat{k}$ autour de $k(n)$ puisque l'amplitude de la rupture est moins forte; cela explique pourquoi on se tourne vers l'estimation non plus de k mais en fait d'une fonction de k et de n. Ce faisant, il est intéressant de noter que la condition ($d^{(n)} \to 0$) ne suffit pas à cerner le problème à cause des problèmes de bords. Deux cas peuvent être distingués : le premier consisterait à envisager une asymptotique avec $\sqrt{n}d^{(n)}$ convergeant vers une limite d, les problèmes de bords rencontrés pour le test du rapport de vraisemblance se retrouvent semblables ici et il conviendrait de renormaliser la vraisemblance pour obtenir des lois limites. Cependant, il s'avérera peu intéressant de le faire pour deux raisons : d'abord, ce cas n'est pas « statistiquement raisonnable » car on cherche alors à estimer une rupture aussi petite que le degré de finesse de l'estimation, rendant impossible l'estimation consistante de l'instant de rupture ; d'autre part, le second cas ($\sqrt{n}d^{(n)}$ non borné) permet une reparamétrisation qui fournit des lois asymptotiques simples à calculer.

3.1. Estimateurs du maximum de vraisemblance d'une rupture.

Nous envisageons donc une suite de modèles avec rupture dont l'amplitude décroît lorsque la taille d'échantillon augmente : on écrit la suite d'hypothèses $H_1(\theta_1^{(n)}, k(n), \theta_2^{(n)})$ avec

$$k(n) \underset{n\to\infty}{\longrightarrow} \infty, \quad n-k(n) \to \infty, \quad \lim_{n\to\infty} \theta_1^{(n)} = \lim_{n\to\infty} \theta_2^{(n)} = \theta_0$$

et il existe α compris strictement entre 0 et 1 tel que $d^{(n)} = \theta_2^{(n)} - \theta_1^{(n)}$ satisfait :

$$\left\{ \frac{k(n)[n-k(n)]}{n} \right\}^{\alpha} {}^t d^{(n)} I(\theta_0) d^{(n)} \underset{n\to\infty}{\longrightarrow} \infty$$

THÉORÈME 3. — Sous la suite d'hypothèses $H_1(\theta_1^{(n)}, k(n), \theta_2^{(n)})$, les estimateurs du maximum de vraisemblance des paramètres sont asymptotiquement indépendants et de lois respectives :

$$\sqrt{k(n)}(\hat{\theta}_1 - \theta_1^{(n)}) \overset{\mathscr{L}}{\to} (0, I(\theta_0)^{-1})$$
$$\sqrt{n - k(n)}(\hat{\theta}_2 - \theta_2^{(n)}) \overset{\mathscr{L}}{\to} (0, I(\theta_0)^{-1})$$
$$\Psi(n)(\hat{k} - k(n)) \overset{\mathscr{L}}{\to} \xi_1$$

où

$$\Psi(n) = \frac{k(n)}{n}\left[1 - \frac{k(n)}{n}\right] {}^t d^{(n)} I(\theta_0) d^{(n)}$$

et ξ_1 est une variable aléatoire réelle de densité par rapport à la mesure de Lebesgue :

$$\frac{3}{2} e^{|s|} \int_{\frac{3}{\sqrt{2}}\sqrt{|s|}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du - \frac{1}{2} \int_{\frac{1}{2}\sqrt{|s|}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

Ce résultat est une conséquence facile d'un théorème d'invariance sous $H_1(\theta_1^{(n)}, k(n), \theta_2^{(n)})$ dans une échelle de temps appropriée ; le nouveau paramétrage traduit les vitesses d'estimation différentes pour chacun des paramètres :

$$\theta_1 \rightsquigarrow \theta_1^{(n)} + \frac{\theta_1^*}{\sqrt{k(n)}}, \quad \theta_2 \rightsquigarrow \theta_2^{(n)} + \frac{\theta_2^*}{\sqrt{n-k(n)}}, \quad k \rightsquigarrow k(n) + \frac{\tau}{\Psi(n)} = K(\tau) \quad (3.0)$$

