Test pour un changement de paramètre de queue pour une série temporelle à mémoire longue à l'aide de statistiques de Hill

Davide Giraudo (en collaboration avec Annika Betken et Rafał Kulik)

Poitiers, 21 février 2019

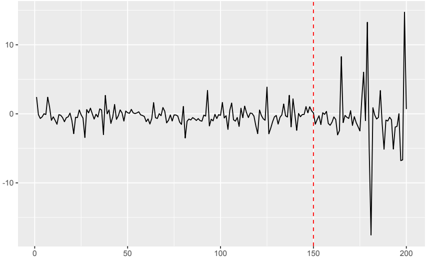
Recherche financée par le *Collaborative Research Center SFB 823*.

Changement de paramètre de queue

Soit $(X_j)_{j\geq 1}$ une série temporelle à variation régulière, *i.e.*, $P(X_j > x) = x^{-\alpha_j} L(x)$, $j = 1, \ldots, n$, où $\alpha_j > 0$ et L est à variation lente ($\lim_{t\to+\infty} L(ct)/L(t) = 1$ pour tout $c > 0$).

Test:

$$H_0 : \alpha_1 = \cdots = \alpha_n \ ,$$
$$H_1 : \alpha_1 = \cdots = \alpha_k \neq \alpha_{k+1} = \cdots = \alpha_n \text{ pour un certain } k \in \{1, \ldots, n-1\} \ .$$



Série temporelle à variation régulière $(X_j)_{j=1}^{200}$ et un changement d'indice de queue en $k = 150$

Estimateur de Hill

Soit $(X_j)_{j\geq 1}$ une suite strictement stationnaire, à variation régulière d'exposant α. La convergence suivante a lieu :

$$\lim_{u\to\infty} \mathrm{E}\left(\log\left(\frac{X_1}{u}\right) \mid X_1 > u\right) = \lim_{u\to\infty} \frac{\mathrm{E}\left(\log\left(\frac{X_1}{u}\right) 1_{\{X_1 > u\}}\right)}{P(X_1 > u)} = \frac{1}{\alpha} =: \gamma.$$

Estimateur de Hill

Soit $(X_j)_{j\geq 1}$ une suite strictement stationnaire, à variation régulière d'exposant α. La convergence suivante a lieu :

$$\lim_{u\to\infty} \mathrm{E}\left(\log\left(\frac{X_1}{u}\right) \mid X_1 > u\right) = \lim_{u\to\infty} \frac{\mathrm{E}\left(\log\left(\frac{X_1}{u}\right) 1_{\{X_1>u\}}\right)}{P(X_1 > u)} = \frac{1}{\alpha} =: \gamma.$$

Estimateur de Hill:

$$\hat{\gamma}_{1,k} := \frac{1}{\frac{1}{k}\sum_{j=1}^{k} 1_{\{X_j>u_n\}}} \frac{1}{k} \sum_{j=1}^{k} \log\left(\frac{X_j}{u_n}\right) 1_{\{X_j>u_n\}} ,$$

où $(u_n)_{n\geq 1}$ est une suite déterministe vérifiant $u_n \to \infty$ et $n\bar{F}(u_n) \to \infty$ (avec $\bar{F}(x) := P(X_1 > x)$).

Estimateur de Hill

Soit $(X_j)_{j\geq 1}$ une suite strictement stationnaire, à variation régulière d'exposant α. La convergence suivante a lieu :

$$\lim_{u\to\infty} \mathrm{E}\left(\log\left(\frac{X_1}{u}\right) \mid X_1 > u\right) = \lim_{u\to\infty} \frac{\mathrm{E}\left(\log\left(\frac{X_1}{u}\right) 1_{\{X_1 > u\}}\right)}{P(X_1 > u)} = \frac{1}{\alpha} =: \gamma.$$

Estimateur de Hill:

$$\hat{\gamma}_{1,k} := \frac{1}{\frac{1}{k}\sum_{j=1}^{k} 1_{\{X_j > u_n\}}} \frac{1}{k} \sum_{j=1}^{k} \log\left(\frac{X_j}{u_n}\right) 1_{\{X_j > u_n\}} \,,$$

où $(u_n)_{n\geq 1}$ est une suite déterministe vérifiant $u_n \to \infty$ et $n\bar{F}(u_n) \to \infty$ (avec $\bar{F}(x) := P(X_1 > x)$).

