

Statistique

Une loi fonctionnelle uniforme du logarithme non standard pour les accroissements du processus empirique multivarié

Davit Varron

ENSAI, 6, rue Blaise-Pascal, 35170 Bruz cedex, France

Reçu le 4 décembre 2004 ; accepté après révision le 22 mai 2006

Disponible sur Internet le 2 août 2006

Présenté par Paul Deheuvels

Résumé

Soit $(Z_i)_{i \geq 1}$ une suite i.i.d. dont la loi commune sur \mathbb{R}^d admet une densité continue et strictement positive sur un ouvert borné O de \mathbb{R}^d . Soit $H \subset O$ un fermé d'intérieur non vide, et soit $(h_n)_{n \geq 1}$ une suite de constantes vérifiant, lorsque $n \rightarrow \infty$, $0 < h_n < 1$, $h_n \downarrow 0$, $nh_n \uparrow \infty$, $nh_n / \log n \rightarrow c$, avec $0 < c < \infty$. Pour tous $z \in H$ $(s_1, \dots, s_d) \in [0, 1)^d$ et pour tout $n \geq 1$, on définit

$$\Delta_n(z, h_n, s) := \frac{\sum_{i=1}^n 1_{[0, s_1) \times \dots \times [0, s_d)}\left(\frac{Z_i - z}{h_n^{1/d}}\right)}{cf(z) \log n}.$$

Nous établissons une loi fonctionnelle uniforme non standard du logarithme pour les processus $\Delta_n(\cdot, h_n, z)$, $z \in H$. Ce résultat étend le résultat de Deheuvels et Mason (1992) au cas multivarié. **Pour citer cet article : D. Varron, C. R. Acad. Sci. Paris, Ser. I 343 (2006).**

© 2006 Académie des sciences. Publié par Elsevier Masson SAS. Tous droits réservés.

Abstract

A nonstandard uniform functional law of the logarithm for the increments of the multivariate empirical process. Let $(Z_i)_{i \geq 1}$ be an i.i.d. sequence being such that Z_1 has a continuous, strictly positive density f on an open subset $O \subset \mathbb{R}^d$. Let $H \subset O$ be a closed subset with nonempty interior. Let $(h_n)_{n \geq 1}$ denote a sequence of positive constants fulfilling, as $n \rightarrow \infty$, $0 < h_n < 1$, $h_n \downarrow 0$, $nh_n \uparrow \infty$, $nh_n / \log n \rightarrow c$, with $0 < c < \infty$. For each $z \in H$ $(s_1, \dots, s_d) \in [0, 1)^d$ and $n \geq 1$, we set

$$\Delta_n(z, h_n, s) := \frac{\sum_{i=1}^n 1_{[0, s_1) \times \dots \times [0, s_d)}\left(\frac{Z_i - z}{h_n^{1/d}}\right)}{cf(z) \log n}.$$

We establish a nonstandard uniform functional law of the logarithm for the processes $\Delta_n(\cdot, h_n, z)$, $z \in H$. This extends the result of Deheuvels and Mason (1992) to the multivariate case. **To cite this article : D. Varron, C. R. Acad. Sci. Paris, Ser. I 343 (2006).**

© 2006 Académie des sciences. Publié par Elsevier Masson SAS. Tous droits réservés.

1. Introduction et résultat

Soit $(Z_i)_{i \geq 1}$ une suite de variables aléatoires i.i.d. à valeurs dans \mathbb{R}^d . On suppose que la loi des Z_i admet une densité f continue et strictement positive sur un ouvert $O \subset \mathbb{R}^d$. Soit $(h_n)_{n \geq 1}$ une suite de constantes strictement positives. On suppose également que cette suite vérifie les hypothèses suivantes

(HVE1) $0 < h_n < 1$, $h_n \downarrow 0$, $nh_n \uparrow \infty$,

(HVE2) $nh_n / \log n \rightarrow c > 0$.