Le processus de vraisemblance $Y_n(\theta_1^*, \tau, \theta_2^*)$ à étudier est alors défini pour θ_1^* et θ_2^* fixés par linéarisation entre les points τ correspondant à $K(\tau) = 1, 2, \ldots, n$:

$$Y_n(\theta_1^*, \tau, \theta_2^*) = \frac{dP_{H_1}\left(\theta_1^{(n)} + \frac{\theta_1^*}{\sqrt{k(n)}}, K(\tau), \theta_2^{(n)} + \frac{\theta^*}{\sqrt{n-k(n)}}\right)}{dP_{H_1}(\theta_1^{(n)}, k(n), \theta_2^{(n)})}$$

Soit $\mathscr{C}_0(\mathrm{R}^{2p+1}; \mathrm{R})$ l'espace des fonctions continues de R^{2p+1} dans R qui tendent vers 0 à l'infini, muni de la norme uniforme.

THÉORÈME 4. — La suite de processus Y_n converge étroitement dans $\mathscr{C}_0(\mathrm{R}^{2p+1}; \mathrm{R})$, sous $\mathrm{H}_1(\theta_1^{(n)}, k(n), \theta_2^{(n)})$ vers le processus Y défini par :

$$\begin{aligned}\operatorname{Log} \mathrm{Y}(\theta_1^*, \tau, \theta_2^*) &= {}^t\theta_1^*\zeta_1 + \mathrm{W}(\tau) + {}^t\theta_2^*\zeta_2 \\ &\quad - \frac{1}{2}\,{}^t\theta_1^*\mathrm{I}(\theta_0)\theta_1^* - \frac{|\tau|}{2} - \frac{1}{2}\,{}^t\theta_2^*\mathrm{I}(\theta_0)\theta_2^*\end{aligned}$$

où ζ_1 et ζ_2 sont deux vecteurs gaussiens $\mathcal{N}(0; \mathrm{I}(\theta_0))$ indépendants et indépendants de W : mouvement brownien sur R attaché en 0.

Remarque. — Le passage du théorème 4 au théorème 3 est immédiat en utilisant les résultats de Shepp [*16*] sur la loi de l'instant de suprémum d'un brownien avec dérive : la densité de ξ_1 est en effet celle de l'instant de suprémum sur R de $\mathrm{W}(\tau) - \dfrac{|\tau|}{2}$.

Démonstration du théorème 4. — En utilisant la symétrie naturelle du problème, nous ne traiterons que la restriction au domaine: $\tau \leqslant 0$ $(\mathrm{K}(\tau) \leqslant k(n))$. Il est alors commode de remarquer que sur $(\tau \leqslant 0)$, on peut regrouper les deux termes extrêmes pour diviser $\mathrm{Y}_n(\theta_1^*, \tau\theta_2^*)$ en deux parties :

$$\mathrm{Z}_{k(n)}\left(\theta_1^*, \frac{\mathrm{K}(\tau)}{k(n)}\right) . \ \tilde{\mathrm{Z}}_{n-k(n)}(\theta_2^*, 1)$$

et
$$\overline{\mathrm{Y}}_{k(n)}(\theta_2^*, \tau) = \frac{\mathrm{Z}_{k(n)}\left(\theta_2^*\sqrt{\dfrac{k(n)}{n-k(n)}} + \sqrt{k(n)}\,.\,d^{(n)}, 1\right)}{\mathrm{Z}_{k(n)}\left(\theta_2^*\sqrt{\dfrac{k(n)}{n-k(n)}} + \sqrt{k(n)}d^{(n)}, \dfrac{\mathrm{K}(\tau)}{k(n)}\right)}$$

où la suite $\mathrm{Z}_{k(n)}$ est construite à partir d'un $k(n)$-échantillon de la loi $\mathrm{P}_{\theta_1^{(n)}}$ et $\tilde{\mathrm{Z}}_{n-k(n)}$ d'un $n-k(n)$-échantillon de la loi $\mathrm{P}_{\theta_2^{(n)}}$ indépendant du précédent. L'effet principal de cette partition est de séparer les paramètres (θ_1^*, θ_2^*) d'une part et τ d'autre part. La partie la plus difficile est le comportement de la suite $\overline{\mathrm{Y}}_n$ que nous résumons dans le lemme suivant

