

Cours de Formation par la Recherche, ENSAE-CREST.

Leçon 2 : Tests de détection de longue portée et estimateurs du paramètre de longue portée

Gilles Teyssière
stats@gillesteysiere.net

Mars–Avril 2007

Test paramétrique

- Pour cette classe de tests, la distribution limite sous l'hypothèse nulle est une distribution standard.

Test paramétrique

- Pour cette classe de tests, la distribution limite sous l'hypothèse nulle est une distribution standard.
- Robinson (1994) considère des processus du type suivant :

$$\phi(L)X_t = (1 - L)^d X_t = \xi_t,$$

Test paramétrique

- Pour cette classe de tests, la distribution limite sous l'hypothèse nulle est une distribution standard.
- Robinson (1994) considère des processus du type suivant :

$$\phi(L)X_t = (1 - L)^d X_t = \xi_t,$$

- ξ_t étant une processus stationnaire du second ordre d'espérance nulle et avec une densité spectrale paramétrique $f(\lambda) = (\sigma^2/2\pi)g(\lambda; \tau)$ dépendant de paramètres inconnus $\tau \in \mathbb{R}^k$ et σ^2 .

Test paramétrique

- Pour cette classe de tests, la distribution limite sous l'hypothèse nulle est une distribution standard.
- Robinson (1994) considère des processus du type suivant :

$$\phi(L)X_t = (1 - L)^d X_t = \xi_t,$$

- ξ_t étant un processus stationnaire du second ordre d'espérance nulle et avec une densité spectrale paramétrique $f(\lambda) = (\sigma^2/2\pi)g(\lambda; \tau)$ dépendant de paramètres inconnus $\tau \in \mathbf{R}^k$ et σ^2 .
- La densité spectrale de $g(\lambda)$ doit être spécifiée : deux processus sont considérés:

Test paramétrique

- Pour cette classe de tests, la distribution limite sous l'hypothèse nulle est une distribution standard.
- Robinson (1994) considère des processus du type suivant :

$$\phi(L)X_t = (1 - L)^d X_t = \xi_t,$$

- ξ_t étant un processus stationnaire du second ordre d'espérance nulle et avec une densité spectrale paramétrique $f(\lambda) = (\sigma^2/2\pi)g(\lambda; \tau)$ dépendant de paramètres inconnus $\tau \in \mathbf{R}^k$ et σ^2 .
- La densité spectrale de $g(\lambda)$ doit être spécifiée : deux processus sont considérés:
 - Processus AR(k), dont le spectre est : $g(\lambda; \tau) = \left| 1 - \sum_{j=1}^k \tau_j e^{ij\lambda} \right|^{-2}$,

Test paramétrique

- Pour cette classe de tests, la distribution limite sous l'hypothèse nulle est une distribution standard.
- Robinson (1994) considère des processus du type suivant :

$$\phi(L)X_t = (1 - L)^d X_t = \xi_t,$$

- ξ_t étant un processus stationnaire du second ordre d'espérance nulle et avec une densité spectrale paramétrique $f(\lambda) = (\sigma^2/2\pi)g(\lambda; \tau)$ dépendant de paramètres inconnus $\tau \in \mathbf{R}^k$ et σ^2 .
- La densité spectrale de $g(\lambda)$ doit être spécifiée : deux processus sont considérés :
 - Processus AR(k), dont le spectre est : $g(\lambda; \tau) = \left| 1 - \sum_{j=1}^k \tau_j e^{ij\lambda} \right|^{-2}$,
 - Processus de Bloomfield, dont le spectre est : $g(\lambda; \tau) = \exp\{2 \sum_{j=1}^k \tau_j \cos(j\lambda)\}$.
 (Processus toujours stationnaire)

Test du score pour $d = d_0$ contre des alternatives fractionnaires

- La statistique du score \hat{r} est définie comme suit :

$$\hat{r} = \frac{T^{1/2}}{\hat{\sigma}^2} \hat{A}^{-1/2} \hat{a}. \quad \text{Under } H_0 : d = d_0, \hat{r} \rightarrow_D \mathcal{N}(0, 1),$$

