



ELSEVIER

Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com



Statistique

Comportement asymptotique de l'estimateur non paramétrique de la fonction de renouvellement associée à des variables aléatoires positives indépendantes et non stationnaires



Asymptotic behavior of the nonparametric estimator for the renewal function associated with independent and nonstationary positive random variables

Michel Harel^{a,b}, Fy Mamenosoa Ravelomanantsoa^c

^a IUFM du Limousin, 209, bd de Vanteaux, 87036 Limoges cedex, France

^b IMT (UMR CNRS 5219), université Paul-Sabatier, 31062 Toulouse cedex, France

^c Université d'Antananarivo, faculté des sciences, département de mathématiques et informatique, BP 906, 101 Antananarivo, Madagascar

INFO ARTICLE

Historique de l'article :

Reçu le 5 juillet 2013

Accepté le 11 juillet 2013

Disponible sur Internet le 20 août 2013

Présenté par Paul Dehevels

RÉSUMÉ

Nous estimons la fonction de renouvellement, associée à des variables aléatoires positives indépendantes et non stationnaires, par une somme de U -statistiques ou de V -statistiques. Nous étudions ses propriétés asymptotiques : normalité asymptotique, convergence presque sûre ainsi que sa convergence faible. Cette dernière sera traitée dans l'espace des fonctions définies sur $[0, 1]$ continues à droite et admettant une limite à gauche, muni de la topologie de Skorokhod.

© 2013 Académie des sciences. Publié par Elsevier Masson SAS. Tous droits réservés.

ABSTRACT

We estimate the renewal function, for independent and nonstationary positive random variables, by considering a sum of U -statistics or V -statistics. We study its asymptotic properties: asymptotic normality, almost sure convergence and weak convergence. This last property will be established on the space of functions defined on $[0, 1]$ that are right-continuous and have left-hand limits, provided with the Skorokhod topology.

© 2013 Académie des sciences. Publié par Elsevier Masson SAS. Tous droits réservés.

1. Introduction

En considérant le cas d'une suite de variables aléatoires stationnaires i.i.d., Frees [3] avait étudié la normalité asymptotique de l'estimateur empirique de la fonction de renouvellement sous forme de somme de U -statistiques. M. Harel, C.A. O'Kinneide et H. Schneider [4], avaient généralisé les résultats dans l'espace de Skorokhod [1]. Ils ont recherché la

Adresses e-mail : michel.harel@unilim.fr (M. Harel), fy.mamenosoa@gmail.com (F.M. Ravelomanantsoa).

convergence faible de cet estimateur, en utilisant une linéarisation de la différence entre l'estimateur empirique et la fonction de renouvellement.

En considérant le cas d'un processus non stationnaire et absolument régulier, Elharfaoui et Harel [2] avaient utilisé la U -statistique pour déterminer un estimateur de la fonction de répartition empirique et avaient étudié les comportements asymptotiques de cet estimateur.

Dans ce travail, nous considérons des variables aléatoires positives indépendantes et non stationnaires ; nous utilisons aussi un estimateur de la fonction de renouvellement sous forme de somme de U -statistiques ou de V -statistiques, car ce sont des estimateurs non paramétriques et sans biais de la fonction de renouvellement. Nous étudions les comportements asymptotiques de cet estimateur.

Soit une suite $(X_i)_{i \in \mathbb{N}}$ de variables aléatoires positives indépendantes et non stationnaires, de fonctions de répartition respectives F_i ($i \geq 1$), telle qu'il existe une suite $(X_i^*)_{i \geq 1}$ de variables aléatoires positives, indépendantes et stationnaires associée à $(X_i)_{i \geq 1}$ et de fonctions de répartition F_0 , telle que F_i converge vers F_0 . Supposons que les F_i ($i \geq 0$) admettent une espérance μ_i finie positive et une variance finie σ_i^2 .

Définition 1.1. Une U -statistique [2] pour une fonctionnelle régulière $\theta(F)$ de degré k est définie par :

$$U_n = \binom{n}{k}^{-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} \psi(X_{i_1} + \dots + X_{i_k})$$

où ψ est une fonction symétrique réelle définie sur \mathbb{R}^k appelée noyau, et :

$$\theta(F) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \psi(x_1, \dots, x_k) dF_{1, \dots, k}(x_1, \dots, x_k)$$

où $F_{1, \dots, k}$ est la fonction de répartition produit de (X_1^*, \dots, X_k^*) .