LEMME 3.1. — *i)* $\forall \mathrm{K}$ compact de R^p, $\forall \tau_1 < 0$, $\forall \varepsilon > 0$, $\exists n_0$ tels que pour $n \geqslant n_0$,

$$\mathrm{P}_{\theta_1^{(n)}}\left\{ \sup_{\theta \in \mathrm{K}} \sup_{\tau < \tau_1} |\log \overline{\mathrm{Y}}_n(\theta, \tau) - \log \overline{\mathrm{Y}}_n(0, \tau)| \geqslant \varepsilon \right\} \leqslant \varepsilon \qquad (3.1)$$

ii) $\forall \tau_1 < 0, \forall \varepsilon > 0, \exists A > 1, \exists n_0$ tels que pour $n \geqslant n_0$,

$$P_{\theta_1^{(n)}}\{ \operatorname*{Sup}_{\tau < \tau_1} \operatorname*{Sup}_{|\theta| \leqslant 2\sqrt{n-k(n)}|d^{(n)}|} \overline{Y}_n(\theta, \tau) > A \} \leqslant \varepsilon \tag{3.2}$$

iii) la suite $\overline{Y}_n(0, \tau)$ converge étroitement, sous $P_{\theta_1^{(n)}}$, dans $\mathscr{C}(R^- ; R)$ muni de la norme uniforme sur les compacts, vers $W(\tau) - \dfrac{|\tau|}{2}$.

iv) $\forall \varepsilon > 0, \forall N > 0, \exists \zeta_1 < 0$ tel que :

$$P_{\theta_1^{(n)}}\{ \operatorname*{Sup}_{\tau \leq \tau_1} \operatorname*{Sup}_{|\theta| \leqslant 2\sqrt{n-k(n)}|d^{(n)}} \overline{Y}_n(\theta, \tau) > \varepsilon \} \leqslant \varepsilon \tag{3.3}$$

$$P_{\theta_1^{(n)}}\{ \operatorname*{Sup}_{K(\tau) \leqslant \frac{k(n)}{2}} \operatorname*{Sup}_{|\theta| \leqslant 2\sqrt{n-k(n)}|d^{(n)}} \overline{Y}_n(\theta, \tau) > k(n)^{-N} \} \leqslant \varepsilon \tag{3.4}$$

v) pour toute suite A_n vérifiant $\lim\limits_n \operatorname{Log} k(n) \cdot A_n^{-1} < \infty$, on a

$$P_{\theta_1^{(n)}}\{ \sup_{\tau \leqslant 0} \sup_{\theta \in R^p} \operatorname{Log} \overline{Y}_n(\theta, \tau) > A_n \} \leqslant \varepsilon \tag{3.5}$$

Nous ne donnons pas ici le détail de la preuve : disons simplement que mises à part les assertions (*iii*) et (*v*) (la partie (*iii*) est une conséquence simple du théorème de Donsker et de la loi des grands nombres ; la partie (*v*) découle directement de la démonstration du théorème 1 de [*4*]), les autres démonstrations s'obtiennent toutes au moyen d'un développement de Taylor, puis utilisation d'inégalités du type maximales ou Birnbaum-Marshall [*3*] comme dans le lemme 4 de [*4*]. Les détails de démonstration figurent dans [*5*].

