

REVUE DE STATISTIQUE APPLIQUÉE

SALIMA TAIBI-HASSANI
ÉLIE YOUNDJÉ

Validation croisée pour l'estimateur lissé de la fonction de hasard : cas des données censurées

Revue de statistique appliquée, tome 51, n° 1 (2003), p. 73-86

http://www.numdam.org/item?id=RSA_2003__51_1_73_0

© Société française de statistique, 2003, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

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

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

VALIDATION CROISÉE POUR L'ESTIMATEUR LISSÉ DE LA FONCTION DE HASARD : CAS DES DONNÉES CENSURÉES

Salima TAIBI-HASSANI⁽¹⁾, Élie YOUNDJÉ⁽²⁾

⁽¹⁾ *E.S.I.T.P.A, 13 rue du Nord 76000 Rouen, France*

⁽²⁾ *Laboratoire Raphaël Salem, UMR-CNRS 6085
Université de Rouen, 76821 Mont Saint Aignan cedex, France
e.mail : staibi@esitpa.org, youndje@univ-rouen.fr*

RÉSUMÉ

Nous considérons dans le cadre des données censurées, un estimateur de la fonction de hasard construit à l'aide d'un estimateur à noyau de la fonction de "survie" de la variable censurée. Dans un premier temps nous donnons la décomposition asymptotique de l'erreur quadratique intégrée, dans un second temps nous proposons une méthode pour sélectionner la largeur de fenêtre; l'optimalité asymptotique de la fenêtre sélectionnée est établie. Une étude par simulation complète ce travail.

Mots-clés : *Estimation de la fonction de Hasard, Noyau, Sélection du Paramètre de Lissage, Censure, Validation-croisée.*

ABSTRACT

In the setting of censored data, we consider an estimator of hazard rate built with a kernel estimator of the "survival" function of the censored variable. In a first attempt we give the asymptotic expression of the Integrated Square Error (ISE), in a second attempt we propose a method based on cross-validation ideas to select the bandwidth; the selected bandwidth is shown to be asymptotically optimal. A small scale simulation completes this paper.

Keywords : *Hazard Estimation, Kernel function, Smoothing Parameter Selection, Censored Data, Cross-validation.*

1. Introduction

L'estimation de la fonction de hasard est un problème intéressant qui apparaît dans l'analyse statistique des durées de vie, notamment l'étude de la survie, de la fiabilité ... Dans le cadre des données complètes (non censurées) la plupart des estimateurs non-paramétriques de la fonction de hasard sont définis comme rapport d'un estimateur de la densité et d'un estimateur de la fonction de survie. La fonction de survie étant généralement estimée à l'aide de la fonction de répartition empirique. Cependant, comme mentionné dans Youndjé *et al.* (1996), estimer la fonction de survie à l'aide d'un estimateur à noyau de la fonction de répartition présente au moins

deux avantages : d'une part, on obtient un estimateur continu, et d'autre part on améliore l'estimation de la fonction de survie en termes d'erreur quadratique. Dans le cadre des données censurées les estimateurs de la fonction de hasard sont aussi en général des rapports d'estimateurs dont le dénominateur est un estimateur de type survie. Donc comme dans le cas des données complètes, il est avantageux d'utiliser un estimateur à noyau au dénominateur.

Comme dans la plupart des problèmes d'estimation fonctionnelle, la performance d'un estimateur à noyau de la fonction de hasard dépend fortement du choix de la largeur de fenêtre. Dans le cadre des données complètes, une procédure de sélection de la largeur de fenêtre a été proposée par Sarda et Vieu (1991), lorsque la fonction de survie est estimée par la fonction de répartition empirique. Cette procédure a été étendue au cas des données censurées par Patil (1993a) et étudiée davantage dans Patil (1993b). Une règle de sélection de la fenêtre a été introduite par Youndjé *et al.* (1996) dans le cadre des données complètes, lorsque la fonction de survie est estimée à l'aide d'un noyau.

Le but de cet article est d'étendre cette dernière règle au cas des données censurées. Plus précisément, nous proposons une méthode basée sur l'idée de validation croisée pour calculer la fenêtre minimisant asymptotiquement l'erreur quadratique intégrée. La Section 2 présente l'estimateur considéré dans l'article, la Section 3 présente l'expression asymptotique de l'erreur quadratique intégrée, la Section 4 présente la règle de sélection de la fenêtre, la Section 5 contient les simulations et le reste de l'article contient les preuves de certains résultats.

