



HAL
open science

Estimation efficace pour un processus autorégressif exponentiel a densité discontinue

Adrian E. Raftery

► **To cite this version:**

Adrian E. Raftery. Estimation efficace pour un processus autorégressif exponentiel a densité discontinue. Annales de l'ISUP, 1980, XXV (1-2), pp.65-91. <hal-04082146>

HAL Id: hal-04082146

<https://hal.science/hal-04082146v1>

Submitted on 26 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

ESTIMATION EFFICACE POUR UN PROCESSUS AUTOREGRESSIF EXPONENTIEL A DENSITE
DISCONTINUE

par Adrian E. RAFTERY

Laboratoire de Statistique Mathématique, Université de Paris VI,
5230 PARIS CEDEX 05.

Résumé :

On considère un modèle linéaire markovien pour des suites de variables aléatoires exponentielles corrélées, qui est analytiquement traitable et facile à simuler. La densité conditionnelle des observations est discontinue. On propose un estimateur des paramètres dont on démontre la consistance forte et l'efficacité asymptotique. On donne également sa relation avec l'estimateur de maximum de vraisemblance. Pour un des paramètres l'estimateur est "superefficace" et pour les autres il est asymptotiquement normal.

Summary :

A linear markovian model for series of correlated exponential random variables is considered. This is analytically tractable and easy to generate, and the conditional probability density of the observations is discontinuous. An estimator is proposed for the parameters, and is shown to be strongly consistent and asymptotically efficient. Its relationship with the maximum likelihood estimator is given. For one of the parameters the estimator is "superefficient" and for the others it is asymptotically normal.

MOTS-CLES : variables aléatoires exponentielles corrélées ; densité discontinue ; consistance forte ; efficacité asymptotique ; superefficacité.

KEY-WORDS : correlated exponential random variables ; discontinuous density ; strong consistency ; asymptotic efficiency ; superefficiency.

1. INTRODUCTION

Nous considérons dans cet article l'estimation des paramètres dans des modèles autorégressifs pour des séries temporelles à loi marginale exponentielle et pour des processus ponctuels stationnaires.

La plus grande partie des travaux actuels sur les séries temporelles est basée sur des modèles linéaires (Box-Jenkins, méthodes de filtrage, etc.). Or, ces modèles ont en général une loi marginale qui est approximativement gaussienne, comme l'a démontré Mallows (1967).

Enfin, une grande partie des techniques actuelles pour l'analyse statistique des séries chronologiques suppose implicitement que la loi des innovations est gaussienne (identification par les autocorrélations, vraisemblance gaussienne, etc.). souvent cette hypothèse n'est pas vérifiée : par exemple une étude récente de Nelson et Granger (1978) sur 21 séries chronologiques économiques a trouvé que dans plus de deux-tiers des cas, l'hypothèse de normalité était fortement rejetée même après application de la transformation de Box-Cox. On travaille beaucoup sur des méthodes pour surmonter cette difficulté dans le cas non-gaussien sans sacrifier la commodité mathématique du modèle gaussien en "adaptant" les données, par exemple par des transformations ou des méthodes robustes.

Or, souvent la loi marginale est elle-même importante. Par exemple, en hydrologie stochastique on s'intéresse souvent à la probabilité que le débit d'une rivière va dépasser une certaine valeur (Lawrance et Kottegoda (1977)). De même, si on analyse des séries de vitesses de vents pour prévoir le rendement de systèmes d'énergie éolienne, on se demande pendant combien de temps la vitesse va dépasser un certain seuil "utile".

Dans un processus ponctuel les intervalles entre les événements sont souvent corrélés et on a besoin de modèles simples pour représenter cette situation. Par exemple en fiabilité on s'attendrait souvent à ce que les intervalles entre les

annes successives d'une même machine soient corrélés : ceci est vérifié pour un ordinateur dans Cox et Lewis (1966). Des files d'attente corrélées ont été étudiées par Jacobs (1978) et les travaux de Lewis et Shedler (1977) sur les exceptions et les demandes dans un système informatique illustrent bien ce besoin.

Dans la section 2 nous considérons les modèles qui ont été proposés pour répondre à ces besoins, en particulier le modèle linéaire markovien de Lawrance (1978) et nous rappelons quelques propriétés de ce modèle. Dans la section 3 nous donnons des estimateurs efficaces pour le modèle et nous énonçons les théorèmes qui décrivent leur comportement asymptotique, qui sont démontrés dans la section 4. Dans la section 5 nous donnons quelques généralisations de ces résultats. Le symbole ■ indique la fin d'une démonstration.

MODELES EXPONENTIELS AUTOREGRESSIFS

On peut considérer que des modèles pour les applications ci-dessus devraient avoir les propriétés suivantes :

-) être faciles à construire en fonction de leur loi marginale et de leur structure d'autocorrélation,
-) utiliser le plus petit nombre de paramètres possible, et
-) être analytiquement traitables et simples à générer sur ordinateur ; il serait donc préférable qu'ils soient linéaires.

En principe, bien sûr, on peut facilement construire une suite de variables aléatoires à loi marginale exponentielle avec dépendance markovienne du moment que l'on dispose d'une variable aléatoire exponentielle bivariée. Les premiers efforts pour ce faire, dûs à Wold (1948) et à Cox (1955) n'ont pas néanmoins donné des résultats traitables. Les variables aléatoires exponentielles bivariées qui ont été découverts depuis lors, par exemple Marshall et Olkin (1967), Downton (1970), Hawkes (1972) et Gaver (1972), ne semblent pas non plus assez simples pour fournir des modèles traitables. De même les nombreux modèles de processus ponctuels qui ont été

le plus souvent développés pour des applications particulières, ne sont pas en général des modèles linéaires et sont souvent peu maniables (voir Lawrance (1972)).

Le succès des modèles ARMA suggère qu'on essaie de les adapter à notre problème. Un premier effort pour ce faire est dû à Linhart (1970) et Linhart et Zucchini (1972) qui ont développé un processus gamma avec la structure de corrélation ARMA ; mais leur modèle n'était pas linéaire et on l'a trouvé difficile à manier. Gaver et Lewis (1980), dans un article qui a ses origines vers 1974, ont considéré l'équation autorégressive d'ordre 1 standard :

$$X_n = \rho X_{n-1} + \varepsilon_n, \quad n \in \mathbb{Z}$$

où $|\rho| < 1$ et $\{\varepsilon_n\}_{n \in \mathbb{Z}}$ est une suite de variables aléatoires indépendantes de même loi. Ils ont trouvé la loi des ε_n pour que X_n ait loi marginale exponentielle avec moyenne m de la façon suivante :

Soient $\phi_X(t) = E(\exp(itX_n))$ et $\phi_\varepsilon(t) = E(\exp(it\varepsilon_n))$. Alors

$$\phi_X(t) = \phi_X(\rho t) \phi_\varepsilon(t)$$

$$\begin{aligned} \text{Donc } \phi_\varepsilon(t) &= \phi_X(t) / \phi_X(\rho t) \\ &= (1 - imt)^{-1} / (1 - im\rho t)^{-1} \\ &= \rho + (1 - \rho) (1 - imt)^{-1} \end{aligned}$$

$$\text{Donc } \varepsilon_n = \begin{cases} 0 & \text{avec pr. } \rho \\ E_n & \text{avec pr. } 1 - \rho \end{cases}$$

où E_n est exponentiel avec moyenne m . Il est clair que pour que X_n soit positif il est nécessaire que $\rho \geq 0$. Ils obtiennent donc le modèle suivant :

$$(2.1) \quad X_n = \begin{cases} \rho X_n & \text{avec pr. } \rho \\ \rho X_n + E_n & \text{avec pr. } 1 - \rho \end{cases}$$

Ce modèle a plusieurs avantages : il est facile à simuler, analytiquement traitable et sa structure d'autocorrélation est indépendante de sa loi marginale ; il a la structure d'autocorrélation du modèle AR(1) standard avec $\text{Corr}(X_n, X_{n+r}) = \rho^r$.

Il peut également être généralisé à une classe de modèles pour des suites de variables aléatoires exponentielles ayant la structure d'autocorrélation du processus ARMA (p,q). Les détails de ces modèles EARMA(p,q) sont donnés par Lawrance et Lewis (1977), Jacobs et Lewis (1977), et Lawrance et Lewis (1980). Il peut aussi être généralisé à d'autres lois marginales - celles qui forment la classe L (Lawrance (1978)).