Statistique de test :

$$\Gamma_n := \max_{1\leq k\leq n-1} \left|\hat{\gamma}_{1,k} - \hat{\gamma}_{1,n}\right| .$$

Processus empirique bi-dimensionnel

On définit

$$T_n(s,t) = \frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{[nt]} 1_{\{X_j > u_n s\}}, \; s \in [1,\infty], \; t \in [0,1].$$

Remarquons que

$$\hat{\gamma}_{1,\lfloor nt \rfloor} = \frac{\frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{[nt]} \log\left(\frac{X_j}{u_n}\right) 1_{\{X_j > u_n\}}}{\frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{\lfloor nt \rfloor} 1_{\{X_j > u_n\}}} = (T_n(1,t))^{-1} \int_1^\infty s^{-1} T_n(s,t) \, \mathrm{d}s \, .$$

Processus empirique bi-dimensionnel

On définit

$$T_n(s,t) = \frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{[nt]} 1_{\{X_j > u_n s\}}\ ,\ s \in [1,\infty],\ t \in [0,1].$$

Remarquons que

$$\hat{\gamma}_{1,\lfloor nt \rfloor} = \frac{\frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{[nt]} \log\left(\frac{X_j}{u_n}\right) 1_{\{X_j > u_n\}}}{\frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{\lfloor nt \rfloor} 1_{\{X_j > u_n\}}} = \left(T_n(1,t)\right)^{-1} \int_1^{\infty} s^{-1} T_n(s,t)\, \mathrm{d}s\ .$$

On considère donc le processus empirique bi-dimensionnel

$$\begin{aligned} e_n(s,t) &:= \{T_n(s,t) - \mathrm{E}\, T_n(s,t)\} \\ &= \frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{[nt]} \left(1_{\{X_j > u_n s\}} - \bar{F}(u_n s)\right)\ ,\ s \in [1,\infty],\ t \in [0,1]. \end{aligned}$$

Volatilité Stochastique à Mémoire Longue

Volatilité **S**tochastique : série temporelle $(X_j)_{j\geq 1}$ définie par

$$X_j = \sigma\left(Y_j\right)\varepsilon_j, \quad j \geq 1\ ,$$

où

- $\left(\varepsilon_j, \eta_j\right)_{j\in\mathbb{Z}}$ est une suite i.i.d., ε_1 est centrée et $(\eta_i)_{i\in\mathbb{Z}}$ est i.i.d. gaussienne centrée ;
- σ est une fonction positive
- $Y_j, j \geq 1$ est une suite gaussienne à mémoire longue.

Mémoire **L**ongue :

$$Y_j = \sum_{k=1}^{\infty} c_k \eta_{j-k}\ , \quad \sum_{k=1}^{\infty} c_k^2 = 1,$$

où $(\eta_i)_{i\in\mathbb{Z}}$ est i.i.d. gaussienne centrée $\operatorname{Var}\eta_1 = 1$, et

$$\gamma_Y(k) = \operatorname{Cov}\left(Y_j, Y_{j+k}\right) = k^{-D} L_\gamma\left(k\right),$$

où $D \in (0,1)$ et L_γ est à variation lente.

Volatilité Stochastique à Mémoire Longue

Volatilité **S**tochastique : série temporelle $(X_j)_{j\geq 1}$ définie par

$$X_j = \sigma\left(Y_j\right)\varepsilon_j, \quad j \geq 1 \ ,$$

où

- $(\varepsilon_j, \eta_j)_{j\in\mathbb{Z}}$ est une suite i.i.d., ε_1 est centrée et $(\eta_i)_{i\in\mathbb{Z}}$ est i.i.d. gaussienne centrée ;
- σ est une fonction positive
- $Y_j, j \geq 1$ est une suite gaussienne à mémoire longue.

Mémoire **L**ongue :

$$Y_j = \sum_{k=1}^{\infty} c_k \eta_{j-k} \ , \ \sum_{k=1}^{\infty} c_k^2 = 1,$$

où $(\eta_i)_{i\in\mathbb{Z}}$ est i.i.d. gaussienne centrée $\operatorname{Var}\eta_1 = 1$, et

$$\gamma_Y(k) = \operatorname{Cov}\left(Y_j, Y_{j+k}\right) = k^{-D} L_\gamma\left(k\right),$$

où $D \in (0,1)$ et L_γ est à variation lente.

On rappelle que

$$X_j = \sigma\left(Y_j\right)\varepsilon_j, \quad Y_j = \sum_{k=1}^{\infty} c_k \eta_{j-k},\ j \geq 1.$$

On observe que

$$\operatorname{Cov}\left(X_1, X_{k+1}\right) = 0,\ k \geq 1, \quad \left(X_j\right)_{j\geq 1} \text{ décorrélée}$$

$$\operatorname{Cov}\left(\left|X_1\right|, \left|X_{k+1}\right|\right) = \left(\mathrm{E}\left|\varepsilon_1\right|\right)^2 \operatorname{Cov}\left(\left|\sigma(Y_1)\right|, \left|\sigma(Y_{k+1})\right|\right), \text{ covariance héritée de } \left(Y_j\right)_{j\geq 1}$$

On rappelle que

$$X_j = \sigma(Y_j)\,\varepsilon_j, \quad Y_j = \sum_{k=1}^{\infty} c_k \eta_{j-k},\ j \geq 1.$$

On observe que

$$\operatorname{Cov}(X_1, X_{k+1}) = 0,\ k \geq 1, \quad (X_j)_{j\geq 1} \text{ décorrélée}$$

$$\operatorname{Cov}(|X_1|, |X_{k+1}|) = (\mathrm{E}\,|\varepsilon_1|)^2 \operatorname{Cov}(|\sigma(Y_1)|, |\sigma(Y_{k+1})|), \text{ covariance héritée de } (Y_j)_{j\geq 1}$$

Hypothèse: $P(\varepsilon_1 > x) = x^{-\alpha} L(x)$ pour un certain $\alpha > 0$ et une fonction à variation lente L.