2. Estimateur à noyau sous censure

Soit X^0 une variable aléatoire réelle représentant une durée de vie de densité f_0 et de fonction de répartition F_0 . Soit C une variable aléatoire réelle de densité g et de fonction de répartition G modélisant la censure. On suppose que X^0 et C sont indépendantes. Soient $(X_i^0, C_i)_{1 \leq i \leq n}$ un n -échantillon i.i.d. de (X^0, C) . Dans le modèle de durées de vie censurées à droite, on observe $(X_i, \Delta_i)_{1 \leq i \leq n}$ où

$$X_i = \min(X_i^0, C_i) \quad \text{et} \quad \Delta_i = I_{[X_i^0 \leq C_i]}.$$

Les $(X_i, \Delta_i)_{1 \leq i \leq n}$ forment un n -échantillon i.i.d. d'un couple aléatoire (X, Δ) , X étant de densité f et de fonction de répartition F , où $1 - F = (1 - F_0)(1 - G)$, de plus on a la relation (cf. Tanner et Wong (1983))

$$E[\Delta / X = x]f(x) = f_0(x)(1 - G(x)). \quad (2.1)$$

Nous poserons pour la suite

$$\phi(x) = f_0(x)(1 - G(x)). \quad (2.2)$$

Nous nous proposons ici d'estimer le taux de hasard $\lambda = f_0/(1 - F_0)$, en utilisant $(X_i, \Delta_i)_{1 \leq i \leq n}$. Comme estimateur de cette fonction, nous retiendrons l'estimateur à

noyau défini par

$$\widehat{\lambda}(x) = \frac{\phi_h(x)}{1 - F_h(x)} \tag{2.3}$$

où compte tenu de (2.1) et (2.2) (voir Collomb (1976))

$$\phi_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) \Delta_i \tag{2.4}$$

est l'estimateur à noyau de ϕ et

$$F_h(x) = \frac{1}{n} \sum_{i=1}^n H\left(\frac{x - X_i}{h}\right) \tag{2.5}$$

est l'estimateur de F (voir par exemple Lejeune et Sarda (1992)), K un noyau, $H(x) = \int_{-\infty}^x K(t)dt$ et $h = h(n)$ est la largeur de fenêtre ou paramètre de régularisation.

De nombreux estimateurs de la fonction de hasard ont été proposés lorsque les durées sont censurées ou non, voir entre autres Singpurwalla et Wong (1983), Hassani *et al.* (1986) et Droesbeke *et al.* (1989) pour des revues bibliographiques. Cependant, à notre connaissance l'estimateur $\widehat{\lambda}$ n'a pas encore été considéré dans la littérature, mais peut être vu comme une extension au cas des données censurées de l'estimateur introduit par Watson et Leadbetter (1964a et 1964b) et étudié ensuite par Murthy (1965), Ahmad (1976) et plus récemment par Youndjé *et al.* (1996). La littérature comporte des résultats intéressants concernant ϕ_h , on peut citer par exemple Colomb (1976), Ango Nzé (1994) (cas des observations mélangeantes). De même Lejeune et Sarda (1992) et Cai et Roussas (1992) (cas des observations mélangeantes) donnent des propriétés asymptotiques de F_h . À partir de ces résultats il est possible de donner des résultats similaires concernant $\widehat{\lambda}$.

3. Décomposition asymptotique de l'ISE

En estimation fonctionnelle, une mesure, très souvent considérée, qui permet d'analyser la performance d'un estimateur est l'erreur quadratique intégrée. Pour l'estimateur $\widehat{\lambda}$ cette quantité est définie par

$$\text{ISE}(\widehat{\lambda}, h) = \int (\widehat{\lambda}(x) - \lambda(x))^2 W(x) dx,$$

W étant une fonction de poids positive. Le lecteur trouvera dans Marron et Padgett (1987) les raisons pour lesquelles il est plus intéressant de considérer l'ISE plutôt que son espérance mathématique, appelée MISE, pour analyser la performance d'un estimateur.

Hypothèses :

Dans toute la suite de l'article nous supposons que :

- La fonction W est à support compact de support S_W et que $S_W^\circ \neq \emptyset$; (H.1)
- La largeur de fenêtre h appartient à l'intervalle d'intérêt

$$H_n = [An^{-\frac{1}{5}-\varepsilon}, Bn^{-\frac{1}{5}+\varepsilon}], \quad 0 < \varepsilon < \frac{1}{5}, \quad 0 < A < B < \infty. \quad (H.2)$$

- K est à support compact, lipschitzien, symétrique et $\int K = 1$; (H.3)
- Les fonctions G et f_0 sont de classe C^2 ; (H.4)
- Il existe $\gamma > 0$ tel que

$$\forall x \in S_W, \quad f(x) \geq \gamma, \quad F_0(x) < 1 - \gamma \quad \text{et} \quad G(x) < 1 - \gamma. \quad (H.5)$$