Test du score pour $d = d_0$ contre des alternatives fractionnaires

- La statistique du score \hat{r} est définie comme suit :

$$\hat{r} = \frac{T^{1/2}}{\hat{\sigma}^2} \hat{A}^{-1/2} \hat{a}. \quad \text{Under } H_0 : d = d_0, \hat{r} \rightarrow_D \mathcal{N}(0, 1),$$

- avec

$$\varphi(\lambda) = \operatorname{Re} \left\{ \log \left(\phi(e^{i\lambda}) \right)'_{d=d_0} \right\} = \log |2 \sin(\lambda/2)|, \quad \lambda \in [-\pi, \pi).$$

$$\tilde{\xi}_t = (1 - L)^{d_0} X_t, \quad I_{\tilde{\xi}}(\lambda) = (1/2\pi N) \left| \sum_{t=1}^N \tilde{\xi}_t e^{it\lambda} \right|^2, \quad \lambda_j = 2\pi j/N,$$

$$\hat{\zeta}(\lambda) = (\partial/\partial\tau) \log g(\lambda; \hat{\tau}),$$

$$\sigma^2(\tau) = \frac{2\pi}{T} \sum_j \frac{I_{\tilde{\xi}}(\lambda_j)}{g(\lambda_j; \tau)}, \quad \hat{\sigma}^2 = \sigma^2(\hat{\tau}), \quad \hat{a} = -\frac{2\pi}{T} \sum_j' \varphi(\lambda_j) \frac{I_{\tilde{\xi}}(\lambda_j)}{g(\lambda_j; \hat{\tau})},$$

$$\hat{A} = \frac{2}{T} \left(\sum_j' |\varphi(\lambda_j)|^2 - \sum_j' \varphi(\lambda_j) \hat{\zeta}(\lambda_j)' \left\{ \sum_j' \hat{\zeta}(\lambda_j) \hat{\zeta}(\lambda_j)' \right\}^{-1} \sum_j' \hat{\zeta}(\lambda_j) \varphi(\lambda_j) \right),$$

Limites des tests paramétriques

- Ces tests sont très puissants, mais très “étroits”, c-à-d, que leur domaine de validité est limité strictement au cadre paramétrique considéré, et sont très sensibles à des écarts par rapport à ce cadre,

Limites des tests paramétriques

- Ces tests sont très puissants, mais très “étroits”, c-à-d, que leur domaine de validité est limité strictement au cadre paramétrique considéré, et sont très sensibles à des écarts par rapport à ce cadre,
- Autre approche, tests semiparamétriques, basés sur le processus des sommes partielles.

Tests semi-paramétriques

Tests basés sur le processus centré des sommes partielles

La dépendance faible est définie par :

Hypothèses

- ① $\sum_{j=-\infty}^{\infty} |\gamma_j| < \infty$, $\gamma_k = \text{Cov}(X_0, X_k)$,
- ② $\frac{1}{\sqrt{T}} \sum_{j=1}^{[Tt]} (X_j - \bar{X}) \xrightarrow{D[0,1]} \sigma W(t)$ as $T \rightarrow \infty$, $\sigma^2 = \sum_{j=-\infty}^{\infty} \gamma_j$,
- ③ Les cumulants du quatrième ordre $\kappa(h, r, s)$ vérifient

$$\sup_h \sum_{r,s=-\infty}^{\infty} |\kappa(h, r, s)| < \infty,$$

$$\begin{aligned} \kappa(h, r, s) = E & \left((X_k - \mu)(X_{k+h} - \mu)(X_{k+r} - \mu)(X_{k+s} - \mu) \right) \\ & - (\gamma_h \gamma_{r-s} + \gamma_r \gamma_{h-s} + \gamma_s \gamma_{h-r}). \end{aligned}$$

Tests semi-paramétriques

L'hypothèse alternative de dépendance forte est définie par :

Hypothèses

① $\gamma_k \sim ck^{2d-1}$

② $\frac{1}{T^{1/2+d}} \sum_{j=1}^{[Tt]} (X_j - \bar{X}) \xrightarrow{D[0,1]} \sigma W_{1/2+d}(t)$ quand $T \rightarrow \infty$,