Lemme de Breiman. Si $\mathbb{E}\left[\sigma^{\alpha+\delta}(Y_1)\right] < \infty$ pour un certain $\delta > 0$, alors

$$P(X_1 > x) \sim \mathrm{E}\,\sigma^{\alpha}(Y_1)\,P(\varepsilon_1 > x),\ x \to \infty$$

i.e. le comportement de la queue de $(X_j)_{j\geq 1}$ est hérité de celui de $(\varepsilon_j)_{j\geq 1}$.

Résultats récents

- Kulik, Soulier (2011): étude du processus empirique à un paramètre

$$\sum_{j=1}^{n}\left(1_{\{X_j>u_n s\}}-\bar{F}\left(u_n s\right)\right),\ s\in[1,\infty].$$

  Convergence en loi dans $D[1,\infty]$ et loi limite de l'estimateur de Hill.

- Betken, Kulik (2019): étude du processus empirique bi-dimensionnel

$$\sum_{j=1}^{\lfloor nt\rfloor}\left(1_{\{\psi(X_j)\leq x\}}-F_{\psi(X_1)}(x)\right),\ x\in[-\infty,\infty],\ t\in[0,1],$$

  Convergence en loi dans $D\left([-\infty,\infty]\times[0,1]\right)$; test de Wilcoxon.

Résultats récents

- Kulik, Soulier (2011): étude du processus empirique à un paramètre
$$\sum_{j=1}^{n}\left(1_{\{X_j>u_n s\}}-\bar{F}\left(u_n s\right)\right)\ ,\ s\in[1,\infty].$$
Convergence en loi dans $D[1,\infty]$ et loi limite de l'estimateur de Hill.
- Betken, Kulik (2019): étude du processus empirique bi-dimensionnel
$$\sum_{j=1}^{\lfloor nt\rfloor}\left(1_{\{\psi(X_j)\leq x\}}-F_{\psi(X_1)}(x)\right),\ x\in[-\infty,\infty],\ t\in[0,1],$$
Convergence en loi dans $D\left([-\infty,\infty]\times[0,1]\right)$; test de Wilcoxon.
- Betken, G., Kulik (2019+): étude du processus empirique bi-dimensionnel
$$\sum_{j=1}^{\lfloor nt\rfloor}\left(1_{\{X_j>u_n s\}}-\bar{F}\left(u_n s\right)\right)\ ,\ s\in[1,\infty],\ t\in[0,1].$$
Convergence en loi dans $D\left([1,\infty]\times[0,1]\right)$ et loi asymptotique de la statistique de Hill.

Decomposition du processus empirique

Rappelons que

$$X_j = \sigma\left(Y_j\right)\varepsilon_j, \quad Y_j = \sum_{k=1}^{\infty} c_k \eta_{j-k},\ j \geq 1\ .$$

Définissons la tribu

$$\mathcal{F}_j = \sigma\left(\varepsilon_k, \eta_k, k \leq j\right).$$

Decomposition du processus empirique

Rappelons que

$$X_j = \sigma\left(Y_j\right)\varepsilon_j, \quad Y_j = \sum_{k=1}^{\infty} c_k \eta_{j-k},\ j \geq 1\ .$$

Définissons la tribu

$$\mathcal{F}_j = \sigma\left(\varepsilon_k, \eta_k, k \leq j\right).$$

On considère la décomposition suivante:

$$e_n(s,t) = M_n\left(s,t\right) + L_n\left(s,t\right),$$

où M_n est le terme en martingale et L_n à mémoire longue :

$$M_n\left(s,t\right) := \frac{1}{n\bar{F}_Z\left(u_n\right)} \sum_{j=1}^{\lfloor nt \rfloor} \left(1_{\left\{X_j > u_n s\right\}} - \mathrm{E}\left(1_{\left\{X_j > u_n s\right\}} \mid \mathcal{F}_{j-1}\right)\right),$$

$$L_n\left(s,t\right) := \frac{1}{n\bar{F}_Z\left(u_n\right)} \sum_{j=1}^{\lfloor nt \rfloor} \left(\mathrm{E}\left(1_{\left\{X_j > u_n s\right\}} \mid \mathcal{F}_{j-1}\right) - \bar{F}(u_n s)\right)$$

Soit φ la densité d'une loi normale centrée réduite.