- La fonction f est bornée. (H.6)

Remarque 3.1. - Les hypothèses (H.1)-(H.6) sont relativement classiques en estimation de la fonction de hasard (voir Patil (1993a)). En ce qui concerne l'hypothèse de compacité du support de W , elle a été introduite d'une part pour des raisons techniques (preuves), et d'autre part pour parer au problème de nullité du dénominateur de $\hat{\lambda}$. Cette condition n'est pas très restrictive, dans le sens où la fonction de poids ne joue pas un rôle prépondérant dans le comportement asymptotique du critère de validation croisée (voir Härdle et Marron (1985)).

THÉORÈME 3.1. - *Sous les hypothèses (H.1) - (H.6) on a :*

$$\sup_{h \in H_n} \left| \frac{\text{ISE}(\hat{\lambda}, h) - C_1(nh)^{-1} - C_2h^4}{C_1(nh)^{-1} + C_2h^4} \right| \rightarrow 0 \quad \text{p.s.} \quad (3.1)$$

où

$$C_1 = \left[\int K^2 \right] \int \frac{\phi(x)}{(1 - F(x))^2} W(x) dx$$

$$\text{et} \quad C_2 = \int (B_\phi(x) + \lambda(x)B_F(x))^2 \frac{W(x)}{(1 - F(x))^2} dx,$$

avec

$$B_\phi(x) = \frac{1}{2} \phi''(x) \left[\int u^2 K(u) du \right] \quad \text{et} \quad B_F(x) = \frac{1}{2} f'(x) \left[\int u^2 K(u) du \right].$$

Remarque 3.2. - Le Théorème 3.1 est une extension au cas des données censurées d'un résultat de Vieu (1991), mais le résultat de Vieu est énoncé sous

α -mélangeance. Ce théorème est à rapprocher du Théorème 3.1 de Marron et Padgett (1987) et du Théorème 3.1 de Patil (1993a). Les expressions respectives de C_1 et C_2 montrent qu'en termes d'erreur quadratique, l'estimateur $\hat{\lambda}$ et les estimateurs $\lambda_{n,1}^0$ et $\lambda_{n,2}^0$ définis dans Patil (1993a) ne sont pas comparables. En fait, tous ces estimateurs ont la même variance, mais des biais différents. Il en découle que selon la loi de la variable en considération, un estimateur peut être meilleur qu'un autre, mais en aucun cas il ne sera meilleur pour toutes les distributions (voir Patil *et al.* (1994) pour une discussion du même type restreinte aux estimateurs $\lambda_{n,1}^0$ et $\lambda_{n,2}^0$). Il faut également noter que l'estimateur étudié par Grégoire (1993) est une généralisation au cas de l'estimation de l'intensité d'un processus de comptage de l'estimateur $\lambda_{n,2}^0$ de Patil (1993a).

4. Le critère de validation croisée

Le Théorème 3.1 montre l'influence de la largeur de fenêtre h sur l'estimateur $\hat{\lambda}$: une largeur trop petite augmente la composante proportionnelle à $(nh)^{-1}$ de l'ISE asymptotique, une largeur trop grande augmente la composante proportionnelle à h^4 . Une méthode usuelle pour choisir la largeur de fenêtre minimisant asymptotiquement l'ISE est la validation croisée. Le critère de validation considéré ici est motivé comme suit :

- L'ISE peut se décomposer comme

$$ISE(\hat{\lambda}, h) = \int \hat{\lambda}^2(x)W(x)dx - 2 \int \left[\frac{\hat{\lambda}(x)W(x)}{1 - F_0(x)} \right] f_0(x)dx + \int \lambda^2(x)W(x)dx.$$

- Le troisième terme est indépendant de h , donc il suffit de choisir h minimisant les deux premiers termes.
- Un estimateur "presque" sans biais du second terme est

$$\frac{2}{n} \sum_{i=1}^n \frac{\hat{\lambda}^{-i}(X_i)W(X_i)\Delta_i}{(1 - F_n(X_i))} \quad \text{avec} \quad \hat{\lambda}^{-i}(x) = \frac{\phi_h^{-i}(x)}{1 - F_h^{-i}(x)}, \quad (4.1)$$

$$\phi_h^{-i}(x) = \frac{1}{(n-1)h} \sum_{\substack{j=1 \\ j \neq i}}^n K\left(\frac{x - X_j}{h}\right) \Delta_j,$$

$$F_h^{-i}(x) = \frac{1}{(n-1)} \sum_{\substack{j=1 \\ j \neq i}}^n H\left(\frac{x - X_j}{h}\right) \quad (4.2)$$