③ Les cumulants $\kappa(h, r, s)$ vérifient $\sup_h \sum_{r,s=-\infty}^T |\kappa(h, r, s)| = O(T^{2d})$,

où $\{W_H(t), 0 \leq t \leq 1\}$ est le processus de Wiener fractionnaire avec coefficient d'Hurst $H = d + 1/2$, c-à-d le processus Gaussien de moyenne zéro et dont les covariances sont

$$EW_H(t_1)W_H(t_2) = \frac{1}{2}(t_1^{2H} + t_2^{2H} - |t_1 + t_2|^{2H}).$$

Tests semi-paramétriques

Cette classe de tests est basée sur le processus des sommes partielles (centrées)

$$S_k^* = \sum_{1 \leq i \leq k} (X_i - \bar{X})$$

et de son comportement asymptotique sous les hypothèses de courte portée et de longue portée :

Tests semi-paramétriques

Cette classe de tests est basée sur le processus des sommes partielles (centrées)

$$S_k^* = \sum_{1 \leq i \leq k} (X_i - \bar{X})$$

et de son comportement asymptotique sous les hypothèses de courte portée et de longue portée :

- 1 L'amplitude standardisée de S_k^* (Mandelbrot et Taqqu, 1979)

$$Q_T(q) = \frac{1}{\hat{\sigma}_T(q)} \left[\max_{1 \leq k \leq T} S_k^* - \min_{1 \leq k \leq T} S_k^* \right] = \frac{\hat{R}_T}{\hat{\sigma}_T(q)}.$$

Tests semi-paramétriques

Cette classe de tests est basée sur le processus des sommes partielles (centrées)

$$S_k^* = \sum_{1 \leq i \leq k} (X_i - \bar{X})$$

et de son comportement asymptotique sous les hypothèses de courte portée et de longue portée :

- 1 L'amplitude standardisée de S_k^* (Mandelbrot et Taquq, 1979)

$$Q_T(q) = \frac{1}{\hat{s}_T(q)} \left[\max_{1 \leq k \leq T} S_k^* - \min_{1 \leq k \leq T} S_k^* \right] = \frac{\hat{R}_T}{\hat{s}_T(q)}.$$

- 2 La variance de S_k^* : (Giraitis *et al.*, 2003a)

$$M_T(q) = \frac{1}{T^2 \hat{s}_T^2(q)} \left[\sum_{k=1}^T (S_k^*)^2 - \frac{1}{T} \left(\sum_{k=1}^T S_k^* \right)^2 \right] = \frac{\hat{V}_T}{T \hat{s}_T^2(q)},$$

Tests semi-paramétriques

Cette classe de tests est basée sur le processus des sommes partielles (centrées)

$$S_k^* = \sum_{1 \leq i \leq k} (X_i - \bar{X})$$

et de son comportement asymptotique sous les hypothèses de courte portée et de longue portée :

- ❶ L'amplitude standardisée de S_k^* (Mandelbrot et Taquq, 1979)

$$Q_T(q) = \frac{1}{\hat{s}_T(q)} \left[\max_{1 \leq k \leq T} S_k^* - \min_{1 \leq k \leq T} S_k^* \right] = \frac{\hat{R}_T}{\hat{s}_T(q)}.$$

- ❷ La variance de S_k^* : (Giraitis *et al.*, 2003a)

$$M_T(q) = \frac{1}{T^2 \hat{s}_T^2(q)} \left[\sum_{k=1}^T (S_k^*)^2 - \frac{1}{T} \left(\sum_{k=1}^T S_k^* \right)^2 \right] = \frac{\hat{V}_T}{T \hat{s}_T^2(q)},$$

Remarque

Le test KPSS est basé sur le second moment du processus des sommes partielles centré. Comme le test V/S est plus puissant, ce test n'est pas présenté.

Estimateur consistant de la variance

- $\hat{\sigma}_T^2(q)$ est un estimateur consistant de la variance du processus :

$$\hat{\sigma}_T^2(q) = T^{-1} \sum_{i=1}^T (X_i - \bar{X})^2 + 2 \sum_{i=1}^q \omega_i(q) \hat{\gamma}_i \quad \text{with} \quad \omega_i(q) = 1 - \frac{1}{q+1},$$

Estimateur consistant de la variance

- $\hat{\sigma}_T^2(q)$ est un estimateur consistant de la variance du processus :

$$\hat{\sigma}_T^2(q) = T^{-1} \sum_{i=1}^T (X_i - \bar{X})^2 + 2 \sum_{i=1}^q \omega_i(q) \hat{\gamma}_i \quad \text{with} \quad \omega_i(q) = 1 - \frac{1}{q+1},$$

- Cet estimateur prend en compte les autocovariances $\hat{\gamma}_i$ du processus jusqu'à l'ordre q .