Théorème

Pour $G \in L^2(\mathbb{R}, \varphi(x)dx)$ et $X \sim \mathcal{N}(0,1)$, la convergence suivante a lieu dans $L^2(\mathbb{R}, \varphi(x)dx)$

$$G(X) = \sum_{r=0}^{\infty} \frac{J_r(G)}{r!} H_r(X),$$

où

$$J_r(G) := \mathrm{E}\, G(X) H_r(X)$$

et H_m est le n-ième polynôme d'Hermite :

$$H_n(x) := (-1)^n \mathrm{e}^{\frac{1}{2}x^2} \frac{d^n}{dx^n} \mathrm{e}^{-\frac{1}{2}x^2}, \ x \in \mathbb{R}.$$

On définit le *rang d'Hermite* de G par.

$$r := \min \{k \geq 1 : J_k(G) \neq 0\}.$$

Théorème limite pour des fonctions de processus gaussien

Théorème (Taqqu (1979))

Soit $(Y_j)_{j\geq 1}$ *une suite strictement stationnaire gaussienne telle que* $\operatorname{Cov}(Y_1, Y_{k+1}) \sim k^{-D}L(k)$, $0 < D < 1$ *et soit* $G \in L^2(\mathbb{R}, \varphi(x)dx)$ *une fonction de rang d'Hermite* r. *Alors*

$$\operatorname{Cov}(G(Y_1), G(Y_{k+1})) \sim J_r^2(G)r!(\gamma(k))^r = k^{-Dr}J_r^2(G)r!L^r(k)$$

et

$$\frac{1}{d_{n,r}}\sum_{j=1}^{\lfloor nt\rfloor} G(Y_j) \overset{D[0,1]}{\Rightarrow} \frac{J_r(G)}{r!}Z_{r,H}(t),\ 0 \leq t \leq 1,$$

où $Z_{r,H}(t)$, $0 \leq t \leq 1$, *est un processus d'Hermite d'ordre* r, *de paramètre* $H = 1 - \frac{D}{2}$ *et*

$$d_{n,r}^2 = \operatorname{Var}\left(\sum_{j=1}^{n} H_r(Y_j)\right) \sim c_r n^{2-rD}L^r(n),\ c_r = \frac{2r!}{(1-Dr)(2-Dr)}.$$

Convergence du terme en “mémoire longue”

On rappelle que $X_j = \sigma\left(Y_j\right)\varepsilon_j$, $Y_j = \sum_{k=1}^{\infty} c_k \eta_{j-k}$, $j \geq 1$, $\mathcal{F}_j = \sigma\left(\varepsilon_j, \varepsilon_{j-1}, \ldots, \eta_j, \eta_{j-1}, \ldots\right)$, et $e_n(s,t) = M_n(s,t) + L_n(s,t)$, où

$$M_n\left(s,t\right) := \frac{1}{n\bar{F}_Z(u_n)} \sum_{j=1}^{\lfloor nt \rfloor} \left(1_{\{X_j > u_n s\}} - \mathrm{E}\left(1_{\{X_j > u_n s\}} \mid \mathcal{F}_{j-1}\right)\right),$$

$$L_n\left(s,t\right) := \frac{1}{n\bar{F}_Z(u_n)} \sum_{j=1}^{\lfloor nt \rfloor} \left(\mathrm{E}\left(1_{\{X_j > u_n s\}} \mid \mathcal{F}_{j-1}\right) - \bar{F}(u_n s)\right) .$$

Convergence du terme en “mémoire longue”

On rappelle que $X_j = \sigma\left(Y_j\right)\varepsilon_j$, $\;Y_j = \sum_{k=1}^{\infty} c_k \eta_{j-k}$, $j \geq 1$, $\mathcal{F}_j = \sigma\left(\varepsilon_j, \varepsilon_{j-1}, \dots, \eta_j, \eta_{j-1}, \dots\right)$, et $e_n(s,t) = M_n(s,t) + L_n(s,t)$, où

$$M_n\left(s,t\right) := \frac{1}{n\bar{F}_Z(u_n)} \sum_{j=1}^{\lfloor nt \rfloor} \left(1_{\{X_j > u_n s\}} - \mathrm{E}\left(1_{\{X_j > u_n s\}} \mid \mathcal{F}_{j-1}\right)\right),$$

$$L_n\left(s,t\right) := \frac{1}{n\bar{F}_Z(u_n)} \sum_{j=1}^{\lfloor nt \rfloor} \left(\mathrm{E}\left(1_{\{X_j > u_n s\}} \mid \mathcal{F}_{j-1}\right) - \bar{F}(u_n s)\right) .$$