En posant $\psi = I(X_{i_1} + \dots + X_{i_k} \leq t)$, alors le processus empirique défini par :

$$\hat{F}_n^{(k)}(t) = \binom{n}{k}^{-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} I(X_{i_1} + \dots + X_{i_k} \leq t)$$

est une U -statistique.

Nous définissons alors la fonction de renouvellement par :

$$H(t) = \sum_{k \geq 1} F^{(k)}(t)$$

où $F^{(k)}(t) = P(X_1^* + \dots + X_k^* \leq t)$.

Pour $t > 0$ fixé, il semble naturel d'estimer $H(t)$ par une somme finie $m = m(n)$, dépendant de n , avec $m \leq n$ et $m \uparrow \infty$ quand $n \uparrow \infty$, des estimateurs empiriques de $F^{(k)}(t)$ définis par :

$$\hat{H}_n(t) = \sum_{k=1}^m \hat{F}_n^{(k)}(t)$$

qui est un estimateur non paramétrique de $H(t)$ sous forme de somme de U -statistiques.

2. Normalité asymptotique de $\hat{H}_n(t)$

L'estimateur défini ci-dessus est une variante de la version de Frees [3] applicable sans la restriction de la positivité des variables aléatoires X_i .

Le processus X_i est à valeurs dans un espace euclidien de dimension finie \mathbb{H} . Pour spécifier nos hypothèses sur le processus, nous devons définir des copies de \mathbb{H} . Dorénavant, notons \mathbb{H}_i ($i \geq 1$) une suite infinie de copies de \mathbb{H} . Le processus au temps i prendra sa valeur dans \mathbb{H}_i et chaque \mathbb{H}_i sera la i -ème composante de \mathbb{H}^∞ . Nous définissons ainsi une suite de variables aléatoires stationnairement asymptotiquement géométrique par :

Définition 2.1. Le processus $(X_i)_{i \geq 1}$, avec sa mesure de probabilité P sur \mathbb{H}^∞ , est stationnairement asymptotiquement géométrique s'il existe un processus stationnaire $(X_i^*)_{i \geq 1}$ et indépendant de $(X_i)_{i \geq 1}$ avec Q comme mesure de probabilité sur \mathbb{H}^∞ , et un réel τ ($0 < \tau < 1$), tel que pour tout $i \geq 1$,

$$|P - Q|_{\sigma_{\mathbb{H}^\infty}} \leq \tau^i \tag{1}$$

où $\mathbb{H}_i^\infty = \bigoplus_{j=i}^\infty \mathbb{H}_j = \mathbb{H}_i \oplus \mathbb{H}_{i+1} \oplus \dots$.

Dans toute la suite, nous identifierons \mathbb{H} avec \mathbb{R} l'espace des réels.
 La normalité asymptotique de $\sqrt{n}(\hat{H}_n(t) - H(t))$ s'énonce comme suit :

Théorème 2.2. *Supposons que la suite de variables aléatoires positives $(X_i)_i$ soit stationnairement asymptotiquement géométrique et vérifie la condition (1). Pour tout $i \geq 1$, supposons que $\sigma_i^2 < \infty$.*

Supposons de plus que, pour $r > 6$ et pour tout $\varepsilon > 0$:

$$\sup_{i \geq 1} E|X_i|^{r+\varepsilon} < \infty, \quad E|X_i| > 0 \quad \text{pour tout } i \geq 1 \text{ et } n = \mathcal{O}(m^{r-4}).$$

Alors, pour chaque $t > 0$:

$$\sqrt{n}(\hat{H}_n(t) - H(t)) \xrightarrow{\mathcal{L}} N(0, \sigma_2^2),$$

et on a

$$\sigma_2^2 = \sum_{s=1}^{\infty} \sum_{r=1}^{\infty} r s \xi_{rs}(1) < \infty$$

où $\xi_{rs}(1) = \text{Cov}(F^{(r-1)}(t - X_1^*), F^{(s-1)}(t - X_1^*))$.

3. Convergence p.s de $\hat{H}_n(t)$

La convergence p.s de la différence entre l'estimateur empirique et la fonction de renouvellement s'énonce comme suit :

Théorème 3.1. *On suppose que la suite (X_i) de variables aléatoires positives est indépendante, stationnairement asymptotiquement géométrique et vérifie la condition (1) pour la suite (X_i^*) , alors :*

$$\frac{1}{n}(\hat{H}_n(t) - H(t)) \xrightarrow{p.s} 0 \quad (n \rightarrow \infty).$$