Le modèle (2.1) a également des désavantages qui sont très graves en pratique. Puisque les ϵ_n sont des variables aléatoires positives, X_n est contraint à être $\geq \rho X_{n-1}$, une restriction qui n'est valable que dans des cas très spéciaux. En plus, la loi conditionnelle de X_n a un atome à ρX_{n-1} , ce qui paraît hautement irréaliste en pratique. Pour résoudre ces problèmes Lawrance (1978) a proposé le modèle alternatif suivant :

$$(2.2) \quad X_n = \epsilon_n + \begin{cases} 0 & \text{avec pr. } 1 - p \\ \alpha X_{n-1} & \text{avec pr. } p \end{cases}$$

où $0 \leq \alpha < 1$, $0 \leq p < 1$ pour $n \in \mathbb{Z}$. On voit que

$$\phi_X(t) = \phi_\epsilon(t) ((1-p) + p \phi_X(\alpha t))$$

Donc

$$\begin{aligned} \phi_\epsilon(t) &= \phi_X(t) / ((1-p) + p \phi_X(\alpha t)) \\ &= (1-\alpha) (1-\alpha+\alpha p)^{-1} (1-imt)^{-1} + \alpha p (1-\alpha+\alpha p)^{-1} (1-im\alpha(1-p)t)^{-1} \end{aligned}$$

quand X_n est exponentiel avec moyenne m. Donc

$$\epsilon_n = \begin{cases} E_n & \text{avec pr. } (1-\alpha) / (1-\alpha+\alpha p) \\ \alpha(1-p)E_n & \text{avec pr. } \alpha p / (1-\alpha+\alpha p) \end{cases}$$

Le processus (2.2) a les avantages de (2.1) sans en avoir les inconvénients.

Comme nous le verrons il a en plus l'avantage que α peut être estimé de façon "superefficace". Il est strictement stationnaire par construction, il est markovien d'ordre 1, et il a la structure d'autocorrélation du modèle AR(1) avec $\text{Corr}(X_n, X_{n+r}) = (\rho\alpha)^r$. On voit donc que le paramètre α n'est pas nécessaire pour préciser la structure d'autocorrélation ; sa présence permet néanmoins de représenter simplement une grande variété de types de transition dans la chaîne de Markov (2.2).

Pour simplifier l'exposition nous allons réduire le nombre de paramètres de trois à deux en mettant $\alpha = p$, ce qui donne le modèle suivant :

$$(2.3) \quad X_n = \varepsilon_n + \begin{cases} 0 & \text{avec pr. } 1-p \\ pX_{n-1} & \text{avec pr. } p \end{cases}$$

$$0 \leq p < 1, \quad 0 < m < \infty, \quad n \in \mathbb{Z}$$

$$\text{où } \varepsilon_n = \begin{cases} E_n & \text{avec pr. } (1-p)(1-p+p^2)^{-1} \\ p(1-p)E_n & \text{avec pr. } p^2(1-p+p^2)^{-1} \end{cases}$$

où les E_n sont indépendants de même loi exponentielle avec moyenne m , et E_n est indépendant de $\{X_i : i \leq n-1\} \forall n$. Dans la section 5 nous verrons que les résultats obtenus pour (2.3) tiennent aussi pour (2.2).

Dans Raftery (1980) nous avons étudié le comportement asymptotique de (2.3). Soient $F_m = \sigma(\dots, X_{m-1}, X_m)$, $G_m = \sigma(X_m, X_{m+1}, \dots)$ et $f \in L^2(F_m)$, $g \in L^2(G_{m+k})$ au sens que $E|f^2| < \infty$; $E|g^2| < \infty$. Alors nous avons démontré qu'il existe une constante C_1 telle que :

$$(2.4) \quad |E(fg) - E(f)E(g)| \leq C_1 p^k E(f^2)^{1/2} E(g^2)^{1/2} \text{ pour } k \geq -\log 15/\log p$$

c'est-à-dire que (2.3) est fortement mélangé au sens de Rosenblatt (1971). En plus, par la section 2 de l'article cité il existe un événement C tel que si $A \in F_n$ et $B \in G_{m+k}$ alors

$$(2.5) \quad |P(B|A) - P(B)| = p^k |P((B|A)|C) - P(B|C)| \leq 2p^k$$

ce qui implique que (2.3) est ϕ -mélangeant au sens de Billingsley (1968), p. 166 avec $\sum_{k=-\infty}^{\infty} k^2 \phi(k)^{1/2} < \infty$. La plupart des résultats de Billingsley (1968), ch. 4 sont donc vérifiés pour ce processus. Ceci est vrai aussi pour le processus (2.2).

3. ESTIMATION ASYMPTOTIQUEMENT EFFICACE

Si on suppose que X_1 est exponentiel avec moyenne m , alors le processus (2.3) est ergodique et la densité de probabilité d'une réalisation $X = (X_1, \dots, X_n)$ est

$$(3.1) \quad L_n(X|\theta) = m^{-1} \exp(-X_1|m) \prod_{k=2}^n f_{k,\theta}$$

où $\theta = (m, p)$ est le paramètre du processus et $f_{k,\theta}$ est la densité conditionnelle de X_k étant donné X_{k-1} sous θ . (3.1) résulte donc de la propriété markovienne de

(2.3). On a

$$f_{k,\theta} = m^{-1} (1-p+p^2)^{-1} \{ (1-p)^2 \exp(-X_k/m) + p \exp(-X_k/mp(1-p)) \\ + \delta_k(p) [p(1-p) \exp(-(X_k - pX_{k-1})/m) + p^2(1-p)^{-1} \\ \exp(-(X_k - pX_{k-1})/mp(1-p))] \}$$

$$\text{où } \delta_k(p) = \begin{cases} 1 & \text{si } X_k \geq pX_{k-1} \\ 0 & \text{sinon} \end{cases}$$

On voit que $f_{k,\theta}$ est discontinue en p et la théorie classique de l'estimation asymptotique ne s'applique pas, puisque les conditions de régularité qu'elle impose sont violées et que les observations ne sont pas indépendantes.

Dans Raftery (1980) nous avons considéré l'estimateur de maximum de vraisemblance (EMV), défini comme une valeur $\tilde{\theta}$ de θ telle que

$$L_n(X|\tilde{\theta}) \geq L_n(X|\theta) \quad \forall \theta \neq \tilde{\theta}$$

Nous avons démontré que quand on sait a priori que $m_0 < M$ pour un M donné alors cet estimateur est consistant au sens que

$$\lim_{n \rightarrow \infty} P_{\theta_0} \left[\frac{1}{n} (\log L_n(X|\theta_0) - \log L_n(X|\tilde{\theta})) > 0 \right] = 1 \quad \forall \theta \neq \theta_0$$

où $\theta_0 = (m_0, p_0)$ est la vraie valeur du paramètre.

Une petite extension des arguments permet de démontrer que l'EMV est fortement consistant, au sens habituel, sans restriction a priori sur m_0 , c'est-à-dire que quand la taille de l'échantillon augmente, l'EMV tend vers θ_0 p.s. En effet, nous

avons démontré dans l'article cité que p.s.

$$\begin{aligned} & \lim_{n \rightarrow \infty} n^{-1} (\log L_n(X|\theta) - \log L_n(X|\theta_0)) \\ &= \lim_{n \rightarrow \infty} n^{-1} \sum_{k=1}^n E_{k-1, \theta_0} \log (p_{k, \theta} / p_{k, \theta_0}) \\ &= E_{\theta_0} \log (p_{k, \theta} / p_{k, \theta_0}) < 0 \end{aligned}$$

Donc

$$(3.2) \quad \lim_{n \rightarrow \infty} n^{-1} (\log L_n(X|\theta) - \log L_n(X|\theta_0)) < 0 \quad \text{p.s.}$$

Il suffit maintenant de démontrer que si $\tilde{\theta}_n$ est l'EMV basé sur un échantillon de taille n , alors

$$\left| \lim_{n \rightarrow \infty} \tilde{\theta}_n - \theta_1 \right| < \varepsilon \quad \text{p.s.} \quad \forall \varepsilon > 0$$

Supposons qu'il existe un point d'adhérence θ' de la suite $\{\tilde{\theta}_n\}$ tel que $|\theta' - \theta_0| > \varepsilon$. Alors

$$\sup_{|\theta - \theta_0| \geq \varepsilon} L_n(X|\theta) \geq L_n(X|\tilde{\theta}_n) \text{ pour un nombre infini de } n.$$

$$\text{Donc } \sup_{|\theta - \theta_0| \geq \varepsilon} L_n(X|\theta) \geq L_n(X|\theta_0) \text{ pour un nombre infini de } n,$$

ce qui a probabilité zéro par (3.2). On a donc

$$(3.3) \quad \lim_{n \rightarrow \infty} \tilde{\theta}_n = \theta_0 \quad \text{p.s.}$$

Nous considérons maintenant l'estimateur alternatif $\hat{\theta} = (\hat{m}, \hat{p})$ défini par

$$(3.4) \quad \begin{cases} \hat{p} \text{ est la valeur de } p \text{ qui maximise} \\ D_n(\theta) = \int_{p-r/n}^{p+r/n} L_n(X|\theta) dp \\ \text{où } r \text{ est un nombre fixe } 0 < r < \infty \\ \text{et } \hat{m} \text{ est la valeur de } m \text{ qui maximise } L_n(X|\theta) \end{cases}$$

Pour simplifier l'exposition nous nous permettons l'abus de notation qui consiste à utiliser p comme variable d'intégration. On reconnaît \hat{p} comme l'estima-

eur de maximum de probabilité (EMP) de p par rapport à un intervalle, et \hat{m} comme l'EMV de m .