Remarquons que

$$\begin{aligned} L_n\left(s,t\right) &= \frac{1}{n} \sum_{j=1}^{\lfloor nt \rfloor} \left(\Psi_s(Y_j) - \mathrm{E}\,\Psi_s(Y_j)\right) \\ &+ \frac{1}{n} \sum_{j=1}^{\lfloor nt \rfloor} \left(\Psi_n\left(Y_j, s\right) - \Psi_s\left(Y_j\right)\right) + \frac{1}{n} \sum_{j=1}^{\lfloor nt \rfloor} \left(\mathrm{E}\,\Psi_s\left(Y_j\right) - \mathrm{E}\left(\Psi_n\left(Y_j, s\right)\right)\right) \end{aligned}$$

avec

$$\Psi_n\left(y,s\right) = \frac{\bar{F}_Z\left(\frac{u_n s}{\sigma(y)}\right)}{\bar{F}_Z(u_n)},\; \Psi_s(y) = s^{-\alpha}\Psi(y),\; \Psi(y) = \sigma^{\alpha}(y) .$$

Convergence du terme en “mémoire longue”

Idée globale:

$$\frac{n}{d_{n,r}} L_n(s,t) = \frac{1}{d_{n,r}} \sum_{j=1}^{\lfloor nt \rfloor} \left(\Psi_s(Y_j) - \operatorname{E} \Psi_s(Y_j) \right) + o_P(1) ,$$

où

$$\Psi_s(y) = s^{-\alpha} \psi(y) \text{ et } \psi(y) = \sigma^{\alpha}(y) .$$

Il nous faut une hypothèse sur $\bar{F}_Z$.

Variation régulière du second ordre

Definition

Soit $\bar{F}(x) = x^{-\alpha} L(x)$ pour une fonction à variation lente L représentée par

$$L(x) = c \exp\left(\int_1^x \frac{\eta(u)}{u} du\right)$$

pour une certaine constante c et une fonction mesurable η. On suppose qu'il existe une fonction décroissante η^* définie sur $[0, \infty)$, telle que $\eta^*(x) = x^{-\rho} L_{\eta^*}(x)$ pour un $\rho \geq 0$ et

$$|\eta(s)| \leq C\eta^*(s),$$

pour une constante C et tout $s \geq 0$. On dit que $\bar{F}$ est *à variation régulière du second ordre* d'indice α et de fonction η^* et on écrit $\bar{F} \in 2\mathrm{RV}(\alpha, \eta^*)$.

Variation régulière du second ordre

Definition

Soit $\bar{F}(x) = x^{-\alpha}L(x)$ pour une fonction à variation lente L représentée par

$$L(x) = c \exp\left(\int_1^x \frac{\eta(u)}{u} du\right)$$

pour une certaine constante c et une fonction mesurable η. On suppose qu'il existe une fonction décroissante η^* définie sur $[0, \infty)$, telle que $\eta^*(x) = x^{-\rho}L_{\eta^*}(x)$ pour un $\rho \geq 0$ et

$$|\eta(s)| \leq C\eta^*(s),$$

pour une constante C et tout $s \geq 0$. On dit que $\bar{F}$ est *à variation régulière du second ordre* d'indice α et de fonction η^* et on écrit $\bar{F} \in 2\mathrm{RV}(\alpha, \eta^*)$.

Lemme (Kulik, Soulier (2011))

Soit $\bar{F}_Z \in 2\mathrm{RV}(\alpha, \eta^*)$. Pour tout $\varepsilon > 0$, il existe une constante C telle que

$$\forall t \geq 1, \forall s > 0, \quad \left|\frac{\bar{F}_Z(st)}{\bar{F}_Z(t)} - s^{-\alpha}\right| \leq C\eta^*(t)\, s^{-\alpha-\rho}\left(s \vee s^{-1}\right)^{\varepsilon}.$$

Hypothèses techniques :

- $\bar{F}_Z \in 2\mathrm{RV}\left(\alpha, \eta^*\right)$;
- $\eta^*(u_n) = o\left(\frac{d_{n,r}}{n}\right)$;
- $\mathrm{E}\left[\sigma(Y_1)^{\alpha+\max\{\rho,\alpha\}+\varepsilon}\right] < \infty$ pour un certain $\varepsilon > 0$.

Hypothèses techniques :

- $\bar{F}_Z \in 2\mathrm{RV}\,(\alpha, \eta^*)$;
- $\eta^*(u_n) = o\left(\frac{d_{n,r}}{n}\right)$;
- $\mathrm{E}\left[\sigma(Y_1)^{\alpha+\max\{\rho,\alpha\}+\varepsilon}\right] < \infty$ pour un certain $\varepsilon > 0$.