Pour obtenir la convergence étroite dans la topologie de la convergence uniforme sur les compacts, il suffit d'adjoindre aux assertions *i*) et *iii*), le lemme suivant qui permet de remplacer $Z_{k(n)}\left(\theta_1^*, \dfrac{K(\tau)}{k(n)}\right)$ par $Z_{k(n)}(\theta_1^*, 1)$; ce lemme est une conséquence du théorème 1 de [*4*] :

LEMME 3.2. — $\forall C$ compact de R^p, $\forall \tau_1 < 0, \forall \varepsilon > 0, \exists n_0$ tels que pour $n > n_0$,

$$P\left\{ \sup_{K(\tau_1) \leqslant k \leqslant k(n)} \sup_{\theta \in C} \left| Z_{k(n)}\left(\theta, \frac{k}{k(n)}\right) - Z_{k(n)}(\theta, 1) \right| > \varepsilon \right\} \leqslant \varepsilon \tag{3.6}$$

Pour ce qui est maintenant de la convergence dans $\mathscr{C}_0$, nous allons toujours séparer les paramètres en (θ_1^*, θ_2^*) d'une part et τ d'autre part, et montrer en fait le

LEMME 3.3. — *i*) $\forall \varepsilon > 0, \exists \tau_1 < 0$ tel que

$$P\{ \operatorname*{Sup}_{\tau \leqslant \tau_1} \operatorname*{Sup}_{(\theta_1^*, \theta_2^*) \in R^{2p}} Y_n(\theta_1^*, \tau, \theta_2^*) > \varepsilon \} \leqslant \varepsilon \tag{3.7}$$

ii) $\forall \varepsilon > 0$, $\exists M > 0$ tel que

$$P\{\operatorname*{Sup}_{\tau_1 \leqslant \tau \leqslant 0} \sup_{|(\theta_1^*, \theta_2^*)| \geqslant M} Y_n(\theta_1^*, \tau, \theta_2^*) > \varepsilon\} \leqslant \varepsilon \tag{3.8}$$

Démonstration du lemme 3.3. — Pour prouver (3.7), il est commode de diviser l'ensemble sur lequel on prend le suprémum en trois domaines :

$$D_1 = \left\{ (\theta_1^*, \theta_2^*, \tau),\ |\theta_2^*| < 2\sqrt{n-k(n)}\,|d^{(n)}|,\ \frac{k(n)}{2} \leqslant K(\tau) \leqslant K(\tau_1) \right\}$$

$$D_2 = \left\{ (\theta_1^*, \theta_2^*, \tau),\ |\theta_2^*| < 2\sqrt{n-k(n)}\,|d^{(n)}|,\ 1 \leqslant K(\tau) < \frac{k(n)}{2} \right\}$$

$$D_3 = \left\{ (\theta_1^*, \theta_2^*, \tau),\ |\theta_2^*| > 2\sqrt{n-k(n)}\,|d^{(n)}|,\ 1 \leqslant K(\tau) < \frac{k(n)}{2} \right\}$$

$$D_3 = \{ (\theta_1^*, \theta_2^*, \tau),\ |\theta_2^*| > 2\sqrt{n-k(n)}\,|d^{(n)}|,\ 1 \leqslant K(\tau) \leqslant K(\tau_1) \}$$

. Sur D_1, les processus $Z_{k(n)}\left(\theta_1^*, \frac{K(\tau)}{k(n)}\right)$ et $\tilde{Z}_{n-k(n)}(\theta_2^*, 1)$ sont bornés en probabilité par le théorème 1 [*4*], le comportement de $\overline{Y}_n(\theta_2^*, \tau)$ est déterminé par l'inégalité (3.3).

. Sur D_2, le processus $\tilde{Z}_{n-k(n)}(\theta_2^*, 1)$ est toujours borné en probabilité; le théorème 1 [*4*] implique que si A_n est une suite de réels vérifiant

$$\lim \operatorname{Log} k(n) \cdot A_n^{-1} < \infty\,,$$

alors le processus $\operatorname{Log} Z_{k(n)}\left(\theta_1^*, \frac{K(\tau)}{k(n)}\right)$ est inférieur à A_n sur D_2 avec une probabilité supérieure à $1 - \varepsilon$, la compensation est alors apportée par $\overline{Y}_n(\theta_2^*, \tau)$ grâce à l'inégalité (3.4).