F_n représentant la fonction de répartition empirique définie par

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{[X_i \leq x]}. \quad (4.3)$$

Soit \hat{h} le minimiseur sur H_n du critère de validation

$$CV(h) = \int \hat{\lambda}^2(x)W(x)dx - \frac{2}{n} \sum_{i=1}^n \frac{\phi_h^{-i}(X_i)\Delta_i W(X_i)}{(1 - F_h^{-i}(X_i))(1 - F_n(X_i))}, \quad (4.4)$$

le résultat ci-dessous montre que \hat{h} minimise asymptotiquement l'ISE.

THÉORÈME 4.1. – *Sous les hypothèses (H.1) – (H.6) on a :*

$$\frac{\text{ISE}(\hat{\lambda}, \hat{h})}{\inf_{h \in H_n} \text{ISE}(\hat{\lambda}, h)} \longrightarrow 1 \quad \text{p.s.} \quad (4.5)$$

Remarque 4.1. – Ce résultat est une extension au cas des données censurées du Théorème 2 de Youndjé *et al.* (1996) et est à rapprocher du Théorème 4.1 (resp. 3.2) de Marron et Padgett (1987) (resp. Patil (1993a)). On pourra également consulter les Propositions 4.1 et 4.2 de Grégoire (1993). Nous pensons qu'en s'inspirant des idées de Patil (1993b), il devrait être possible de préciser la vitesse de convergence de la fenêtre optimale \hat{h} . On pourra également consulter Fermanian (1996), qui propose un sélecteur de fenêtre ayant une vitesse de convergence en $O(\frac{1}{\sqrt{n}})$; ce résultat concerne l'estimateur $\lambda_{n,2}^0$ de Patil (1993a).

5. Applications

Afin d'examiner le comportement de l'estimateur $\hat{\lambda}$ et confirmer l'intérêt du Théorème 4.1 nous avons mis en oeuvre une étude par simulations. Nous avons illustré nos résultats en considérant la valeur absolue d'une loi logistique et un mélange de lois de Weibull. La variable de censure dans ces exemples suit une loi exponentielle.

Nous posons pour la suite

$$h_0 = \arg \min \text{ISE}(\hat{\lambda}, h) \quad \text{et} \quad \hat{h} = \arg \min CV(\hat{\lambda}, h)$$

Pour les différents exemples, le noyau K choisi est le noyau d'Epanechnikov (1969) connu pour ses qualités optimales et défini par : $K(x) = \frac{3}{4}(1 - x^2)$ si $|x| \leq 1$ et 0 sinon. Il est bien connu (voir Härdle et Marron (1985)) que la qualité de l'estimateur dépend peu de la fonction de poids W , elle est définie pour les besoins de l'étude par $W(x) = I_{[0.25;3]}(x)$.

Exemple 1

Nous considérons ici que :

$X^0 = |Y|$ où Y suit la loi logistique dont la fonction de répartition est donnée par

$$\hat{F}_Y(t) = \frac{1}{1 + \exp(-t)}$$

La fonction de hasard de X^0 est dans ce cas :

$$\lambda(t) = \begin{cases} \frac{1}{1+\exp(-t)} & \text{si } t \geq 0, \\ 0 & \text{si } t < 0. \end{cases}$$

La censure C dans cet exemple suit la loi exponentielle de paramètre 0.05. Conformément aux hypothèses, les observations de la variable C sont générées indépendamment de celles de Y .

La figure 1 représente les courbes du critère de validation croisée (CV) et de l'erreur quadratique intégrée (ISE) sur un échantillon de taille $n = 1000$. On remarque qu'elles ont la même allure et que le minimum est atteint approximativement au même point, ce qui illustre bien le Théorème 4.1.