Théorème (Betken, G., Kulik (2019+))

Sous les hypothèses précedentes et le modèle VSML, si $\frac{n}{d_{n,r}} = o\left(\sqrt{n\bar{F}(u_n)}\right)$ *alors*

$$\frac{n}{d_{n,r}} \frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{[nt]} \left(1_{\{X_j > u_n s\}} - \bar{F}(u_n s)\right) \Rightarrow \frac{s^{-\alpha}}{\mathrm{E}\left(\sigma^\alpha(Y_1)\right)} \frac{J_r(\Psi)}{r!} Z_{r,H}(t),$$

où $\Rightarrow$ *désigne la convergence en loi dans* $D\left([1,\infty] \times [0,1]\right)$, $\Psi(y) = \sigma^\alpha(y)$, r *est le rang d'Hermite de* Ψ *et* $Z_{r,H}$ *est un processus d'Hermite avec* $H = 1 - \frac{D}{2}$.

Obstacle à une application directe d'une fonctionnelle continue

On rappelle que

$$\hat{\gamma}_{1,\lfloor nt \rfloor} = \frac{\frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{[nt]} \log\left(\frac{X_j}{u_n}\right) 1_{\{X_j > u_n\}}}{\frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{\lfloor nt \rfloor} 1_{\{X_j > u_n\}}} = (T_n(1,t))^{-1} \int_1^\infty s^{-1} T_n(s,t) ds \,,$$

où $T_n(s,t) = \frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{[nt]} 1_{\{X_j > u_n s\}}$, $s \in [1,\infty]$, $t \in [0,1]$.

Obstacle à une application directe d'une fonctionnelle continue

On rappelle que

$$\hat{\gamma}_{1,\lfloor nt \rfloor} = \frac{\frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{[nt]} \log\left(\frac{X_j}{u_n}\right) 1_{\{X_j > u_n\}}}{\frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{\lfloor nt \rfloor} 1_{\{X_j > u_n\}}} = (T_n(1,t))^{-1} \int_1^{\infty} s^{-1} T_n(s,t) ds \ ,$$

où $T_n(s,t) = \frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{[nt]} 1_{\{X_j > u_n s\}}$, $s \in [1,\infty]$, $t \in [0,1]$. On voudrait utiliser la fonctionnelle

$$H\colon D([1,\infty] \times [0,1]) \longrightarrow D[0,1],\ f \mapsto (f(1,t))^{-1} \int_1^{\infty} s^{-1} f(s,t)\, ds.$$

Obstacle à une application directe d'une fonctionnelle continue

On rappelle que

$$\hat{\gamma}_{1,\lfloor nt\rfloor} = \frac{\frac{1}{n\bar{F}(u_n)}\sum_{j=1}^{[nt]}\log\left(\frac{X_j}{u_n}\right)1_{\{X_j>u_n\}}}{\frac{1}{n\bar{F}(u_n)}\sum_{j=1}^{\lfloor nt\rfloor}1_{\{X_j>u_n\}}} = (T_n(1,t))^{-1}\int_1^{\infty}s^{-1}T_n(s,t)ds\ ,$$

où $T_n(s,t) = \frac{1}{n\bar{F}(u_n)}\sum_{j=1}^{[nt]}1_{\{X_j>u_ns\}}$, $s\in[1,\infty]$, $t\in[0,1]$. On voudrait utiliser la fonctionnelle

$$H\colon D([1,\infty]\times[0,1])\longrightarrow D[0,1],\ f\mapsto (f(1,t))^{-1}\int_1^{\infty}s^{-1}f(s,t)\,ds.$$

On utilise le

Lemme

Pour tout $\varepsilon>0$, la convergence suivante a lieu:

$$\lim_{R\to+\infty}\limsup_{n\to+\infty}\mathbb{P}\left\{\sup_{t\in[0,1]}\left|\int_R^{+\infty}s^{-1}\frac{n}{d_{n,r}}\left\{T_n(s,t)-\mathrm{E}\,T_n(s,t)\right\}\mathrm{d}s\right|>\varepsilon\right\}=0.$$

Autres hypothèses techniques

Nuos avons utilisé les hypothèses :

- $\bar{F}_Z \in 2\mathrm{RV}\left(\alpha, \eta^*\right)$;
- $\eta^*(u_n) = o\left(\frac{d_{n,r}}{n}\right)$;
- $\mathrm{E}\left[\sigma(Y_1)^{\alpha+\max\{\rho,\alpha\}+\varepsilon}\right] < \infty$ pour un certain $\varepsilon > 0$.

Afin de garantir la convergence du terme en martingale, nous aurons besoin des hypothèses :

- $\mathbb{E}\left[\sigma\left(Y_1\right)^{2(\alpha-\rho)}\left(\sigma\left(Y_0\right) \vee \left(\sigma\left(Y_1\right)\right)^{-1}\right)^{\varepsilon_0}\right] < \infty$ pour un certain $\varepsilon_0 > 0$;
- $\mathbb{E}\left[\left(\sigma\left(Y_1\right)\right)^{-2-\delta}\right] < +\infty$ pour un certain $\delta > 0$;
- $\eta^*(u_n) = o\left(\frac{d_n}{n} + \frac{1}{\sqrt{n\bar{F}(u_n)}}\right)$.