. Sur D_3, le théorème 1 [*4*] et l'inégalité (3.5) permettent de borner en probabilité par A_n le produit $Z_{k(n)}\left(\theta_1^*, \frac{K(\tau)}{k(n)}\right) \cdot \overline{Y}_n(\theta_2^*, \tau)$, la compensation est cette fois apportée par $\tilde{Z}_{n-k(n)}(\theta_2^*, 1)$ grâce au lemme 2 de [*4*].

Pour démontrer l'inégalité (3.8), on divise l'espace en deux domaines :

$$B_1 = \Big\{ (\theta_1^*, \theta_2^*, \tau),\quad \tau_1 \leqslant \tau \leqslant 0,\quad |(\theta_1^*, \theta_2^*)| > M,\ \left| \sqrt{\frac{k(n)}{n-k(n)}}\,\theta_2^* + \sqrt{k(n)}\,d^{(n)} \right| \leqslant 2\sqrt{k(n)}\,|d^{(n)}| \Big\}$$

$$B_2 = \Big\{ (\theta_1^*, \theta_2^*, \tau),\quad \tau_1 \leqslant \tau \leqslant 0,\quad \left| \sqrt{\frac{k(n)}{n-k(n)}}\,\theta_2^* + \sqrt{k(n)}\,d^{(n)} \right| > 2\sqrt{k(n)}\,|d^{(n)}| \Big\}.$$

. Sur B_1, le processus $\overline{Y}_n(\theta_2^*, \tau)$ est borné en probabilité grâce à l'inégalité (3.2), la compensation est alors apportée soit par $Z_{k(n)}\left(\theta_1^*, \frac{K(\tau)}{k(n)}\right)$ soit par $Z_{n-k(n)}(\theta_2^*, 1)$ suivant que $|\theta_1^*|$ ou $|\theta_2^*|$ est supérieur à M, grâce au théorème 1 de [*4*].

. Sur B_2, le processus $\overline{Y}_n(\theta_2^*, \tau)$ est contrôlé par l'inégalité (3.5), le processus $Z_{k(n)}\left(\theta_1^*, \frac{K(\tau)}{k(n)}\right)$ est borné en probabilité (théorème 1 [*4*]), la compensation étant apportée par $\tilde{Z}_{n-k(n)}(\theta_2^*, 1)$ grâce au lemme 2 de [*4*].

3.2. Estimateurs bayésiens d'un modèle avec rupture.

Lorsque des informations *a priori* sur les paramètres θ_1, k, θ_2 d'une rupture conduisent à supposer une loi *a priori* de densité sur $R^p \times [0, 1] \times R^p$: $q\left(\theta_1, \frac{k}{n}, \theta_2\right)$ continue et strictement positive au point limite $(\theta_0, t_0, \theta_0)$ de la suite $\left(\theta_1^{(n)}, \frac{k(n)}{n}, \theta_2^{(n)}\right)$, les estimateurs bayésiens de chacun des paramètres s'expriment sous $H_1(\theta_1^{(n)}, k(n), \theta_2^{(n)})$ à l'aide du processus de vraisemblance Y_n.

Avec la fonction de perte quadratique, les estimateurs sont donnés par les formules suivantes :

$$\tilde{\theta}_1 = A^{-1} \int \theta_1 Q(\theta_1, t, \theta_2 ; X_1, \ldots, X_n) d\theta_1 dt d\theta_2$$

$$\tilde{k} = n \cdot A^{-1} \int t Q(\theta_1, t, \theta_2 ; X_1, \ldots, X_n) d\theta_1 dt d\theta_2$$

$$\tilde{\theta}_2 = A^{-1} \int \theta_2 Q(\theta_1, t, \theta_2 ; X_1, \ldots, X_n) d\theta_1 dt d\theta_2$$

où $Q(\theta_1, t, \theta_2 ; X_1, \ldots, X_n)$ désigne la densité jointe

$$q(\theta_1, t, \theta_2) \prod_{i=1}^{[nt]} g(X_i, \theta_1) \prod_{i=[nt]+1}^{n} g(X_i, \theta_2)$$

et

$$A = \int Q(\theta_1, t, \theta_2 ; x_1, \ldots, X_n) d\theta_1 dt d\theta_2 .$$