Pour tout entier $n \geq 1$, pour tout $z \in O$ et pour tout $s = (s_1, \dots, s_d) \in [0, 1)^d$, on note

$$\Delta_n(z, h_n, s) := \frac{\sum_{i=1}^n 1_{[0, s_1) \times \dots \times [0, s_d)} \left(\frac{Z_i - z}{h_n^{1/d}} \right)}{cf(z) \log n}.$$

On note $\mathcal{B}([0, 1)^d)$ l'ensemble des fonctions bornées sur $[0, 1)^d$, et $\|\cdot\|$ la norme sup sur $[0, 1)^d$, c'est à dire $\|f\| := \sup\{|f(s)|, s \in [0, 1)^d\}$. En faisant varier s dans $[0, 1)^d$, les $\Delta_n(z, h_n, \cdot)$ sont des processus stochastiques à trajectoires presque-sûrement dans $\mathcal{B}([0, 1)^d)$, que l'on appelle « accroissements de la fonction de répartition empirique multivariée ». Pour une fonction réelle $g \in \mathcal{B}([0, 1)^d)$ admettant une densité g' par rapport à la mesure de Lebesgue λ , on pose

$$I(g) := \int_{[0, 1)^d} h(g') d\lambda, \quad \text{avec}$$

$$h(x) := x \log x - x + 1, \quad x \in]0, \infty[, \quad h(0) = \infty.$$

Pour toute autre fonction $g \in \mathcal{B}([0, 1)^d)$, on pose $I(g) := \infty$. On définit ensuite pour tout $a > 0$,

$$\Gamma_a := \{g \in \mathcal{B}([0, 1)^d), h(0) = 0, I(g) \leq 1/a\}.$$

Cette Note a pour but de présenter la loi fonctionnelle uniforme du logarithme suivante :

Théorème 1.1 (LFUL non standard). Soient O un ouvert de \mathbb{R}^d et $H \subset O$ un compact d'intérieur non vide, et soit $(Z_i)_{i \geq 1}$ une suite de variables aléatoires i.i.d. dont la loi sur \mathbb{R}^d admet une densité f continue et strictement positive sur O . Soit $(h_n)_{n \geq 1}$ une suite de constantes vérifiant (HVE1) et (HVE2). Alors on a presque sûrement

$$(i) \forall z \in H, \forall g \in \Gamma_{cf(z)}, \quad \lim_{n \rightarrow \infty} \inf\{\|\Delta_n(z', h_n, \cdot) - g\|, z' \in H\} = 0,$$

$$(ii) \lim_{n \rightarrow \infty} \sup_{z \in H} \inf\{\|\Delta_n(z, h_n, \cdot) - g\|, g \in \Gamma_{cf(z)}\} = 0.$$

Remarques. Dans un résultat fondateur, Deheuvels et Mason [3] ont montré notre Théorème 1.1 dans le cas où les Z_i sont univariés et suivent une loi uniforme sur $[0, 1]$. La contribution de cette Note est donc essentiellement l'extension de ce résultat au cas multivarié. En corollaire d'un résultat plus général, Mason [5] a établi un comportement limite analogue au Théorème 1.1, mais avec des ensembles limite de type gaussien, dans le cas où $nh_n / \log n \rightarrow \infty$, $\log(1/h_n) / \log \log n \rightarrow \infty$.

2. Principe de la preuve

On note η_n une variable aléatoire de Poisson d'espérance n , indépendante de l'échantillon $(Z_i)_{i \geq 1}$. On définit, pour tout entier $n \geq 1$, les versions « Poissonisées » des $\Delta_n(z, h_n, \cdot)$, $z \in H$, c'est-à-dire :

$$\Delta \Pi_n(z, h_n, s) := \frac{\sum_{i=1}^{\eta_n} 1_{[0, s_1) \times \dots \times [0, s_d)} \left(\frac{Z_i - z}{h_n^{1/d}} \right)}{cf(z) \log n}.$$

Notre preuve repose sur un principe de grandes déviations uniforme pour les $\Delta \Pi_n(z, h_n, \cdot)$, $z \in H$. En voici l'énoncé :

Proposition 2.1. *On munit $\mathcal{B}([0, 1]^d)$ de la norme $\| \cdot \|$. Pour tout ouvert $O \subset \mathcal{B}([0, 1]^d)$, on a*

$$\liminf_{n \rightarrow \infty} \inf_{z \in H} (cf(z) \log n)^{-1} \log(\mathbb{P}(\Delta \Pi_n(z, h_n, \cdot) \in O)) \geq - \inf_O I. \tag{1}$$

De plus, pour tout fermé $F \subset \mathcal{B}([0, 1]^d)$, on a

$$\limsup_{n \rightarrow \infty} \sup_{z \in H} (cf(z) \log n)^{-1} \log(\mathbb{P}(\Delta \Pi_n(z, h_n, \cdot) \in F)) \leq - \inf_F I. \tag{2}$$