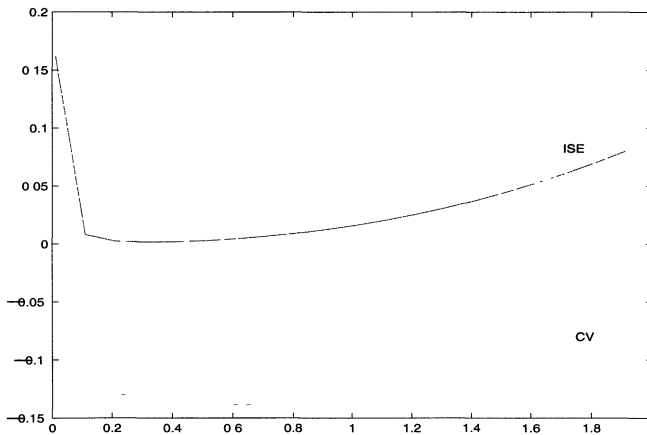


FIGURE 1
ISE(h) et CV(h) pour un échantillon de taille n = 1000

Sur la figure 2 sont décrits, l'estimateur $\hat{\lambda}$ (ligne en pointillé) avec pour fenêtre celle obtenue par validation croisée ($\hat{h} = 0.71$) et la fonction de hasard réelle λ (ligne continue). La taille de l'échantillon est 1000, l'approximation est assez bonne car l'estimateur ne s'écarte de la courbe réelle qu'au voisinage de 0.5 où il y a peu d'observations, et où l'on constate l'effet du phénomène de bord.

Le tableau 1 résume les résultats obtenus sur 100 replications pour cet exemple. La remarque faite sur la figure 1 sur un seul échantillon est confirmée à partir des replications lorsqu'on observe les moyennes et les médianes de la table 1. Il convient de dire un mot sur l'extrême "variabilité" (donnée par la colonne des écart-types) des quantités $ISE(h_0)$ et $ISE(\hat{h})$. En fait, malgré le Théorème 3.1 qui présente une expression asymptotique de l'ISE (très régulière), Marron et Wand (1992) dans le cadre de l'estimation d'une densité ont montré que pour certains exemples il faut parfois des échantillons de taille 1.000.000 pour stabiliser l'ISE. Nous observons

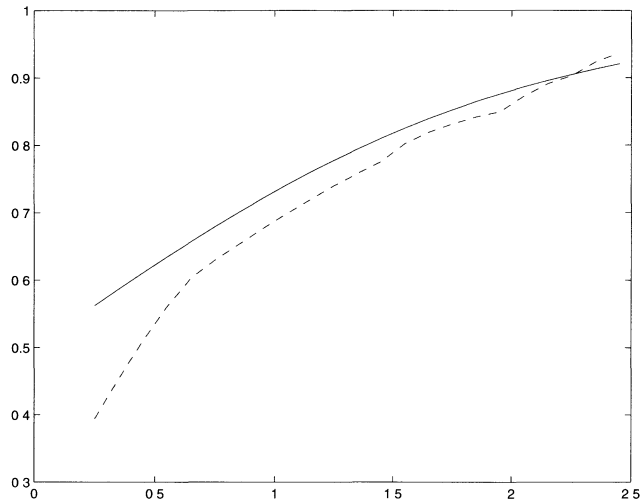


FIGURE 2

Fonction de hasard estimée (- - -) pour un échantillon de taille $n = 1000$ et $\hat{h} = 0.71$. Fonction de hasard (—)

donc ici ce phénomène et d'ailleurs on remarque bien que les écart-types diminuent lorsqu'on augmente la taille des échantillons. Quant à l'ISE(\hat{h}) il faut noter que \hat{h} est un estimateur (v.a.) de h_0 qui est lui-même une variable aléatoire, il s'ensuit donc une plus grande dispersion. Tout ce qui est dit au sujet de la variabilité pour la table 1 est valable aussi pour les tables 2 et 3 présentées ci-dessous.

TABLEAU 1

Paramètres de centralité et de dispersion de l'erreur quadratique intégrée en \hat{h} et en h_0 . La censure suit la loi exponentielle de paramètre 0.05

	ISE(\hat{h})			ISE(h_0)		
	moyenne	médiane	écart type	moyenne	médiane	écart type
$n = 300$	0.0166	0.0102	0.0173	0.0086	0.0070	0.0079
$n = 500$	0.0181	0.0094	0.0227	0.0071	0.0057	0.0052
$n = 1000$	0.0059	0.0046	0.0045	0.0037	0.0033	0.0024

Exemple 2

Ce cas est identique à l'exemple 1 à l'exception de la loi suivie par la censure. C'est la loi exponentielle de paramètre 0.2, ainsi la variable C aura tendance à prendre de plus petites valeurs que dans l'exemple précédent et provoquera donc plus de censure. En effet les calculs montrent que l'on passe d'un taux de censure (en moyenne

TABLEAU 2

Paramètres de centralité et de dispersion de l'erreur quadratique intégrée en \hat{h} et en h_0 . La censure suit la loi exponentielle de paramètre 0.2

	$ISE(\hat{h})$			$ISE(h_0)$		
	moyenne	médiane	écart type	moyenne	médiane	écart type
$n = 300$	0.0248	0.0156	0.0316	0.0101	0.0086	0.0077
$n = 500$	0.0249	0.0121	0.0396	0.0083	0.0067	0.0058
$n = 1000$	0.0094	0.0057	0.0112	0.0044	0.0037	0.0028

par échantillon) de 6% dans l'exemple 1 à 22% dans celui-ci. La table 2 récapitule les résultats pour cet exemple et confirme le fait qu'un taux de censure élevé conduit à des résultats moins bons.