L'EMP peut être considéré l'estimateur de Bayes avec loi a priori uniforme et fonction de perte $\psi(\hat{p}, p_0) = I_{[-r/n, r/n]}(\hat{p} - p_0)$. On trouve une exposition détaillée de l'estimation de MP dans Weiss et Wolfowitz (1974). En fait il s'ensuit de Heyde (1978) que puisque $f_{k, \theta}$ est trois fois dérivable en m , alors

$$\lim_{n \rightarrow \infty} P_{\theta_0} [\sqrt{n}(\hat{m} - m_0) \in [-r, r]] \geq \lim_{n \rightarrow \infty} P_{\theta_0} [\sqrt{n}(T_n - m_0) \in [-r, r]]$$

pour tout estimateur T_n de m , et nous allons voir que ceci implique que (3.4) est en fait asymptotiquement équivalent à l'EMP de θ .

On note que puisque $L_n(X|\theta)$ est dérivable en m , \hat{m} est une solution de l'équation

$$3.5) \quad \frac{\partial \log L_n(X|\theta)}{\partial m} = 0$$

Nous énonçons maintenant les résultats qui décrivent le comportement asymptotique de l'estimateur (3.4).

Théorème 1 : $\hat{\theta}$ est un estimateur fortement consistant de θ , c'est-à-dire que si $\hat{\theta}_n$ est l'estimateur défini par (3.4) pour un échantillon de taille n , alors

$$3.6) \quad \lim_{n \rightarrow \infty} \hat{\theta}_n = \theta_0 \quad \text{p.s}$$

Théorème 2 : (i) \hat{p} est un estimateur asymptotiquement efficace de p au sens que

$$3.7) \quad \lim_{n \rightarrow \infty} P_{\theta_0} [-r/n \leq \hat{p} - p_0 \leq r/n] = 1$$

(ii) $\sqrt{n}(\hat{m} - m_0)$ suit asymptotiquement une loi gaussienne centrée avec variance $E \left(\left(\frac{\partial \log f_k}{\partial m} \Big|_{\theta_0} \right)^{-2} \right)$, c'est-à-dire que \hat{m} est un estimateur asymptotiquement

normal et asymptotiquement efficace de m .

(iii) \hat{m} et \hat{p} sont asymptotiquement indépendants et la loi limite de \hat{m} est la même que la loi limite de l'EMV de m quand p_0 est connu, et la loi limite de \hat{p} est la même que la loi limite de l'EMP de p quand m_0 est connu.

Corollaire : Si $\tilde{\theta}$ est l'EMV, alors il existe une constante C_2 telle que pour n suffisamment grand on aît en probabilité

$$(3.8) \quad (\hat{m} - \tilde{m}) = C_2(\hat{p} - \tilde{p}) + o(\hat{p} - \tilde{p}), \text{ et}$$

$$(3.9) \quad (\hat{p} - \tilde{p}) = O(n^{-1})$$

4. DEMONSTRATIONS DES RESULTATS

Tous les résultats sont démontrés à l'aide d'un développement de Taylor de la log - vraisemblance. Le théorème 1 est une conséquence de la consistance de l'EMV fait que certaines dérivées de la log-vraisemblance sont finies partout. La démonstration du théorème 2 utilise en plus l'existence de bornes locales pour les dérivés d'ordres deux et trois, la loi forte des grands nombres pour les processus, et pour la partie (ii), le théorème central limite pour les martingales de Lindeberg-Levy énoncé par Billingsley (1961).

Pour $k \geq 2$ quand $p \geq p_0$ nous pouvons écrire

$$\begin{aligned} \log f_{k,\theta} = & a_{0k} + (p-p_0) a_{1k} + (m-m_0) a_{2k} + \frac{1}{2} (p-p_0)^2 a_{3k} \\ & + (p-p_0)(m-m_0) a_{4k} + \frac{1}{2} (m-m_0)^2 a_{5k} - (\delta_k(p_0) - \delta_k(p)) \alpha_k(\theta) \end{aligned}$$

$$\text{où } a_{0k} = \log f_{k,\theta_0} \qquad a_{3k} = \left. \frac{\partial^2 \log f'_k}{\partial p^2} \right|_{\theta'}$$

$$a_{1k} = \left. \frac{\partial \log f_k}{\partial p} \right|_{\theta_0} \qquad a_{4k} = \left. \frac{\partial^2 \log f'_k}{\partial p \partial m} \right|_{\theta'}$$

$$a_{2k} = \left. \frac{\partial \log f_k}{\partial m} \right|_{\theta_0} \qquad a_{5k} = \left. \frac{\partial^2 \log f'_k}{\partial m^2} \right|_{\theta'}$$

$$\text{où } f'_{k,\theta} = (1 - \delta_k(p_0)) g_{1k,\theta} + \delta_k(p_0) (g_{1k,\theta} + g_{2k,\theta})$$

$$\theta' = (m', p') \text{ où } p_0 < p' < p \text{ et } m' \text{ est entre } m \text{ et } m_0$$

$$\alpha_k(\theta) = \log(g_{1k,\theta} + g_{2k,\theta}) - \log g_{1k,\theta}$$

en écrivant f_k pour $f_{k,\theta}$ quand le contexte ne permet pas d'ambiguïté, en notant $f_k = g_{1k} + \delta_k(p)g_{2k}$ et en adoptant la convention que $h|_\theta$ veut dire la fonction h évaluée quand le paramètre prend la valeur θ . Les i_k souscrits sont bien sûr des couples et non des produits.

Puisque $\alpha_k(\theta)$ est régulier en θ nous pouvons écrire

$$(\delta_k(p_0) - \delta_k(p)) \alpha_k(\theta) = a_{6k} + (p-p_0) a_{7k} + (m-m_0) a_{8k}$$

où $a_{6k} = (\delta_k(p_0) - \delta_k(p)) \alpha_k(\theta_0)$

$$a_{7k} = (\delta_k(p_0) - \delta_k(p)) \left. \frac{\partial \alpha_k}{\partial p} \right|_{\theta'}$$

$$a_{8k} = (\delta_k(p_0) - \delta_k(p)) \left. \frac{\partial \alpha_k}{\partial m} \right|_{\theta'}$$

Donc quand $p \geq p_0$ on a

$$(4.1) \quad n^{-1} \log L_n(X|\theta) = d_0 + (p-p_0)d_1 + (m-m_0)d_2 + \frac{1}{2} (p-p_0)^2 d_3 + (p-p_0)(m-m_0)d_4 + \frac{1}{2} (m-m_0)^2 d_5 - (p-p_0)d_6 - (p-p_0)^2 d_7 - (p-p_0)(m-m_0)d_8$$

où $d_0 = n^{-1} (-\log m_0 - X_1/m_0 + \sum_{k=2}^n a_{0k})$

$$d_2 = n^{-1} (m_0^{-1} (X_1/m_0 - 1) + \sum_{k=2}^n a_{2k})$$

$$d_5 = n^{-1} (-(m')^2 (2X_1/m' - 1) + \sum_{k=2}^n a_{5k})$$

$$d_i = n^{-1} \sum_{k=2}^n a_{ik}, \quad i = 1, 3, 4$$

$$(p-p_0) d_i = n^{-1} \sum_{k=2}^n a_{ik}, \quad i = 6, 7, 8$$

De même, quand $p < p_0$ on a pour $k \geq 2$:

$$\log f_{k,\theta} = a_{0k} + (p-p_0) a_{1k} + (m-m_0) a_{2k} + \frac{1}{2} (p-p_0)^2 a'_{3k} + (p-p_0)(m-m_0) a'_{4k} + \frac{1}{2} (m-m_0)^2 a'_{5k} + a'_{6k} + (p-p_0) a'_{7k} + (m-m_0) a'_{8k}$$

$$\begin{aligned} \text{où } a'_{3k} &= \left. \frac{\partial^2 \log f'_k}{\partial p^2} \right|_{\theta''} & a'_{6k} &= (\delta_k(p) - \delta_k(p_0)) \alpha_k(\theta_0) \\ a'_{4k} &= \left. \frac{\partial^2 \log f'_k}{\partial p \partial m} \right|_{\theta''} & a'_{7k} &= (\delta_k(p) - \delta_k(p_0)) \frac{\partial \alpha_k}{\partial p} \Big|_{\theta''} \\ a'_{5k} &= \left. \frac{\partial^2 \log f'_k}{\partial m^2} \right|_{\theta''} & a'_{8k} &= (\delta_k(p) - \delta_k(p_0)) \frac{\partial \alpha_k}{\partial m} \Big|_{\theta''} \end{aligned}$$

où $\theta'' = (m'', p'')$, m'' est entre m et m_0 et $p < p'' < p_0$.