Estimateur consistant de la variance

- $\hat{\sigma}_T^2(q)$ est un estimateur consistant de la variance du processus :

$$\hat{\sigma}_T^2(q) = T^{-1} \sum_{i=1}^T (X_i - \bar{X})^2 + 2 \sum_{i=1}^q \omega_i(q) \hat{\gamma}_i \quad \text{with} \quad \omega_i(q) = 1 - \frac{1}{q+1},$$

- Cet estimateur prend en compte les autocovariances $\hat{\gamma}_i$ du processus jusqu'à l'ordre q .
- La fenêtre q est telle que

$$\frac{1}{T} + \frac{q}{T} \longrightarrow 0 \quad T \rightarrow \infty$$

Estimateur consistant de la variance

- $\hat{\sigma}_T^2(q)$ est un estimateur consistant de la variance du processus :

$$\hat{\sigma}_T^2(q) = T^{-1} \sum_{i=1}^T (X_i - \bar{X})^2 + 2 \sum_{i=1}^q \omega_i(q) \hat{\gamma}_i \quad \text{with} \quad \omega_i(q) = 1 - \frac{1}{q+1},$$

- Cet estimateur prend en compte les autocovariances $\hat{\gamma}_i$ du processus jusqu'à l'ordre q .
- La fenêtre q est telle que

$$\frac{1}{T} + \frac{q}{T} \rightarrow 0 \quad T \rightarrow \infty$$

Remarque

Pas de règle pour choisir q : (on voudrait une statistique qui ne dépende pas trop de q ce qui exclut la statistique R/S)

Test semi-paramétrique I

- Sous les hypothèses de dépendance faible, la statistique R/S a la distribution asymptotique suivante :

$$(q/T)T^{-\frac{1}{2}}R/S(q) \xrightarrow{d} \max_{0 \leq t \leq 1} W^0(t) - \min_{0 \leq t \leq 1} W^0(t),$$

c-à-d l'amplitude d'un pont Brownien sur l'intervalle unité

$$W^0(t) = W(t) - tW(1),$$

- tandis que pour la statistique V/S

$$M_T(q) \xrightarrow{d} U_{V/S} = \int_0^1 (W^0(t))^2 dt - \left(\int_0^1 W^0(t) dt \right)^2.$$

Test semi-paramétrique II

- Sous l'hypothèse de dépendance forte :

$$T^{-\frac{1}{2}} R/S(q) \xrightarrow{d} \sup_{0 \leq t \leq 1} W_{1/2+d}^0(t) - \inf_{0 \leq t \leq 1} W_{1/2+d}^0(t),$$

c-à-d l'amplitude d'un pont Brownien fractionnaire, avec :

$$W_{1/2+d}^0(t) = W_{1/2+d}(t) - tW_{1/2+d}(1),$$

- tandis que pour la statistique V/S

$$(q/T)^{2d} M_T \xrightarrow{d} \int_0^1 \left(W_{1/2+d}^0(t) \right)^2 dt - \left(\int_0^1 W_{1/2+d}^0(t) dt \right)^2.$$

Remarque

L'intérêt d'un tel centrage est que la puissance de la statistique V/S est moins sensible au choix de q que la statistique R/S . Le choix de q est jusqu'ici arbitraire. ($q \in [T^{1/3}, T^{1/2}]$)

Courbes taille-puissance

Introduites par Wilk et Gnanadesikan (1968)

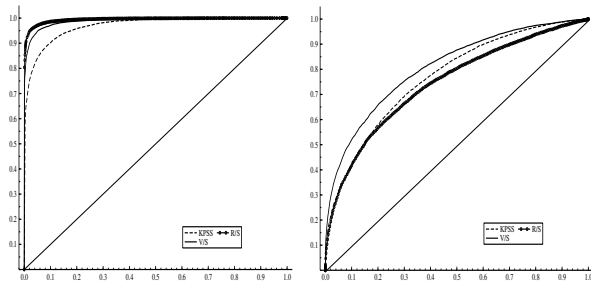


Figure: FARIMA(0, d ,0); $d = 0.30$ $q = 0$ (à gauche) $q = 30$ (à droite)

Remarque

La puissance de la statistique R/S tend rapidement vers sa taille quand q augmente. La puissance de la statistique V/S est bien moins sensible au choix de q .