Exemple 3

Ici la loi de X_0 est un mélange de deux lois de Weibull, dont la fonction de densité est définie par

$$f_0(x) = \frac{1}{3}w_{\frac{1}{2}}(x) + \frac{2}{3}w_2(x)$$

où

$$w_\gamma(x) = \gamma x^{\gamma-1} \exp(-x^\gamma) \text{ si } x > 0 \text{ et } w_\gamma(x) = 0 \text{ si } x \leq 0.$$

est la densité d'une loi de Weibull de paramètre γ . La fonction de hasard est alors définie par

$$\lambda(x) = \frac{\sqrt{x} \exp(-\sqrt{x}) + 8x^2 \exp(-x^2)}{2x (\exp(-\sqrt{x}) + 2 \exp(-x^2))}, \quad x > 0.$$

La censure suit la loi exponentielle de paramètre 0.05. Ici le taux moyen de censure par échantillon est de 5.5%.

Les courbes représentatives de la fonction de hasard estimée, avec $\hat{h} = 0.301$ obtenue par validation croisée, et la fonction de hasard réelle concordent assez bien dans la région concave (figure 3).

La table 3 ci-après résume les paramètres obtenus dans le cas de 200 replications. On peut tirer les mêmes conclusions que dans les exemples 1 et 2, sauf qu'ici pour obtenir une relative stabilité de l'ISE nous avons été obligés de travailler avec des échantillons de tailles plus grandes.

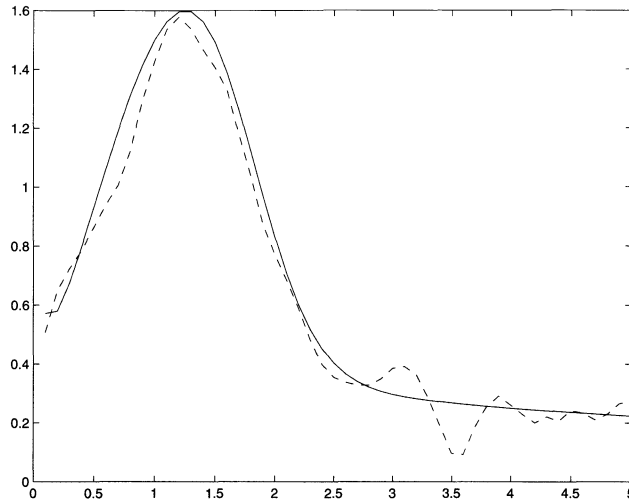


FIGURE 3

Fonction de hasard estimée (- - -) pour un échantillon de taille $n = 1000$ et $\hat{h} = 0.301$. Fonction de hasard (—)

TABLEAU 3

Paramètres de centralité et de dispersion de l'erreur quadratique intégrée en \hat{h} et en h_0 . La censure suit la loi exponentielle de paramètre 0.05

	$ISE(\hat{h})$			$ISE(h_0)$		
	moyenne	médiane	écart type	moyenne	médiane	écart type
$n = 1000$	0.0155	0.0105	0.0181	0.0063	0.0041	0.0063
$n = 2000$	0.0092	0.0056	0.0102	0.0041	0.0029	0.0041
$n = 3000$	0.0059	0.0040	0.0060	0.0027	0.0019	0.0024

6. Preuve du Théorème 3.1

Puisque $\hat{\lambda}$ est un rapport d'estimateurs δ -suites, en suivant les idées de Marron et Härdle (1986), pour analyser l'ISE il est intéressant d'introduire les critères d'erreur supplémentaires suivants :

$$ISE^*(\hat{\lambda}, h) = \int (\hat{\lambda}(x) - \lambda(x))^2 \left(\frac{1 - F_h(x)}{1 - F(x)} \right)^2 W(x) dx$$

$$\text{et } MISE^*(\hat{\lambda}, h) = \mathbb{E}[ISE^*(\hat{\lambda}, h)].$$

Marron et Härdle (1986) ont montré que, sous les hypothèses du Théorème 3.1 on a :

$$\sup_{h \in H_n} \left| \frac{\text{ISE}(\widehat{\lambda}, h) - \text{ISE}^*(\widehat{\lambda}, h)}{\text{ISE}(\widehat{\lambda}, h)} \right| \longrightarrow 0 \quad \text{p.s.} \quad (6.2.1)$$

et

$$\sup_{h \in H_n} \left| \frac{\text{ISE}^*(\widehat{\lambda}, h) - \text{MISE}^*(\widehat{\lambda}, h)}{\text{MISE}^*(\widehat{\lambda}, h)} \right| \longrightarrow 0 \quad \text{p.s.} \quad (6.2.2)$$

Il résulte de (6.2.1) et (6.2.2) que la preuve du Théorème 3.1 découle du Lemme 1 ci-dessous dont la preuve est classique et repose sur les Propositions 3 et 4 données dans l'annexe.