Donc quand $p < p_0$ on a

$$(4.2) \quad \begin{aligned} n^{-1} \log L_n(X|\theta) &= d_0 + (p-p_0)d_1 + (m-m_0)d_2 + \frac{1}{2} (p-p_0)^2 d'_3 \\ &+ (p-p_0)(m-m_0)d'_4 + \frac{1}{2} (m-m_0)^2 d'_5 + (p-p_0)d'_6 + (p-p_0)^2 d'_7 \\ &+ (p-p_0)(m-m_0)d'_8 \end{aligned}$$

$$\begin{aligned} \text{où } d'_i &= n^{-1} \sum_{k=2}^n a'_{ik}, & i &= 3, 4 \\ d'_5 &= n^{-1} (-(m'')^{-2} (2X_1/m''-1) + \sum_{k=2}^n a'_{5k}) \\ (p-p_0)d'_0 &= n^{-1} \sum_{k=2}^n a'_{ik}, & i &= 6, 7, 8 \end{aligned}$$

De même pour $k \geq 2$

$$\begin{aligned} \frac{\partial \log f_k}{\partial m} \Big|_{\theta_0} &= a_{2k} + (m-m_0)a_{9k} + (p-p_0) a_{10k} + \frac{1}{2} (m-m_0)^2 a_{11k} \\ &+ (m-m_0)(p-p_0) a_{12k} + \frac{1}{2} (p-p_0)^2 a_{13k} + a_{14k} \\ \text{où } a_{9k} &= \left. \frac{\partial^2 \log f_k}{\partial m^2} \right|_{\theta_0} & a_{12k} &= \left. \frac{\partial^3 \log f'_k}{\partial m^2 \partial p} \right|_{\theta''''} \\ a_{10k} &= \left. \frac{\partial^2 \log f_k}{\partial p \partial m} \right|_{\theta_0} & a_{13k} &= \left. \frac{\partial^3 \log f'_k}{\partial m \partial p^2} \right|_{\theta''''} \\ a_{11k} &= \left. \frac{\partial^3 \log f'_k}{\partial m^3} \right|_{\theta''''} & a_{14k} &= (\delta_k(p) - \delta_k(p_0)) \frac{\partial \alpha_k}{\partial m} \Big|_{\theta''''} \end{aligned}$$

où $\theta'''' = (m''', p''')$, m''' est entre m et m_0 , et p''' est entre p et p_0 .

Donc pour $k \geq 2$

$$(4.3) \quad \frac{1}{n} \left. \frac{\partial \log L_n(X|\theta)}{\partial m} \right|_{\theta} = d_2 + (m-m_0) d_9 + (p-p_0) d_{10} + \frac{1}{2} (m-m_0)^2 d_{11} \\ + (m-m_0)(p-p_0) d_{12} + \frac{1}{2} (p-p_0)^2 d_{13} + (p-p_0) d_{14}$$

$$\text{où } d_9 = n^{-1} (m_0^{-2} - 2 X_1/m_0^2 + \sum_{k=2}^n a_{9k}) \\ d_{11} = n^{-1} (4X_1/m_0^3 - 2 m_0^{-2} + \sum_{k=2}^n a_{11k}) \\ (p-p_0) d_{14} = n^{-1} \sum_{k=2}^n a_{14k} \\ d_i = n^{-1} \sum_{k=2}^n a_{ik}, \quad i = 10, 12, 13$$

Dans tout ce qui précède, les θ' , θ'' , θ''' ne sont pas nécessairement tous évalués au même point.

Avant de démontrer le théorème 1, nous avons besoin du résultat technique suivant :

Lemme 1 : Pour tout θ , $\theta_0 \in (0, \infty) \times [0, 1)$ on a

- (i) $|E_{\theta_0}(a_{ik})| < \infty$ et $E_{\theta_0}(a_{ik}^2) < \infty$, $i = 1, 2, 3$
- (ii) $|E_{\theta_0}(a'_{3k})| < \infty$ et $E_{\theta_0}((a'_{3k})^2) < \infty$
- (iii) $|(p-p_0)^{-1} E_{\theta_0}(a_{ik})| < \infty$ et $(p-p_0)^{-2} E_{\theta_0}(a_{ik}^2) < \infty$, $i = 6, 7$
- (iv) $|(p-p_0)^{-1} E_{\theta_0}(a'_{ik})| < \infty$ et $(p-p_0)^{-2} E_{\theta_0}((a'_{ik})^2) < \infty$, $i = 6, 7$

Démonstration : On note que pour a_{6k} et a'_{6k} le résultat suit de (3.4) de Raftery (1980) et de l'inégalité qui le précède. On va se contenter ici de vérifier le lemme pour a_{1k} ; la vérification est analogue dans les autres cas.

$$\text{On a } |E_{\theta_0}(a_{1k})| = |E_{\theta_0} \left(\frac{\partial f_k / \partial p}{f_k} \Big|_{\theta_0} \right)| \\ = \left| \int_0^\infty \int_0^\infty \frac{\partial f_k}{\partial p} \Big|_{\theta_0} dx_k m_0^{-1} \exp(-x_{k-1}/m_0) dx_{k-1} \right| \\ = p_0 < \infty$$

Aussi

$$(4.4) \quad E_{\theta_0} (a_{1k}^2) = \int_0^\infty \int_0^\infty \frac{(\partial f_k / \partial p)^2}{f_k} \Big|_{\theta_0} dX_k m_0^{-1} \exp(-X_{k-1} / m_0) dX_{k-1}$$

Mais

$$(4.5) \quad \int_0^\infty \frac{(\partial f_k / \partial p)^2}{f_k} \Big|_{\theta_0} dX_k$$

$$= \int_0^{p_0 X_{k-1}} \frac{(\partial g_{1k} / \partial p)^2}{f_k} \Big|_{\theta_0} dX_k + \int_{p_0 X_{k-1}}^\infty \frac{(\partial(g_{1k} + g_{2k}) / \partial p)^2}{f_k} \Big|_{\theta_0} dX_k$$

$$\leq 2 \int_0^{p_0 X_{k-1}} \frac{(\partial g_{1k} / \partial p)^2}{f_k} \Big|_{\theta_0} dX_k + 2 \int_{p_0 X_{k-1}}^\infty \frac{(\partial g_{1k} / \partial p)^2 + (\partial g_{2k} / \partial p)^2}{f_k} \Big|_{\theta_0} dX_k$$

puisque $(a+b)^2 \leq 2(a^2+b^2)$ et $f_k > 0$

$$= 2 \int_0^\infty \frac{(\partial g_{1k} / \partial p)^2}{f_k} \Big|_{\theta_0} dX_k + 2 \int_{p_0 X_{k-1}}^\infty \frac{(\partial g_{2k} / \partial p)^2}{f_k} \Big|_{\theta_0} dX_k$$

Or, des calculs simples montrent que

$$\frac{(\partial g_{1k} / \partial p)^2}{f_k} \Big|_{\theta_0} \leq m_0^{-1} \{ (1-p_0)^2 \exp(-X_k / m_0) - 2(C_3 + C_4 X_k / m_0) \exp(-X_k / m_0 p_0 (1-p_0)) + (1-p_0)^{-2} (C_3 + C_4 X_k / m_0)^2 \exp(-\frac{X_k}{m_0} (\frac{2}{p_0(1-p_0)} - 1)) \}$$

$$\text{où } C_3 = 1 + p_0 - 2 p_0^2$$

$$C_4 = (1-2p_0) (1-p_0)$$

Donc il existe clairement une constante C_5 qui dépend seulement de m_0 et p_0 telle

$$\text{que } \int_0^\infty \frac{(\partial g_{1k} / \partial p)^2}{f_k} \Big|_{\theta_0} dX_k \leq C_5$$

En mettant $u = X_k - p_0 X_{k-1}$ on voit par exactement les mêmes manipulations qu'il existe une constante C_6 indépendante de X_{k-1} telle que

$$\int_{p_0 X_{k-1}}^{\infty} \frac{(\partial g_{2k}/\partial p)^2}{f_k} \Big|_{\theta_0} dX_k \leq C_6$$

Donc par (4.4) et (4.5)

$$E_{\theta_0}(a_{1k}^2) \leq 2(C_5 + C_6) < \infty \quad \blacksquare$$

Démonstration du théorème 1 : Nous démontrons d'abord que

$$(4.6) \quad \lim_{n \rightarrow \infty} (n^{-1}(\log L_n(m_0, p_1) - \log L_n(m, p_2))) > 0 \quad \text{p.s.}$$

pour tout m et pour tout p_1 et p_2 tels que $p_0 - r/n \leq p_1 \leq p_0 + r/n$ et $p_2 = p_0 + \varepsilon$, où $\varepsilon \in [-p_0, 0) \cup (0, 1-p_0)$.