4. Convergence faible de $H_n(t)$ par rapport à la topologie de Skorokhod

En revanche, nous montrerons la convergence faible par rapport à la topologie de Skorokhod de l'estimateur sous forme d'une somme infinie de V -statistiques défini par Schneider [5] :

$$H_n(t) = \sum_{k=1}^{\infty} \hat{F}_{ne}^{(k)}(t)$$

où $\hat{F}_{ne}^{(k)}(t)$ est un estimateur défini par :

$$\hat{F}_{ne}^{(k)}(t) = n^{-k} \sum_{1 \leq i_1 \leq \dots \leq i_k \leq n} I(X_{i_1} + \dots + X_{i_k} \leq t).$$

Mais, avant de passer à la convergence faible de $\sqrt{n}(H_n(t) - H(t))$, nous étudions, dans un premier temps, la convergence faible du processus $\sqrt{n}(H_n(t) - \tilde{H}_n(t))$, où $\tilde{H}_n(t)$ est définie par :

$$\tilde{H}_n(t) = \sum_{k=1}^{\infty} \tilde{F}_{ne}^{(k)}(t)$$

avec

$$\tilde{F}_{ne}^{(k)}(t) = n^{-k} \sum_{1 \leq i_1 \leq \dots \leq i_k \leq n} F_{i_1} * \dots * F_{i_k}(t)$$

où $*$ est le produit de convolution.

Nous noterons \mathbf{D} l'espace des fonctions continues à droite admettant une limite à gauche sur $[0, 1]$ dotée de la topologie de Skorokhod et \mathbf{C} l'espace des fonctions continues sur $[0, 1]$. Le processus $H_n(t)$ sera à valeurs dans \mathbf{D} .

En généralisant la linéarisation d'Harel, O'Cinneide et Schneider [4], nous avons :

$$\sqrt{n}(H_n(t) - \tilde{H}_n(t)) = \sqrt{n}(\hat{F}_{ne}^{(1)} - \tilde{F}_{ne}^{(1)}) * \tilde{G}_n(t) + \sqrt{n}(\hat{F}_{ne}^{(1)} - \tilde{F}_{ne}^{(1)})^2 * H_n * \tilde{G}_n(t)$$

où $\tilde{G}_n(t) = \tilde{H}_n * \tilde{H}_n(t)$.

En utilisant cette décomposition, on montre le théorème suivant :

Théorème 4.1. *Supposons que la suite $(X_i)_i$ de variables aléatoires soit stationnairement asymptotiquement géométrique, vérifie la condition (1), et que les fonctions F_i admettent une densité f_i ($i \in \mathbb{N}^*$) vérifiant la condition suivante : il existe un réel $K > 0$ tel que :*

$$\sup_{i \in \mathbb{N}^*} \sup_{t \in \mathbb{R}} f_i(t) \leq K,$$

alors le processus $\sqrt{n}(H_n(t) - H(t))$ ($\in \mathbf{D}$) converge faiblement vers $W_\infty^{(1)} * G$ ($\in \mathbf{C}$) par rapport à la topologie de Skorokhod.

Le processus $W_\infty^{(1)}$ est un processus gaussien centré de covariance, pour $0 \leq t_1 \leq t_2 \leq 1$:

$$\text{Cov}(W_\infty^{(1)}(t_1), W_\infty^{(1)}(t_2)) = F_0(t_1)(1 - F_0(t_2)),$$

$G = \lim_{n \rightarrow \infty} G_n$ existe et $W_\infty^{(1)} * G$ a pour covariance :

$$\text{Cov}(W_\infty^{(1)} * G(t_1), W_\infty^{(1)} * G(t_2)) = \int_0^{t_1} G(t_1 - s)G(t_2 - s) dF_0(s) - g(t_1)g(t_2)$$

où $g(u) = \int_0^u G(u - s) d(F_0(s))$ et $u \in [0, 1]$.

Remerciement

Les auteurs remercient la Coopération française pour ses contributions financières.

Références

- [1] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.
- [2] E. Elharfaoui, M. Harel, Central limit theorem of the smoothed empirical distribution functions for asymptotically stationary absolutely regular stochastic processes, *J. Appl. Math. Stoch. Anal.* (2008), Article ID 735436 (18 p.).
- [3] E.W. Frees, Nonparametric renewal function estimation, *Ann. Stat.* 14 (4) (1986) 1366–1378.
- [4] M. Harel, C.A. O’Cinneide, H. Schneider, Asymptotics of the sample renewal function, *J. Math. Anal. Appl.* 189 (1995) 240–255.
- [5] H. Schneider, B. Lin, C.A. O’Cinneide, Comparison of nonparametric estimators of the renewal function, *Appl. Stat.* 39 (1987) 55–61.