Cette proposition, essentielle à notre preuve, est prouvée en appliquant les techniques décrites par Arcones [1,2] et par Lynch et Sethuraman [4]. Ainsi, on établit d’abord un principe de grandes déviations pour les processus $\Delta \Pi_n(z, h_n, \cdot)$ discrétisés sur une grille dyadique d’ordre arbitraire. Cela revient, à une transformation près, à établir un tel principe pour les mesures aléatoires associées aux $\Delta \Pi_n(z, h_n, \cdot)$, prises sur la partition dyadique correspondante. Nous mettons alors à profit la propriété d’indépendance classique des mesures aléatoires de Poisson, que nous combinons avec un principe de tensorisation des grandes déviations, largement inspiré de Lynch et Sethuraman [4]. Puis nous appliquons le Théorème 2.1 dans [1], en établissant une inégalité exponentielle pour les oscillations des $\Delta \Pi_n(z, h_n, \cdot)$. Cette inégalité est obtenue via la méthode de Cramér–Chernoff, combinée à la structure inductive des rectangles de type $[0, s_1] \times \dots \times [0, s_d]$.

Nous donnons le schéma grossier de la preuve du point (ii) dans le théorème énoncé. On fixe $\epsilon > 0$ et on pose, pour tout entier $k \geq 1$, $n_k := [(1 + \gamma)^k]$, où $\gamma > 0$ est un paramètre qui sera choisi suffisamment petit à la fin de la preuve. On définit, pour tout entier k suffisamment grand, $N_k := \{n_{k-1} + 1, \dots, n_k\}$. Pour tout ensemble $A \subset \mathcal{B}([0, 1]^d)$, on note A^ϵ le ϵ -dilaté de A pour la norme $\| \cdot \|$. Par semi-continuité inférieure de I dans $(\mathcal{B}([0, 1]^d), \| \cdot \|)$, on sait que, pour tout $z \in H$, il existe $\alpha_z > 0$ tel que l’on ait $\inf\{I(g), g \notin \Gamma_{cf(z)}^\epsilon\} > 1 + 3\alpha_z$. Comme f est continue et strictement positive sur O , et par régularité de la fonctionnelle I , on peut trouver un cube ouvert H_z non vide contenant z , et vérifiant les quatre points suivants :

$$H_z \subset O, \quad \inf_{z_1, z_2 \in H_z} \frac{f(z_1)}{f(z_2)} \geq \frac{1 + \alpha_z}{1 + 2\alpha_z}, \quad \bigcup_{z' \in H_z} \Gamma_{cf(z')} \subset \Gamma_{cf(z)}^\epsilon, \quad \mathbb{P}\left(Z_1 \in \bigcup_{z \in H_z} \{z + h_{n_k}[0, 1]^d\}\right) \leq 1/2. \tag{3}$$

On extrait un recouvrement fini de ce recouvrement de H par des ouverts, que l’on note $H \subset \bigcup_{l=1}^L H_{z_l} \subset O$. Le réel $\epsilon > 0$ ayant été choisi arbitrairement, montrer le point (ii) du théorème énoncé revient à montrer que, pour tout $1 \leq l \leq L$, on a presque sûrement

$$\limsup_{n \rightarrow \infty} \sup_{z \in H_{z_l}} \inf_{g \in \Gamma_{cf(z_l)}} \|\Delta_n(z, h_n, \cdot) - g\| \leq 10\epsilon.$$

Nous prouvons ce résultat pour $1 \leq l \leq L$ fixé. Pour cela, on discrétise H_{z_l} de façon usuelle (voir par exemple [5]). Soit $\delta > 0$ un paramètre de discrétisation suffisamment petit. Pour tout entier $n \geq 1$, on construit un recouvrement de H_{z_l} par des pavés disjoints deux à deux

$$H_{z_l} \subset \bigcup_{1 \leq i \leq m_k} \{z_{i, n_k} + (\delta h_{n_k})^{1/d} [0, 1]^d\} =: \bigcup_{1 \leq i \leq m_k} C_{i, k},$$

m_k désignant le nombre minimal de cubes disjoints nécessaire pour faire ce recouvrement. On pose, pour tout entier k suffisamment grand et pour tout $n \in N_k$,

$$H_n(z, s) := \frac{\sum_{i=1}^n 1_{[0, s_1] \times \dots \times [0, s_d]} \left(\frac{Z_i - z}{h_{n_k}^{1/d}}\right)}{cf(z) \log n_k}.$$