Supposons d'abord que $p_1 = p_0 + \delta/n$ et $p_2 = p_0 + \varepsilon$, où $0 \leq \delta \leq r$ et $0 < \varepsilon < 1 - p_0$. Alors par (4.1)

$$\begin{aligned} & \lim_{n \rightarrow \infty} (n^{-1}(\log L_n(X|m_0, p_1) - \log L_n(X|m, p_2))) \\ &= \lim_{n \rightarrow \infty} ((\delta/n - \varepsilon)d_1 - (m-m_0)d_2 + \frac{1}{2}(\delta^2/n^2 - \varepsilon^2)d_3 - \varepsilon(m-m_0)d_4 \\ & - \frac{1}{2}(m-m_0)^2d_5 - (\delta/n - \varepsilon)d_6 - (\delta^2/n^2 - \varepsilon^2)d_7 - \varepsilon(m-m_0)d_8) \\ &= \lim_{n \rightarrow \infty} (d_0 - (d_0 + \varepsilon d_1 + (m-m_0)d_2 + \frac{1}{2}\varepsilon^2 d_3 + \varepsilon(m-m_0)d_4 \\ & + \frac{1}{2}(m-m_0)^2d_5 - \varepsilon d_6 - \varepsilon^2 d_7 - \varepsilon(m-m_0)d_8)) \\ &= \lim_{n \rightarrow \infty} (n^{-1}(\log L_n(X|m_0, p_0) - \log L_n(X|m, p_2))) \\ &> 0 \quad \text{p.s. par (3.2)} \end{aligned}$$

(4.6) est donc satisfait dans ce cas.

Un raisonnement analogue basé sur (4.2) montre que (4.6) est vérifié quand p_1 et p_2 sont tous les deux plus petits que p_0 .

De même, si $p_1 = p_0 + \delta/n$ et $p_2 = p_0 - \epsilon$, où $0 \leq \delta \leq r$, $0 < \epsilon \leq p_0$, on a

$$\begin{aligned}
 (4.7) \quad & \lim_{n \rightarrow \infty} (n^{-1} (\log L_n(X|m_0, p_1) - \log L_n(X|m, p_2))) \\
 &= \lim_{n \rightarrow \infty} ((\delta/n + \epsilon)d_1 - (m-m_0)d_2 + \frac{1}{2} (\delta^2 d_3/n^2 - \epsilon^2 d_3') + \epsilon d_4' \\
 &\quad - \frac{1}{2}(m-m_0)^2 d_5' - (\delta d_6/n - \epsilon d_6') - (\delta^2 d_7/n^2 - \epsilon^2 d_7') + \epsilon(m-m_0)d_8') \\
 &= \lim_{n \rightarrow \infty} ((\epsilon d_1 - (m-m_0)d_2 - \epsilon^2 d_3' + \epsilon d_4' - \frac{1}{2}(m-m_0)^2 d_5' + \epsilon d_6' \\
 &\quad - \epsilon d_7' + \epsilon(m-m_0)d_8') + (\delta/n)(d_1 + \frac{1}{2}\delta d_3/n - d_6 - \delta d_7/n))
 \end{aligned}$$

Or, par la stationarité de (2.3), a_{ik} , $i=1,3,6,7$ et a'_{ik} , $i=3,6,7$ sont des fonctions de (X_k, X_{k-1}) qui sont les mêmes pour tout k et qui appartiennent à $L^2(X_k, X_{k-1})$ par le Lemme 1. Donc par (2.4) il existe des constantes K_i , $i=1,3,6,7$ et K'_i , $i=3,6,7$ indépendantes de j telles que

$$\text{Corr}(a_{ik}, a_{i,k+j}) \leq K_i p^j, \quad i=1,3,6,7$$

$$\text{Corr}(a'_{ik}, a'_{i,k+j}) \leq K'_i p^j, \quad i=3,6,7, \text{ pour } j \geq -\log 15/\log p$$

$\{a_{ik}\}_{k \in \mathbb{N}}$ $i=1,3,6,7$ et $\{a'_{ik}\}_{k \in \mathbb{N}}$ $i=3,6,7$ satisfont donc à la loi forte des grands

nombre pour les processus stationnaires (Doob(1953), p. 489). Donc

$$(4.8) \quad \lim_{n \rightarrow \infty} ((\delta/n) (d_1 + \frac{1}{2} \delta d_3/n - d_6 - \delta d_7/n)) = 0 \text{ p.s.}$$

Donc par (4.7) on a p.s.

$$\begin{aligned}
 & \lim_{n \rightarrow \infty} n^{-1} (\log L_n(X|m_0, p_1) - \log L_n(X|m, p_2)) \\
 &= \lim_{n \rightarrow \infty} (d_0 - (d_0 - \epsilon d_1 + (m-m_0)d_2 + \frac{1}{2} \epsilon^2 d_3' - \epsilon d_4' + \frac{1}{2} (m-m_0)^2 d_5' \\
 &\quad - \epsilon d_6' + \epsilon^2 d_7' - \epsilon(m-m_0)d_8')) \\
 &= \lim_{n \rightarrow \infty} n^{-1} (\log L_n(X|m_0, p_0) - \log L_n(X|m, p_2))
 \end{aligned}$$

> 0 par (3.2), et (4.6) est vérifié dans ce cas.

Un raisonnement analogue tient quand $p_1 < p_0$ et $p_2 > p_0$. Donc (4.6) est vérifié dans tous les cas possibles.

D'autre part

$$D_n(\theta_0) = \int_{p_0 - r/n}^{p_0 + r/n} L_n(X|m_0, p) dp \geq (2r/n) \min_{p_0 - r/n \leq p_1 \leq p_0 + r/n} L_n(X|m_0, p_1)$$

$$D_n(\theta) = \int_{p_2 - r/n}^{p_2 + r/n} L_n(X|m, p) dp \leq (2r/n) \max_{p_2 - r/n \leq p \leq p_2 + r/n} L_n(X|m, p)$$

Donc il existe $p_1 \in [p_0 - r/n, p_0 + r/n]$ et $p \in [p_2 - r/n, p_2 + r/n]$ tels que

$$D_n(\theta_0)/D_n(\theta) \geq L_n(X|m_0, p_1) / L_n(X|m, p)$$

Donc si l'on écrit $\alpha(\theta, \theta_0) = \lim_{n \rightarrow \infty} \log (L_n(X|\theta_0)/L_n(X|\theta)) > 0$ p.s.

par (4.6), on a

$$\begin{aligned} \lim_{n \rightarrow \infty} D_n(\theta_0)/D_n(\theta) &\geq \lim_{n \rightarrow \infty} L_n(X|m_0, p_1)/L_n(X|m, p_2) \\ &= \lim_{n \rightarrow \infty} \exp(n\alpha(\theta, \theta_0)) = \infty \text{ p.s.} \end{aligned}$$

où $\theta = (m, p_2)$, par la démonstration de (4.6). Donc

$$\lim_{n \rightarrow \infty} n^{-1}(\log D_n(\theta_0) - \log D_n(\theta)) > 0 \text{ p.s.}$$

et le théorème est vérifié par le raisonnement qui mène à (3.3). ■

Pour démontrer le théorème 2 on a besoin de deux autres lemmes techniques.

Lemme 2 : Soient $\delta = (\delta_1, \delta_2)$ et $\varepsilon = (\varepsilon_1, \varepsilon_2)$, où $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$ sont des réels avec

$$0 \leq \delta_1 < m_0, 0 \leq \delta_2 < \infty, 0 \leq p_0 - \varepsilon_1 \leq p_0 + \varepsilon_2 < 1 \text{ et notons :}$$

$$\mathcal{J}(\theta_0, \delta, \varepsilon) = \{(m, p) : m_0 - \delta_1 \leq m \leq m_0 + \delta_2 \text{ et } p_0 - \varepsilon_1 \leq p \leq p_0 + \varepsilon_2\}$$

Alors il existe des fonctions u_i, u_i' , $i = 3, 4, 5, 7, 8, 11, 12, 13$, de X_k, X_{k-1}, m et p et

les constantes $M_i(\theta_0, \delta, \varepsilon), M_i'(\theta_0, \delta, \varepsilon)$, $i = 3, 4, 5, 7, 8, 11, 12, 13$ telles que

$$(4.9) \quad \begin{cases} \max_{\theta \in B(\theta_0, \delta, \epsilon)} |E(u_i)| < M_i(\theta_0, \delta, \epsilon) \\ \max_{\theta \in B(\theta_0, \delta, \epsilon)} E(u_i^2) < \infty, \quad i = 3, 4, 5, 7, 8, 11, 12, 13 \end{cases}$$

et

$$(4.10) \quad \begin{cases} |a_{ik}| < u_i, \quad i = 3, 4, 5, 11, 12, 13 \\ |a_{ik}| < |p - p_0| u_i, \quad i = 7, 8 \end{cases}$$

et que (4.9) et (4.10) sont vrais aussi quand les u_i, M_i et a_{ik} sont remplacés par les u'_i, M'_i et a'_{ik} correspondants pour $i = 3, 4, 5, 7, 8$.