Le nouveau paramétrage défini en (3.0) permet d'écrire ces estimateurs bayésiens respectivement sous la forme :

$$\theta_1^{(n)} + \frac{\tilde{\theta}_1^*}{\sqrt{k(n)}}, \qquad k(n) + \frac{\tilde{\tau}}{\Psi(n)}, \qquad \theta_2^{(n)} + \frac{\tilde{\theta}_2^*}{\sqrt{n-k(n)}}$$

avec

$$\tilde{\theta}_1^* = \frac{\displaystyle\int \theta_1^* q\left(\theta_1^{(n)} + \frac{\theta_1^*}{\sqrt{k(n)}}, \frac{k(n)}{n} + \frac{\tau}{n\Psi(n)}, \theta_2^{(n)} + \frac{\theta_2^*}{\sqrt{n-k(n)}}\right). \mathrm{Y}_n(\theta_1^*, \tau, \theta_2^*) d\theta_1^* d\tau d\theta_2^*}{\displaystyle\int q\left(\theta_1^{(n)} + \frac{\theta_1^*}{\sqrt{k(n)}}, \frac{k(n)}{n} + \frac{\tau}{n\Psi(n)}, \theta_2^{(n)} + \frac{\theta_2^*}{\sqrt{n-k(n)}}\right). \mathrm{Y}_n(\theta_1^*, \tau, \theta_2^*) d\theta_1^* d\tau d\theta_2^*}$$

et de même pour $\tilde{\tau}$ et $\tilde{\theta}_2^*$. Sous cette forme, on voit clairement que le théorème 4 de convergence du processus Y_n fournit la loi limite des estimateurs bayésiens :

THÉORÈME 5. — Sous la suite d'hypothèse $\mathrm{H}_1(\theta_1^{(n)}, k(n), \theta_2^{(n)})$, les estimateurs bayésiens des paramètres sont asymptotiquement indépendants et de lois respectives :

$$\sqrt{k(n)}(\tilde{\theta}_1 - \theta_1^{(n)}) \xrightarrow{\mathscr{L}} \mathcal{N}(0, \mathrm{I}(\theta_0)^{-1})$$
$$\sqrt{n-k(n)}(\tilde{\theta}_2 - \theta_2^{(n)}) \xrightarrow{\mathscr{L}} \mathcal{N}(0, \mathrm{I}(\theta_0)^{-1})$$
$$\Psi(n)(\tilde{k} - k(n)) \xrightarrow{\mathscr{L}} \xi_2$$

où ξ_2 est la variable aléatoire $\dfrac{\displaystyle\int \tau \cdot e^{\mathrm{W}(\tau) - \frac{|\tau|}{2}} d\tau}{\displaystyle\int e^{\mathrm{W}(\tau) - \frac{|\tau|}{2}} d\tau}$.

On constate que les estimateurs du maximum de vraisemblance et bayésiens sont asymptotiquement équivalents en ce qui concerne les paramètres statistiques classiques θ_1 avant rupture et θ_2 après; par contre, les lois de ξ_1 et de ξ_2 qui sont les lois limites des estimateurs du maximum de vraisemblance et bayésien de l'instant de rupture ne sont pas directement comparables. Cependant, Ibragimov-Khasminskii [*12*] ont obtenu les mêmes lois ξ_1 et ξ_2 comme lois limites des estimateurs de paramètre de discontinuité pour la dérive d'un brownien; ils ont montré que :

$$19{,}5 \simeq \mathrm{E}\xi_2^2 < \mathrm{E}\xi_1^2 = 26,$$

ce qui conduit à conclure, comme dans le cas d'Ibragimov-Khasminskii, que les estimateurs bayésiens sont ici meilleurs que l'estimateur du maximum de vraisemblance avec une fonction de perte quadratique. Malheureusement, ils présentent l'inconvénient de ne pas fournir une loi limite calculable autrement que par simulation. En revanche, la loi limite de l'estimateur du maximum de vraisemblance est de densité explicite et permet de définir des intervalles de confiance approchés. D'autre part, l'optimalité asymptotique des estimateurs bayésiens repose fondamentalement sur le critère du risque quadratique et on peut se demander si celui-ci est bien adapté dans ce cas particulier où la loi limite n'est pas gaussienne et si les intervalles de confiance construits à partir des différents estimateurs ne seraient pas équivalents.