LEMME. – *Sous les hypothèses du Théorème 3.1 on a :*

$$\text{MISE}^*(\widehat{\lambda}, h) = \frac{C_1}{nh} + C_2 h^4 + o\left(\frac{1}{nh} + h^4\right)$$

où C_1 et C_2 sont données dans le Théorème 3.1.

La preuve du Théorème 4.1 est assez technique, néanmoins elle peut être obtenue auprès des auteurs.

Annexe

Dans cette section nous allons rappeler quelques résultats de base concernant les estimateurs ϕ_h de ϕ et F_h de F , utiles dans la démonstration du Théorème 3.1.

PROPOSITION 1 (cf. Collomb (1976)). – *Sous les hypothèses (H.1) – (H.4) on a :*

$$\begin{aligned} (i) \quad & \sup_{x \in S_W} |\phi_h(x) - \mathbb{E}\phi_h(x)| = O\left(\frac{\log n}{nh}\right)^{\frac{1}{2}} \quad \text{p.s.}, \\ (ii) \quad & \sup_{x \in S_W} |\mathbb{E}\phi_h(x) - \phi(x)| = O(h^2) \end{aligned}$$

On a des résultats très proches concernant F_h .

PROPOSITION 2 (cf. Winter (1979)). – *Sous les hypothèses (H.1) – (H.4) on a :*

$$\begin{aligned} (i) \quad & \sup_{x \in S_W} |F_h(x) - \mathbb{E}F_h(x)| = O\left(\frac{\log n}{n}\right)^{\frac{1}{2}} \quad \text{p.s.}, \\ (ii) \quad & \sup_{x \in S_W} |\mathbb{E}F_h(x) - F(x)| = O(h^2) \end{aligned}$$

Les Propositions 3 et 4 ci-dessous donnent la décomposition asymptotique de l'erreur quadratique moyenne ponctuelle des estimateurs ϕ_h et F_h .

PROPOSITION 3 (cf. Collomb (1976) et Lejeune et Sarda (1992)). – *Sous les hypothèses (H.1) – (H.4) on a :*

$$\begin{aligned} (i) \quad \mathbb{E}(\phi_h(x) - \mathbb{E}\phi_h(x))^2 &= \frac{\phi(x)}{nh} \left[\int K^2 \right] + o\left(\frac{1}{nh}\right), \\ (ii) \quad \mathbb{E}(F_h(x) - \mathbb{E}F_h(x))^2 &= \frac{F(x)(1-F(x))}{n} - \frac{hf(x)}{n} \int 2uK(u)H(u)du + o\left(\frac{h^2}{n}\right). \end{aligned}$$

PROPOSITION 4 (cf. Collomb (1976) et Lejeune et Sarda (1992)). – *Sous les hypothèses (H.1) – (H.4) on a :*

$$\begin{aligned} (i) \quad \mathbb{E}\phi_h(x) - \phi(x) &= \frac{h^2}{2} \phi''(x) \int u^2 K(u) du + o(h^2), \\ (ii) \quad \mathbb{E}F_h(x) - F(x) &= \frac{h^2}{2} f'(x) \int u^2 K(u) du + o(h^2). \end{aligned}$$

COROLLAIRE 1. – *Il découle des Propositions 1 et 2 que sous les hypothèses (H.1) – (H.4) on a :*

$$\begin{aligned} (i) \quad \sup_{i=1, \dots, n} \sup_{x \in S_W} |\phi_h^{-i}(x) - \mathbb{E}\phi_h(x)| &= O\left(\frac{\log n}{nh}\right)^{\frac{1}{2}} \quad \text{p.s.}, \\ (ii) \quad \sup_{i=1, \dots, n} \sup_{x \in S_W} |F_h^{-i}(x) - \mathbb{E}F_h^{-i}(x)| &= O\left(\frac{\log n}{n}\right)^{\frac{1}{2}} \quad \text{p.s.} \end{aligned}$$

Remerciements. – Nous tenons à remercier les rapporteurs pour la lecture détaillée de l'article ainsi que pour leurs suggestions pertinentes.