La démonstration de ce lemme est composée de calculs standards qui sont omis. Néanmoins, nous avançons la justification heuristique suivante : les a_{ik} et les a'_{ik} en question sont bornés partout sauf pour p dans un voisinage de 1 et m dans un voisinage de 0. C'est pourquoi nous devons nous contenter de ces bornes locales, qui sont suffisantes pour démontrer le théorème 2 grâce à la consistance de $\hat{\theta}$.

Lemme 3 : Si $k_1 = E_{\theta_0}(a_{1k})$ et $k_6 = \lim_{p \rightarrow p_0} (p - p_0)^{-1} E_{\theta_0}(a_{6k})$

alors $k_6 > k_1$ p.s.

Démonstration : $k_1 = p_0$ par la démonstration du lemme 1.

$$\begin{aligned} (p - p_0)k_6 &= E_{\theta_0} [(\delta_k(p_0) - \delta_k(p))\alpha_k(\theta_0)] \\ &= \int_0^\infty \int_{p_0 X_{k-1}}^{p X_{k-1}} \log(1 + g_{2k, \theta_0} / g_{1k, \theta_0}) (g_{1k, \theta_0} + g_{2k, \theta_0}) dX_k \\ &\quad m_0^{-1} \exp(-X_{k-1}/m_0) dX_{k-1} \end{aligned}$$

Or pour $x > 0$, $\log(1+x) = x/(1+x) + \frac{1}{2}(x/(1+x))^2 + \frac{1}{3}(x/(1+x))^3 + \dots > x/(1+x)$

$$\begin{aligned} \text{Donc } (p-p_0)k_6 &> \int_0^\infty \int_{p_0 X_{k-1}}^{p X_{k-1}} g_{2k} dX_k m_0^{-1} \exp(-X_{k-1}/m_0) dX_{k-1} \\ &= (p-p_0)p_0(1-p_0+p_0^2)^{-1} \int_0^\infty (X_k/m_0)(1-p_0+p_0(1-p_0)^{-1})m_0^{-1} \exp(-X_{k-1}/m_0) dX_{k-1} \\ &\quad + O(p-p_0)^2 \end{aligned}$$

$$\text{Donc } k_6 > p_0 (1-2p_0 + 2p_0^2)/(1-2p_0+2p_0^2-p_0^3) > p_0 = k_1 \quad \blacksquare$$

démonstration du théorème 2 :

(i) On pose $\phi = n(\hat{p}-p_0)$. Si on écrit $L_n(\phi)$ pour $L_n(X|\hat{\theta})$ on a de (4.1) et (4.2) que

$$(4.11) \quad \log L_n(\phi) = \begin{cases} nd_0 - \phi(d_0-d_1) + n(\hat{m}-m_0)d_2 + \frac{1}{2} \phi(\hat{p}-p_0)(d_3-2d_7) \\ \quad + \phi(\hat{m}-m_0)(d_4-d_8) + \frac{1}{2} n(\hat{m}-m_0)^2 d_5, & \phi \geq 0 \\ nd_0 + \phi(d'_0+d_1) + n(\hat{m}-m_0)d_2 + \frac{1}{2} \phi(\hat{p}-p_0)(d'_3+2d'_7) \\ \quad + \phi(\hat{m}-m_0)(d'_4+d'_8) + \frac{1}{2} n(\hat{m}-m_0)^2 d'_5, & \phi < 0 \end{cases}$$

Il s'ensuit des lemmes 1 et 2 et du théorème 1 par le raisonnement qui mène à 4.8) que d_1, d_6 et d'_6 satisfont à la loi forte des grands nombres et que $|d_i|$ et $|d'_i|$ ont bornés par des fonctions qui y satisfont pour $i = 2, 3, 4, 5, 7, 8$. Supposons que

$$\langle \epsilon < \min \{ (k_0-k_1)/(6+M_3/2+M_4+M_7+M_8), (k'_6+k_1)/(6+M'_3/2+M'_4+M'_7+M'_8), p_0, 1-p_0, m_0 \}$$

$$\text{ù } k'_6 = \lim_{p \rightarrow p_0} (p_0-p)^{-1} E_{\theta_0}(a'_{6k})$$

et où nous écrivons M_i pour $M_i(\theta_0, \delta, \epsilon)$ et M'_i pour $M'_i(\theta_0, \delta, \epsilon)$.

Alors il existe une suite $\{\delta_n\}_{n \in \mathbb{N}}$ avec $\delta_n \rightarrow 0$ telle que pour n suffisamment grand chacune des affirmations suivantes est vraie avec probabilité $> 1-\delta_n$ sous θ_0 :

$$|\hat{p}-p_0| < \epsilon \quad \text{et} \quad |\hat{m}-m_0| < \epsilon$$

$$|d_1 - k_1| < \epsilon$$

$$|d_6 - k_6| < \epsilon$$

$$|d'_6 - k'_6| < \epsilon$$

$$|d_i| < M_i + \epsilon, \quad i = 2, 3, 4, 5, 7, 8$$

Donc avec probabilité $\geq 1 - 16 \delta_n$ on a par (4.11)

$$(4.12) \quad \begin{cases} nd_0 - \phi(k_6 - k_1 + C_7 \varepsilon) - nC_8 \varepsilon \leq \log L_n(\phi) \leq nd_0 - \phi(k_6 - k_1 - C_7 \varepsilon) \\ \quad + nC_8 \varepsilon, \quad \phi \geq 0 \\ nd_0 + \phi(k'_6 + k_1 - C_9 \varepsilon) - nC_{10} \varepsilon \leq \log L_n(\phi) \leq nd_0 + \phi(k'_6 + k_1 + C_9 \varepsilon) \\ \quad + nC_{10} \varepsilon, \quad \phi < 0 \end{cases}$$

$$\text{où } C_7 = 2 + M_3/2 + M_4 + M_7 + M_8 + 9\varepsilon/2$$

$$C_8 = M_2 + \varepsilon + (\varepsilon/2)(M_5 + \varepsilon)$$

$$C_9 = 2 + M'_3/2 + M'_4 + M'_7 + M'_8 + 9\varepsilon/2$$

$$C_{10} = M_2 + \varepsilon + (\varepsilon/2)(M'_5 + \varepsilon)$$

Par définition $D_n(p_0 + z/n, \hat{m}) = \int_{z-r}^{z+r} L_n(\phi) d\phi$. On considère séparément cette intégrale pour $z \in [r, \infty)$, $z \in (-\infty, r]$, et éventuellement pour $z \in [-r, r]$. Tout ce qui suit est vrai avec probabilité $\geq 1 - 16 \delta_n$ sous θ_0 :

Si $z \in [r, \infty)$, par (4.12)

$$\begin{aligned} & (k_6 - k_1 + C_7 \varepsilon)^{-1} \exp(n(d_0 - C_8 \varepsilon)) [\exp(-(z-r)(k_6 - k_1 + C_1 \varepsilon)) \\ & \quad - \exp(-(z+r)(k_6 - k_1 + C_7 \varepsilon))] \leq D_n(p_0 + z/n, \hat{m}) \\ & \leq (k_6 - k_1 - C_7 \varepsilon)^{-1} \exp(n(d_0 + C_8 \varepsilon)) [\exp(-(z-r)(k_6 - k_1 - C_1 \varepsilon)) \\ & \quad - \exp(-(z+r)(k_6 - k_1 + C_7 \varepsilon))] \end{aligned}$$

Par le lemme 3 on sait que $k_6 - k_1 > 0$ pour n suffisamment grand, et puisque $0 < \varepsilon < (k_6 - k_1)/C_7$ on a

$$(4.13) \quad \max_{z \in [r, \infty)} D_n(p_0 + z/n, \hat{m}) \text{ est atteint quand } z = r.$$

De même quand $z \in (-\infty, -r]$:

$$\begin{aligned} & (k'_6 + k_1 + C_9 \varepsilon)^{-1} \exp(n(d_0 - C_{10} \varepsilon)) [\exp((z+r)(k'_6 + k_1 + C_9 \varepsilon)) \\ & \quad - \exp((z-r)(k'_6 + k_1 + C_9 \varepsilon))] \leq D_n(p_0 + z/n, \hat{m}) \\ & \leq (k'_6 + k_1 - C_9 \varepsilon)^{-1} \exp(n(d_0 + C_{10} \varepsilon)) [\exp((z+r)(k'_6 + k_1 - C_9 \varepsilon)) \\ & \quad - \exp((z-r)(k'_6 + k_1 - C_9 \varepsilon))] \end{aligned}$$

$k'_6 > 0$ par sa définition et donc $k'_6 + k_1 > 0$.