Convergence du “terme martingale”

On rappelle que

$$M_n(s,t) = \frac{1}{n\bar{F}_Z(u_n)} \sum_{j=1}^{\lfloor nt \rfloor} \left(1_{\{X_j > u_n s\}} - \mathrm{E}\left(1_{\{X_j > u_n s\}} \mid \mathcal{F}_{j-1} \right) \right).$$

Théorème (Betken, G., Kulik, 2019+)

Sous les hypothèses principales, la convergence en loi suivante a lieu dans $D([1,R] \times [0,1])$ *pour tout* $R > 1$ *:*

$$\left(\sqrt{n\bar{F}_Z(u_n)} M_n(s,t) \right)_{1 \le s \le R, 0 \le t \le 1} \to \sqrt{\mathbb{E}\left[\sigma\left(Y_1\right)^{\alpha}\right]} \left(B_{s^{-\alpha},t}\right)_{1 \le s \le R, 0 \le t \le 1},$$

où $B_{u,v}$ *est un drap brownien standard.*

Convergence de l'estimateur de Hill

Pour établir la convergence de l'estimateur de Hill, on utilise la fonctionnelle continue $x \mapsto \int_1^R s^{-1} x(s,t)\,\mathrm{d}s$ et on montre que la contribution de $\int_R^{+\infty}$ est négligeable.

Lemme

Pour tout ε, la convergence suivante a lieu:

$$\lim_{R\to+\infty} \limsup_{n\to+\infty} \mathbb{P}\left\{ \sup_{t\in[0,1]} \left| \int_R^{+\infty} s^{-1} \sqrt{n\bar{F}_Z(u_n)} M_n(s,t)\,\mathrm{d}s \right| > \varepsilon \right\} = 0.$$

Idée de démonstration

Idée globale.
On utilise un critère de Davydov et Zitikis (2008) garantissant la convergence en loi d'une suite de processus $(\xi_n(s,t))_{n\geq 1}$ indexés par $[0,1]^2$.

1. Convergence des lois fini-dimensionnelles ;
2. Contrôle des moments des différences $|\xi_n(s_1,t_1) - \xi_n(s_2,t_2)|$ où $\|(s_1,t_1) - (s_2,t_2)\|_\infty \geq a_n$ et $a_n \to 0$.
3. Convergence en probabilité d'un maximum mettant en jeu les différences $|\xi_n(s_1,t_1) - \xi_n(s_2,t_2)|$ lorsque $\|(s_1,t_1) - (s_2,t_2)\|_\infty \leq a_n$.

Idée de démonstration (2)

Soient ξ_n, $n \geq 1$ des processus stochastiques indexés par $[0,1]^2$, à trajectoires dans $D([0,1] \times [0,1])$. Supposons que

1. le processus ξ_n peut s'exprimer comme la différence de deux processus croissants ξ_n° et ξ_n^*;
2. les lois fini-dimensionnelles de ξ_n convergence vers celles d'un processus ξ à trajectoire continue.

Idée de démonstration (3)

S'il existe une suite $a_n \downarrow 0$ vérifiant:

1. Condition sur les moments: il existe des constantes $\gamma \geq \beta > 2$, $c > 0$ telles que pour tout $n \geq 1$, $n \geq 1$, $\mathrm{E} \left| \xi_n (0) \right|^{\gamma} \leq c$ et
$$\mathrm{E} \left| \xi_n (s_1, t_1) - \xi_n (s_2, t_2) \right|^{\gamma} \leq c \left\| (s_1, t_1) - (s_2, t_2) \right\|_{\infty}^{\beta}$$
lorsque $\left\| (s_1, t_1) - (s_2, t_2) \right\|_{\infty} \geq a_n$.

2. La convergence (en probabilité) suivante a lieu:
$$\max_{0 \leq j_1, j_2 \leq k_n - 1} \left| \xi_n^* \left(\frac{j_1 + 1}{k_n}, \frac{j_2}{k_n} \right) - \xi_n^* \left(\frac{j_1}{k_n}, \frac{j_2}{k_n} \right) \right| \to 0,$$
$$\max_{0 \leq j_1, j_2 \leq k_n - 1} \left| \xi_n^* \left(\frac{j_1}{k_n}, \frac{j_2 + 1}{k_n} \right) - \xi_n^* \left(\frac{j_1}{k_n}, \frac{j_2}{k_n} \right) \right| \to 0,$$
où $k_n := [1/a_n]$,

alors $\xi_n \to \xi$ dans $D\left([0,1]^2\right)$.