3.3. Cohérence des procédures de test et d'estimation.

Il est facile de rendre cohérentes les méthodes de test d'existence d'une rupture et d'estimation des paramètres de la rupture : il suffit d'employer le test du rapport de vraisemblance renormalisé par une fonction φ et, lorsque l'hypothèse H_0 est rejetée, de considérer les estimateurs du maximum de cette vraisemblance renormalisée ; lorsqu'une rupture a effectivement eu lieu, ces estimateurs ont le même comportement que les estimateurs du maximum de vraisemblance ordinaire, en particulier la consistance de l'estimateur $\frac{\hat{k}}{n}$ de l'instant de rupture implique que les lois asymptotiques des estimateurs de tous les paramètres sont identiques.

4. CONCLUSION

On a vu que les théorèmes d'invariance sur le processus de vraisemblance démontrés dans un cadre général, ont permis de résoudre globalement les problèmes posés par l'utilisation des méthodes de vraisemblance dans les modèles de rupture et apportent en particulier une réponse aux questions posées en introduction :

— La statistique du test de vraisemblance n'a pas asymptotiquement de loi limite et il est absurde de la renormaliser globalement. En revanche, nous suggérons une pondération de la statistique par l'intermédiaire d'une fonction φ pénalisant les ruptures éventuelles aux bords. Le choix de cette fonction φ peut être fait dans une classe assez large et aucun critère d'optimalité ne peut être mis en évidence. Toutefois, cet arbitraire permettra à l'utilisateur de régler lui-même la sensibilité du test aux changements intervenant en début ou fin d'observation au gré de ses contraintes ou de son information *a priori* : si la modélisation a pour but la prévision (en économie par exemple), il est important de bien détecter une rupture en fin d'observation ; par contre dans un problème de détection de panne, on peut vouloir ne pas tenir compte de dérèglements éventuels se produisant en début d'observation où le système n'est pas encore stabilisé.

— Le comportement asymptotique des estimations des paramètres d'une rupture est caractérisée globalement par l'expression :

$$\frac{k(n-k)}{n} | \theta_2 - \theta_1 |^2 .$$

L'asymptotique adoptée ($d_n \to 0$) a permis de lever les inconvénients des calculs, par méthode de fluctuations, de la loi de l'estimateur de l'instant de rupture (calculs longs qui doivent être refaits dans chaque modèle particulier et qui n'ont été résolus que pour les modèles gaussiens et binomiaux [*8*]-[*9*]). Cette asymptotique signalée par Hinkley avait été mise en œuvre par Bhattacharya [*1*] sous des conditions restrictives sur le modèle ne permettant pas de traiter l'estimation par le maximum de vraisemblance.

Nous obtenons comme loi limite de l'instant de rupture une loi classique dans les problèmes avec discontinuité du modèle par rapport au paramètre. Comme Ibragimov-Khasminskii [*12*], nous constatons qu'un estimateur bayésien est de risque quadratique asymptotique inférieur.

Mini-guide du praticien :

Le test de changement de paramètre dans un modèle statistique classique (par exemples, changement de moyenne pour des variables gaussiennes ; changement de paramètre d'échelle pour des variables gaussiennes ou exponentielles, ...) donnera de bons résultats en pondérant la statistique de vraisemblance aux bords de l'intervalle d'observation par l'intermédiaire d'une fonction φ : le seuil peut alors être approximé en utilisant la loi limite qui est celle de :

$$\operatorname*{Sup}_{t \in]0,1[} \varphi(t)\varphi(1-t) \frac{{}^t[\mathrm{W}_p(t) - t\mathrm{W}_p(1)][\mathrm{W}_p(t) - t\mathrm{W}_p(1)]}{2t(1-t)}$$

Cette loi est bien connue lorsque $\varphi(t) = t$ et $p = 1$, elle est tabulée dans [*5*] pour $p > 1$ et est tabulable pour d'autres fonctions φ.