On va d’abord montrer que, quel que soit les choix de $\delta > 0$ et $\gamma > 0$, on a presque sûrement

$$\limsup_{n \rightarrow \infty} \sup_{1 \leq i \leq m_k} \inf_{g \in \Gamma_{cf(z_i)}} \|H_n(z_{i, n_k}, \cdot) - g\| \leq 2\epsilon. \tag{4}$$

Pour cela, considérons les probabilités suivantes, pour tout entier k suffisamment grand :

$$\mathbb{P}_k := \mathbb{P}\left(\bigcup_{1 \leq i \leq m_k} \bigcup_{n \in N_k} H_n(z_{i, n_k}, \cdot) \notin \Gamma_{cf(z_l)}^{2\epsilon}\right).$$

Des arguments d'association par bloc classiques associés à une majoration grossière de probabilités ainsi qu'avec une technique de Poissonisation classique (voir par exemple [3]) montrent que, dès lors que $0 < \gamma < \epsilon/8$, on a asymptotiquement en k

$$\mathbb{P}_k \leq 4 \sum_{i=1}^{m_k} \mathbb{P}(\Delta \Pi_n(z_{i,n_k}, h_{n_k}, \cdot) \notin \Gamma_{cf(z_i)}^\epsilon) =: 4 \sum_{i=1}^{m_k} \mathbb{P}_{i,k}.$$

En appliquant (2) ainsi que (3), il vient que, pour tout choix de $\delta > 0$ et de $0 < \gamma < \epsilon/8$, on a, pour tout entier k suffisamment grand,

$$\mathbb{P}_k \leq 4m_k \max_{1 \leq i \leq m_k} \mathbb{P}_{i,k} \leq 4m_k \max_{1 \leq i \leq m_k} \exp\left(-\frac{f(z_{i,n_k})}{f(z_i)}(1 + 2\alpha_{z_i}) \log n_k\right) \leq 4m_k \exp(-(1 + \alpha_{z_i}) \log n_k).$$

Or, pour tout choix de $\delta > 0$ on a $m_k = h_{n_k}^{-1+o(1)}$ lorsque $k \rightarrow \infty$. On en déduit que \mathbb{P}_k est sommable en k . Ainsi, pour tout choix de $\delta > 0$ et pour tout choix de $0 < \gamma < \epsilon/8$, l'assertion (4) est vérifiée. Ensuite, par un choix de $\gamma > 0$ et $\delta > 0$ suffisamment petits, on contrôle la différence entre les $\Delta_n(z, h_n, \cdot)$, $z \in H$ et les $H_n(z_{i,n_k}, \cdot)$, $1 \leq i \leq m_k$ en se servant de la continuité de f sur O ainsi que du principe de grandes déviations cité plus haut (voir (1) et (2)). On obtient ainsi que, pour tout choix de $\gamma > 0$ et $\delta > 0$ suffisamment petits, on a presque sûrement

$$\limsup_{k \rightarrow \infty} \sup_{z \in H_{z_i}} \min_{1 \leq i \leq m_k} \max_{n \in N_k} \|H_n(z_{i,n_k}, \cdot) - \Delta_n(z, h_n, \cdot)\| \leq 7\epsilon. \quad (5)$$

En combinant ensuite (4) et (5), on aboutit au résultat voulu.

Références

- [1] M.A. Arcones, The large deviation principle of stochastic processes, Part 1, *Teor. Veroyatnost. i Primenen.* 47 (1) (2002) 122–150.
- [2] M.A. Arcones, The large deviation principle of stochastic processes, Part 2, *Teor. Veroyatnost. i Primenen.* 48 (4) (2002) 1727–1746.
- [3] P. Deheuvels, D. Mason, Functional laws of the iterated logarithm for the increments of empirical and quantile processes, *Ann. Probab.* 20 (1992) 1248–1287.
- [4] J. Lynch, J. Sethuraman, Large deviations for processes with independent increments, *Ann. Probab.* 15 (2) (1987) 610–627.
- [5] D. Mason, A uniform functional law of the iterated logarithm for the local empirical process, *Ann. Probab.* 32 (2) (2004) 1391–1418.