Références

- [1] AHMAD I.A. (1976), Uniform strong convergence of the generalized failure rate estimate. *Bull. Math. Statist.*, **17**, pp. 77–84.
- [2] ANGO-NZÉ P. (1994), *Critères d'ergodicité de modèles markoviens. Estimation non paramétrique sous hypothèses de dépendance*. Thèse à Univ Paris IX, France.

- [3] CAI Z. et ROUSSAS G.G. (1992), Uniform strong estimation under α -mixing, with rates. *Stat. and Prob. Lett.*, pp. 47–55.
- [4] COLLOMB G. (1976), *Estimation non paramétrique de la régression par la méthode du noyau*. Thèse à Univ. P. Sabatier, Toulouse, France.
- [5] DROESBEKE J.J., FICHET B. et TASSI P. (1989), *Analyse statistique des durées de vie*. Ed. Economica.
- [6] EPANECHNIKOV V. (1969), Nonparametric estimates of a multivariate probability density. *Theory Probab. Appl.* **14**, pp. 153–158.
- [7] FERMANIAN J.D. (1996), Choix d'une fenêtre optimale en $n^{\frac{1}{2}}$ pour l'estimateur non paramétrique des fonctions de hasard, *XVIIème Rencontre Franco-Belge de Statisticiens, Marne-la-Vallée*.
- [8] GRÉGOIRE G. (1993), Least squares Cross-Validation for counting process intensities. *Scan. J. of Statist.*, **20**, pp. 343–360.
- [9] HASSANI S., SARDA P. et VIEU P. (1986), Approche non paramétrique en théorie de la fiabilité. *Rev. Statist. Appl.*, **35**, pp. 27–41.
- [10] HÄRDLE W. et MARRON J.S. (1985), Optimal bandwidth selection in nonparametric regression function. *Ann. Statist.*, **13**, pp. 1465–1481.
- [11] LEJEUNE M. et SARDA P. (1992), Smooth estimation of distribution and density function. *Comp. Statist. Data Analysis*, **14**, pp. 457–471.
- [12] MARRON J.S. et PADGETT W.J. (1987), Asymptotically optimal bandwidth selection for kernel density estimators from randomly right censored samples. *Ann. Statist.*, **15**, pp. 1520–1535.
- [13] MARRON J.S. et HÄRDLE W. (1986), Random approximations of some measure of accuracy in nonparametric curve estimation. *J of Multiv. Ana.*, **19**, pp. 1–13.
- [14] MARRON J.S. et WAND M.P. (1992), Exact mean integrated squared errors. *Ann. Statist.*, **20**, pp. 712–736.
- [15] MURTHY V.K. (1965), Estimation of jumps, reliability and hazard rate. *Ann. Statist.*, **36**, pp. 1032–1040.
- [16] PATIL P.N. (1993a), Bandwidth choice for nonparametric hazard rate estimation. *J. Statist. Plann. Inference*, **35**, pp. 15–30.
- [17] PATIL P.N. (1993b), On the least squares cross-validation bandwidth in hazard rate estimation. *Ann. Statist.*, **21**, pp. 1792–1810.
- [18] PATIL P.N., WELLS M.T. et MARRON J.S. (1994), Some heuristic of kernel based estimators of ratio functions. *Nonparametric Statist.*, **4**, pp. 203–209.
- [19] SARDA P. et VIEU P. (1991), Smoothing parameter selection in hazard estimation. *Stat. and Prob. Lett.*, pp. 429–434.
- [20] SINGPURWALLA N.D. et WONG M.Y. (1983), Kernels estimators of the failure rate and density estimation : an analogy. *J.A.S.A* vol. **78**, n° 382.
- [21] TANNER M.A. et WONG W. H. (1983), The estimation of the hazard function from randomly censored data by kernel method. *Ann. Statist.*, **11**, pp. 989–993.

- [22] VIEU P. (1991), Quadratic errors for nonparametric estimates under dependence. *J. of Multiv. Ana.*, vol. **39**, pp. 324–347.
- [23] WATSON G.S. et LEADBETTER M.R. (1964a), Hazard analysis I. *Biometrika*, **51**, pp. 175–184.
- [24] WATSON G.S. et LEADBETTER M.R. (1964b), Hazard analysis II. *Sankhya*, Ser. A, **26**, pp. 110–116.
- [25] WINTER B.B. (1979), Convergence rate of perturbed empirical distribution function. *J. Appl. Prob.*, **16**, pp. 163–173.
- [26] YOUNDJÉ É., SARDA P. et VIEU P. (1996), Optimal smooth hazard estimates. *Test*, vol **5**, N.2, pp. 374–379.