Alors puisque $0 < \epsilon < (k'_6 + k_1) / C_9$ on a

$$(4.14) \quad \max_{z \in (-\infty, -r]} D_n(p_0 + z/n, \hat{m}) \text{ est atteint quand } z = -r.$$

Il s'ensuit de (4.13) et (4.14) que

$$P_{\theta_0} [\max_{z \in \mathbb{R}} D_n(p_0 + z/n, \hat{m}) \text{ est atteint quand } z \in [-r, r]] \geq 1 - 16\delta_n$$

c'est-à-dire que

$$P_{\theta_0} [-r/n \leq \hat{p} - p_0 \leq r/n] \geq 1 - 16\delta_n \rightarrow 1 \text{ quand } n \rightarrow \infty$$

et (3.7) est vérifié.

(ii) Par (3.5) et (4.3)

$$(4.15) \quad n^{1/2}(\hat{m} - m_0) = \frac{-n^{1/2}(d_2 + (\hat{p} - p_0)d_{10} + \frac{1}{2}(\hat{p} - p_0)^2 d_{13} + (\hat{p} - p_0)d_{14})}{d_9 + \frac{1}{2}(\hat{m} - m_0)d_{11} + (\hat{p} - p_0)d_{12}}$$

f_k est trois fois dérivable en m et donc par des arguments standards (Kendall et Stuart (1973), p.8) $E_{\theta_0}(a_{2k} | X_{k-1}) = 0$, c'est-à-dire que d_{2k} est un martingale et $E_{\theta_0}(a_{2k}^2) = -E_{\theta_0}(a_{9k})$ que nous notons $J(\theta_0)$.

Par le lemme 1, $J(\theta_0) < \infty$ et puisque $\{X_k\}_{k \in \mathbb{N}}$ est stationnaire et ergodique,

$\{a_{2k}\}_{k \in \mathbb{N}}$ l'est aussi et les hypothèses du théorème de Lindeberg-Lévy pour les martingales (Billingsley (1961)) sont satisfaits. Donc $n^{1/2} d_2 \rightarrow N(0, J(\theta_0))$.

Par le théorème 1, $(\hat{m} - m_0) \rightarrow 0$ en probabilité et par (3.7), $n^{1/2}(\hat{p} - p_0) \rightarrow 0$ en probabilité. d_{10} est borné en probabilité puisque d_4 et d'_4 le sont par le lemme 2, et d_{14} l'est puisque d_8 et d'_8 le sont par le lemme 2. d_{11} , d_{12} et d_{13} sont bornés en probabilité par le lemme 2. d_2 satisfait à la loi forte des grands nombres par l'argument qui mène à (4.8). Donc par (4.15) $n^{1/2}(\hat{m} - m_0) \rightarrow N(0, 1/J(\theta_0))$.

(iii) Il est clair de (ii) que la loi limite de \hat{m} ne dépend pas de \hat{p} . D'autre part, par (4.11) et (4.12) pour $z \in [-r, r]$

$$\begin{aligned} & (d_6 - d_1)^{-1} \exp(n(d_0 - C_8 \epsilon)) (1 - \exp(-(z+r)(d_6 - d_1))) \\ & + (d_6' + d_1)^{-1} \exp(n(d_0 - C_{10} \epsilon)) (1 - \exp((z-r)(d_6' + d_1))) \leq D_n(p_0 + z/n, \hat{m}) \\ & \leq (d_6 - d_1)^{-1} \exp(n(d_0 + C_8 \epsilon)) (1 - \exp(-(z+r)(d_6 - d_1))) \\ & + (d_6' + d_1)^{-1} \exp(n(d_0 - C_{10} \epsilon)) (1 - \exp((z-r)(d_6' + d_1))) \end{aligned}$$

$D_n(p_0 + z/n, \hat{m})$ atteint donc son maximum au point

$$(4.16) \quad z = r \frac{(d_6' + d_1) - (d_6 - d_1)}{(d_6' + d_1) + (d_6 - d_1)} \quad \text{que nous notons } z = r\omega$$

Donc $n(\hat{p} - p_0)$ suit asymptotiquement en probabilité la loi de $r((1/\omega)V - 1)$ et la loi limite de \hat{p} est indépendante de \hat{m} . ■

Démonstration du corollaire : Il s'ensuit de (4.15) que

$$(4.17) \quad (\hat{m} - \check{m}) = (\hat{p} - \check{p})(d_{10} + d_{14})/d_9 + o(\hat{p} - \check{p})$$

et puisque d_{10}, d_{14} et d_9 satisfont à la loi forte des grands nombres, (3.8) est vérifié.

D'autre part, supposons que $\hat{p} - p = O(n^{-a})$, $0 < a < 1$. Nous savons du théorème

2(i) que $\hat{p} - p_0 = (\delta/n) + o(n^{-1})$, $-r \leq \delta \leq r$.

Ecrivons $\check{p} - \hat{p} = \epsilon n^{-a} + o(n^{-a})$ où $\epsilon \in (-\infty, 0) \cup (0, \infty)$

Alors de (4.1) et (4.17)

$$(4.18) \quad n^{-1} (\log L_n(X|\hat{\theta}) - \log L_n(X|\check{\theta})) = \begin{cases} \epsilon n^{-a} (d_6 - d_1 + d_2(d_{10} + d_{14})/d_9) + o(n^{-a}) & \epsilon \geq 0 \\ \epsilon n^{-a} (-d_6' - d_1 + d_2(d_{10} + d_{14})/d_9) + o(n^{-a}) & \epsilon < 0 \end{cases}$$

Or, par le lemme 3, $\lim_{n \rightarrow \infty} P_{\theta_0} [d_6 - d_1 > 0] = 1$, et par la démonstration du théorème

2(ii), $\lim_{n \rightarrow \infty} P_{\theta_0} [|d_2| < \gamma] = 1 \quad \forall \gamma > 0$.

On sait également que pour n suffisamment grand, $E_{\theta_0}(d'_6 + d_1) > 0$, et $d'_6 + d_1$ satisfait à la loi forte des grands nombres par le raisonnement qui mène à (4.8).

$$\lim_{n \rightarrow \infty} P_{\theta_0} [d'_6 + d_1 > 0] = 1.$$

onc par (4.18)

$$\lim_{n \rightarrow \infty} P_{\theta_0} [n^{-(1-a)} (\log L_n(X|\hat{\theta}) - \log L_n(X|\tilde{\theta})) > 0] = 1$$

ce qui contredit la définition de $\tilde{\theta}$. Donc $\hat{p} - p \leq O(\frac{1}{n})$

en plus, par (4.16),

$$\hat{p} = p_0 + n^{-1} (r(1/\lambda\omega)\psi - 1) + o(n^{-1})$$

ce qui ne maximise pas la fonction de vraisemblance.

onc $(\hat{p} - p) \geq O(n^{-1})$ et (3.9) est vérifié. ■

Dans le cas où les observations sont indépendantes avec densité discontinue, par Ibragimov et Hasminskii (1972) équation (5.3) on a pour r suffisamment petit :

$$EMP = EMV \pm r/n + O(n^{-1}) .$$

Il semble par (4.16) et le Corollaire que ce résultat ne se généralise pas au cas où les observations sont dépendantes, du moins lorsque les discontinuités dans la densité d'une observation dépendent des observations précédentes.

Pour calculer la loi exacte de \hat{p} nous notons que a_{1k} , a_{6k} et a'_{6k} ne dépendent que de X_k et de X_{k-1} et donc $\{a_{1k}\}_{k \in \mathbb{N}}$, $\{a_{6k}\}_{k \in \mathbb{N}}$ et $\{a'_{6k}\}_{k \in \mathbb{N}}$ sont des suites ϕ -mélangeantes par (2.5). Par un raisonnement analogue à celui utilisé pour démontrer (2.5), les $\phi(k)$ pour ces suites satisfont à $\sum_{k=-\infty}^{\infty} \phi(k)^{1/2} < \infty$ et donc d_1, d_6 et d'_6 satisfont au théorème 20.1 de Billingsley (1968) et suivent asymptotiquement une loi normale. Il s'ensuit de (4.16) que $\omega = \omega_1/\omega_2$ où (ω_1, ω_2) est un couple gaussien non-centré avec $0 \leq E(\omega_1) < E(\omega_2)$.