Les lois limites des estimateurs du maximum de vraisemblance (ou de vraisemblance renormalisée) définies par le théorème 3 permettent de calculer des intervalles de confiance approchés pour chacun des paramètres d'une rupture ; ceux des paramètres du modèle avant et après rupture sont obtenus classiquement à partir de la loi de Gauss en remplaçant, si nécessaire, l'information de Fisher par une estimation ; pour l'instant de rupture, la densité explicite de la variable ξ_1 donnée au théorème 3 permet de dresser une table.

L'étude faite ici pour une rupture intervenant dans un modèle statistique classique se généralise à d'autres situations (modèles de régression linéaire, modèles autorégressifs, processus de Poisson... voir [5]).

BIBLIOGRAPHIE

[1] P. K. BHATTACHARYA, « *Estimation of change-point in the distribution of independant random variables* ». Preprint of the University of Arizona, 1981.

[2] P. K. BHATTACHARYA, P. J. BROCKWELL, « The minimum of an additive process with applications to signal estimation and storage theory », *Zeit. Wahr. und veriv. Gebiete*, t. **37**, 1976, p. 51-75.

[3] Z. W. BIRNBAUM, A. W. MARSALL, « Some multivariate Chebyshev inequalities with extensions to continuous parameter processes », *A. M. S.*, t. **32**, 1961, p. 687-703.

[4] J. DESHAYES, D. PICARD, « Principe d'invariance sur le processus de vraisemblance », *Annales de l'I. H. P.*, t. **20**, 1984, p. 1-20.

[5] J. DESHAYES, D. PICARD, « *Ruptures de modèles en statistique* ». *Thèses d'État*. Université Paris-Sud, 1983.

[6] P. I. FEDER, « The log-likelihood ratio in segmented regression », *Ann. Stat.*, t. **3**, 1975, p. 84-97.

[7] D. M. HAWKINS, « Testing a sequence of observations for a shift in location », *J. A. S. A.*, t. **72**, 1977, p. 180-186.

[8] D. V. HINKLEY, « Inference about the change-point in a sequence of random variables », *Biometrika*, t. **57**, 1970, p. 1-17.

[9] D. V. HINKLEY, E. A. HINKLEY, « Inference about the change-point in a sequence of binomial variables », *Biometrika*, t. **57**, 1970, p. 477-488.

[10] D. V. HINKLEY, P. CHAPMAN, G. RUNGER, « *Change-point problems* », Invited conference at the 11th European Meeting of Statisticians, Brighton, 1980.

[11] I. A. IBRAGIMOV, R. Z. KHASMINKII, « Asymptotic behaviour of statistical estimators in the smooth case-I-study of the likelihood ratio ». *Theory of Prob. and Applic.*, t. **17**, 1972, p. 445-462.

[12] I. A. IBRAGIMOV, R. Z. KHASMINSKII, « Statistical estimation-asymptotic theory », Springer-Verlag. *Applications of Mathematics*, t. **16**, 1981.

[13] L. LECAM, « *Limits of experiments* », Proceedings of the VIth Berkeley Symposium-I, 1972, p. 245-261.

[14] I. B. MCNEIL, « Tests for change of parameter at unknown time and distributions of some related functionals on Brownian motion », *Ann. Stat.*, t. **2**, 1974, p. 950-962.

[15] R. E. QUANDT, « Tests of the hypothesis that a linear system obeys two separate regimes », *J. A. S. A.*, t. **55**, 1960, p. 324-330.

[16] L. A. SHEPP, « The joint density of the maximum and its location for a Wiener process with drift », *J. of Appl. Prob.*, t. **16**, 1979, 423-427.

(Manuscrit reçu le 2 mars 1984)

(reçu corrigé le 10 juillet 1984)