Idée de démonstration (4)

On utilise le résultat précédent dans le contexte suivant:

$$\xi_n(s,t) := \frac{1}{\sqrt{n\bar{F}_Z(u_n)}} \sum_{j=1}^{[nt]} \left(\mathbf{1}\left\{X_j \leq u_n(1+Rs)\right\} \right. \\ \left. - \mathbb{E}\left[\mathbf{1}\left\{X_j \leq u_n(1+R)\right\} \mid \mathcal{F}_{j-1}\right] \right),$$

qui peut s'exprimer comme la différence de deux processus croissants ξ_n° et ξ_n^* :

$$\xi_n^\circ(s,t) := \frac{1}{\sqrt{n\bar{F}_Z(u_n)}} \sum_{j=1}^{[nt]} \mathbf{1}\left\{X_j \leq u_n(1+Rs)\right\} \text{ et}$$

$$\xi_n^*(s,t) := \frac{1}{\sqrt{n\bar{F}_Z(u_n)}} \sum_{j=1}^{[nt]} \mathbb{E}\left[\mathbf{1}\left\{X_j \leq u_n(1+Rs)\right\} \mid \mathcal{F}_{j-1}\right].$$

Idée de démonstration (5)

- Lois fini-dimensionnelles : par le théorème de Cramer-Wold, il suffit de traiter la convergence des combinaisons linéaires. On utilise un théorème limite central pour des tableaux d'accroissement de martingale.

- Condition sur les moments : inégalité de Burkholder et contrôle des variances conditionnelles à l'aide de l'hypothèse de variation régulière du second ordre de $\bar{F}_Z$.

- Contrôle des accroissements : ces-derniers ds'expriment à l'aide de $\bar{F}_Z$, pour laquelle on utilise encore hypothèse de variation régulière du second ordre.

Théorème (Betken, G., Kulik (2019+))

Sous les hypothèses concernant $\bar{F}_Z$ et $\sigma(Y_1)$:

- *Si $n^{\frac{D}{2}} = o\left(\sqrt{n\bar{F}(u_n)}\right)$, alors*

$$\frac{1}{n^{1-\frac{D}{2}}}\sum_{j=1}^{[nt]}\left(\frac{1}{\bar{F}(u_n)}1_{\{X_j>u_n s\}} - s^{-\alpha}\right) \Rightarrow K_{r,H}\frac{s^{-\alpha}}{\mathrm{E}\left(\sigma^\alpha(Y_1)\right)}Z_{r,H}(t),$$

  où $Z_{r,H}$ est un processus d'Hermite $H = 1-\frac{D}{2}$, r l'indide d'Hermite de σ^α et $K_{r,H}$ une constante.

- *Si $\sqrt{n\bar{F}(u_n)} = o\left(n^{\frac{D}{2}}\right)$, alors*

$$\frac{\sqrt{\bar{F}(u_n)}}{\sqrt{n}}\sum_{j=1}^{[nt]}\left(\frac{1}{\bar{F}(u_n)}1_{\{X_j>u_n s\}} - s^{-\alpha}\right) \Rightarrow \sqrt{\mathrm{E}\,\sigma^\alpha(Y_1)}B_{s^{-\alpha},t},$$

  où B est un drap brownien standard.

($\Rightarrow$ désigne la convergence en loi dans $D([1,\infty]\times[0,1])$.)

Soit

$$\hat{\gamma}(t) = \frac{\sum_{j=1}^{[nt]} \log\left(\frac{X_j}{u_n}\right) 1_{\{X_j > u_n\}}}{\sum_{j=1}^{\lfloor nt \rfloor} 1_{\{X_j > u_n\}}}, \quad \gamma = \alpha^{-1}$$

Corollaire

Sous les hypothèses du résultat principal

$$a_n\left(\hat{\gamma}(t) - \gamma\right) \Rightarrow \frac{1}{t}\int_1^\infty s^{-1} W(s,t)ds - \alpha^{-1}\frac{1}{t}W(1,t)$$

où $\Rightarrow$ *désigne la convergence en loi dans* $D[0,1]$*,*

$$a_n = \begin{cases} \frac{n}{d_n}, & si\ \frac{n}{d_n} = o\left(\sqrt{n\bar{F}(u_n)}\right) \\ \sqrt{n\bar{F}(u_n)}, & si\ \sqrt{n\bar{F}(u_n)} = o\left(\frac{n}{d_n}\right) \end{cases}$$

et

$$W(s,t) = \begin{cases} \frac{s^{-\alpha}}{\mathrm{E}(\sigma^\alpha(Y_1))}\frac{J_q(\Psi)}{q!}Z_q(t), & si\ \frac{n}{d_n} = o\left(\sqrt{n\bar{F}(u_n)}\right) \\ \sqrt{\mathbb{E}\left[\sigma\left(Y_1\right)^\alpha\right]}B_{s^{-\alpha},t}, & si\ \sqrt{n\bar{F}(u_n)} = o\left(\frac{n}{d_n}\right) \end{cases}.$$

Perspectives

- Démonstration alternative de la tension du terme en martingale avec des arguments de chaînage : le supremum est approximé par le maximum sur des partitions.
- Simulations.