5. GENERALISATIONS

On peut considérer \hat{p} comme l'estimateur de Bayes pour p avec loi a priori uniforme et fonction de perte $\psi(\hat{p}, p_0) = 1_{[-r/n, r/n]}(\hat{p} - p_0)$. Il est possible de construire également des estimateurs qui sont Bayes pour d'autres fonctions de perte à l'instar de Weiss et Wolfowitz (1974), ch. 4. Par exemple, par rapport à la fonction de perte quadratique $\psi(\hat{p}, p_0) = n^2(\hat{p} - p_0)^2$ on peut définir l'EMP de p comme la valeur de d qui minimise la quantité

$$\int_{(d-n)^{1/2} \vee 0}^{(d+n)^{1/2} \wedge 1} (d-p)^2 L_n(X|\theta) dp$$

La vérification des conditions (4.4), (4.5) et (4.6) de Weiss et Wolfowitz (1974) suit directement de notre section 4, et donc par leurs théorèmes 4.1 et 4.2, cet estimateur est asymptotiquement efficace pour la fonction de perte qui le définit. \hat{m} est défini de la même façon que dans (3.4) et les parties (ii) et (iii) de notre théorème 2 sont vérifiées.

Dans cet article nous nous sommes bornés pour simplifier l'exposition à considérer le modèle (2.3) qui a deux paramètres. En fait des résultats analogues sont vrais pour le modèle plus général (2.3) qui a trois paramètres. Si $\hat{\theta} = (\hat{m}, \hat{p}, \hat{\alpha})$ est défini par

$$\left\{ \begin{array}{l} \hat{\alpha} \text{ est la valeur de } \alpha \text{ qui maximise} \\ D_n(\theta) = \int_{\alpha-r/n}^{\alpha+r/n} L_n(X|\theta) d\alpha \\ \text{où } L_n(X|\theta) \text{ est la fonction de vraisemblance de (2.2) et } (\hat{m}, \hat{p}) \text{ est la valeur} \\ \text{de } (m, p) \text{ qui maximise } L_n(X|\theta) \end{array} \right.$$

alors théorème 1 reste vrai et on a

Théorème 2' :

(i) $\hat{\alpha}$ est un estimateur asymptotiquement efficace pour α au sens que

$$\lim_{n \rightarrow \infty} p_{\theta_0} [-r/n \leq \hat{\alpha} - \alpha_0 \leq r/n] = 1$$

(ii) $(n^{1/2}(\hat{m}-m_0), n^{1/2}(\hat{p}-p_0))$ suit asymptotiquement une loi normale bivariée centrée

avec matrice de variance-covariance B^{-1}

$$B = - E_{\theta_0} \begin{pmatrix} \frac{\partial^2 \log f_k}{\partial m^2} & \frac{\partial^2 \log f_k}{\partial m \partial p} \\ \frac{\partial^2 \log f_k}{\partial m \partial p} & \frac{\partial^2 \log f_k}{\partial p^2} \end{pmatrix}$$

c'est-à-dire que (\hat{m}, \hat{p}) est un estimateur asymptotiquement normal et asymptotiquement efficace de (m, p)

(iii) $\hat{\alpha}$ et (\hat{m}, \hat{p}) sont asymptotiquement indépendants.

La démonstration est parallèle à celle du théorème 2. En plus, si on vérifie (iii) (qui ne dépend pas de (ii)), alors (ii) est une conséquence immédiate de Crowder (1976), section 3. Un résultat analogue au Corollaire est également vérifié à savoir essentiellement que $(\hat{\alpha} - \tilde{\alpha}), (\hat{p} - \tilde{p})$ et $(\hat{m} - \tilde{m})$ sont tous $O(n^{-1})$.

Les mêmes résultats asymptotiques tiennent également quand la réalisation X est la suite des intervalles entre les évènements dans un processus ponctuel où le premier évènement n'est pas observé à l'instant zéro, (c'est-à-dire que le processus est en mode asynchrone), et où X_1 est le temps entre l'instant zéro et le premier évènement. En ce cas les $f_k, k \geq 3$ restent inchangés et f_1, f_2 doivent être changés suivant la formule (4.21) de Cox et Lewis (1966). Il est clair que l'effet de ce changement est négligeable pour les grands échantillons.

-:-:-:-:-:-:-:-

Le travail dont rend compte cet article fait partie de la préparation d'une thèse de doctorat sous la direction de Monsieur Paul DEHEUVELS que je tiens à remercier de son aide constante.

BIBLIOGRAPHIE

- 1 Billingsley P. (1961). "The Lindeberg-Lévy theorem for martingales", Proc. Amer. Math. Soc. 12, 788-792.
- 2 Billingsley P. (1968). Convergence of probability measures, Wiley, New York.
- 3 Cox, D.R. (1955). "Some statistical methods connected with series of events", J. Roy. Statist. Soc. B 17, 129-164.
- 4 Cox, D.R., Lewis, P.A.W. (1966). The statistical analysis of series of events, Methuen, Londres.
- 5 Crowder, M.J. (1976). "Maximum likelihood estimation for dependent observations", J. Roy. Statist. Soc. B 38, 45-53.
- 6 Doob, J.L. (1953). Stochastic processes, Wiley, New York.
- 7 Downton, F. (1970) "Bivariate exponential distributions of reliability theory", J. Roy. Statist. Soc. B 32, 408-417.
- 8 Gaver, D.P. (1972). "Point process problems in reliability", in Stochastic point processes (éd. Lewis), Wiley, New York, 775-800.
- 9 Gaver, D.P., Lewis, P.A.W. (1980) "First order autoregressive sequences", J. Appl. Prob. A paraître.
- 10 Hawkes, A.G. (1977) "Mutually exciting point processes", in Stochastic point processes (éd. Lewis) Wiley, New York, 261-271.
- 11 Heyde, C.C. (1978). "On an optimal asymptotic property of the maximum likelihood estimator of a parameter from a stochastic process", Stoch. Proc. Appl. 8, 1-9.
- 12 Ibragimov, I.A., Hasminskii, R.Z. (1972). "Asymptotic behaviour of statistical estimates for samples with a discontinuous density", Math. USSR Sbornik 16, 573-606 (traduit du russe).
- 13 Jacobs, P.A. (1978). "A closed cyclic queueing network with dependent exponential service times", J. Appl. Prob. 15, 573-589.
- 14 Jacobs, P.A., Lewis, P.A.W. (1977). "A mixed autoregressive-moving average exponential sequence and point process EARMA(1,1)", Adv. Appl. Prob 9, 87-104.
- 15 Kendall, M.G., Stuart, A. (1973). The advanced theory of statistics, vol. II (3e édition), Griffin, Londres.
- 16 Lawrance, A.J. (1972). "Some models for stationary series of univariate events in Stochastic point processes (éd. Lewis), Wiley, New York, 199
- 17 Lawrance, A.J. (1980). "Some autoregressive models for point processes", Proc. Bolyai Janos Math. Soc. Colloq. on point processes and queueing theory, North Holland, A paraître.

- 8 Lawrance, A.J., Kottegoda, N.T. (1977). "Stochastic modelling of river flow time series", J. Roy. Statist. Soc. A 140, 1-47.
- 9 Lawrance, A.J., Lewis, P.A.W. (1977). "An exponential moving-average sequence and point process", J. Appl. Prob. 14, 98-113.
- 0 Lawrance, A.J., Lewis, P.A.W. (1980). "An exponential autoregressive-moving average process EARMA(p,q) : definition and correlational properties". J. Roy. Statist. Soc. B. A paraître.
- 1 Lewis, P.A.W., Shedler, G.S. (1978). "Analysis and modelling of point processes in computer systems", Proc. 41 st session, Int. Statist. Inst.
- 2 Linhart, H. (1970). "The gamma point process", S. Afr. Statist. J. 4, 1-17.
- 3 Linhart, H, Zucchini, W. (1972). "Parametric models for point processes", S.Afr. Statist. J. 6, 19-26.
- 4 Marshall, A.W., Olkin, I. (1967). "A generalised bivariate exponential distribution", J. Appl. Prob. 4, 291-302.
- 5 Mallows, C.L. (1967). "Linear processes are nearly Gaussian", J. Appl. Prob. 4, 313-329.
- 6 Nelson, H.L., Granger, C.W.J. (1978). "Experience with using the Box-Cox transformation when forecasting economic time series", Economics Dept., Univ. Calif., San Diego.
- 7 Raftery, A.E. (1980). "Un processus autoregressif a loi marginale exponentielle : propriétés asymptotiques et estimation de maximum de vraisemblance", Ann. Sci. Univ. Clermont. A paraître.
- 8 Rosenblatt, M. (1971). Markov processes. Structure and asymptotic behaviour, Springer-Verlag, Berlin.
- 9 Weiss, L., Wolfowitz, J. (1974). Maximum probability estimators and related topics, Lecture notes in mathematics 424, Springer-Verlag, Berlin.
- 0 Wold, H. (1948). "Sur les processus stationnaires ponctuels", Coll. Int. C.N.R.S. 13, 75-86.

Reçu en Janvier 1980