Estimation semiparamétrique du paramètre de longue portée d

- Par méthodes semiparamétriques, nous entendons des méthodes qui requièrent des hypothèses minimales sur le processus,

Estimation semiparamétrique du paramètre de longue portée d

- Par méthodes semiparamétriques, nous entendons des méthodes qui requièrent des hypothèses minimales sur le processus,
- En particulier, nous ne caractérisons pas le processus étudié ni par son spectre, ni par sa densité, etc.

Estimation semiparamétrique du paramètre de longue portée d

- Par méthodes semiparamétriques, nous entendons des méthodes qui requièrent des hypothèses minimales sur le processus,
- En particulier, nous ne caractérisons pas le processus étudié ni par son spectre, ni par sa densité, etc.
- Nous nous contentons d'hypothèses minimales sur la forme fonctionnelle du spectre au voisinage de la fréquence zéro,

Estimation semiparamétrique du paramètre de longue portée d

- Par méthodes semiparamétriques, nous entendons des méthodes qui requièrent des hypothèses minimales sur le processus,
- En particulier, nous ne caractérisons pas le processus étudié ni par son spectre, ni par sa densité, etc.
- Nous nous contentons d'hypothèses minimales sur la forme fonctionnelle du spectre au voisinage de la fréquence zéro,
- Nous commençons par une spécification bien générale du spectre du processus étudié au voisinage de la fréquence zéro

$$f_Y(\lambda) \sim c_f \lambda^{-2d}, \quad \lambda \rightarrow 0_+, \quad c_f \in (0, \infty),$$

Estimation semiparamétrique du paramètre de longue portée d

- Par méthodes semiparamétriques, nous entendons des méthodes qui requièrent des hypothèses minimales sur le processus,
- En particulier, nous ne caractérisons pas le processus étudié ni par son spectre, ni par sa densité, etc.
- Nous nous contentons d'hypothèses minimales sur la forme fonctionnelle du spectre au voisinage de la fréquence zéro,
- Nous commençons par une spécification bien générale du spectre du processus étudié au voisinage de la fréquence zéro

$$f_Y(\lambda) \sim c_f \lambda^{-2d}, \quad \lambda \rightarrow 0_+, \quad c_f \in (0, \infty),$$

- Ce qui est la définition d'un processus à longue portée. (voir leçon précédente)

Estimateur local de Whittle

Défini à partir de l'estimateur de Whittle

$$\hat{d}_{LW} = \arg \min_d G(d, m) := \left\{ \ln \left(\frac{1}{m} \sum_{j=1}^m \frac{I_Y(\lambda_j)}{\lambda_j^{-2}} \right) - \frac{2d}{m} \sum_{j=1}^m \ln(\lambda_j) \right\},$$

- $I_Y(\lambda_j)$ est le périodogramme évalué aux m fréquences de Fourier $\lambda_j = 2\pi j/T, j = 1, \dots, m \ll [T/2]$,

Estimateur local de Whittle

Défini à partir de l'estimateur de Whittle

$$\hat{d}_{LW} = \arg \min_d G(d, m) := \left\{ \ln \left(\frac{1}{m} \sum_{j=1}^m \frac{I_Y(\lambda_j)}{\lambda_j^{-2d}} \right) - \frac{2d}{m} \sum_{j=1}^m \ln(\lambda_j) \right\},$$

- $I_Y(\lambda_j)$ est le périodogramme évalué aux m fréquences de Fourier $\lambda_j = 2\pi j/T, j = 1, \dots, m \ll [T/2]$,
- m est une fenêtre qui tend vers l'infini avec la taille de l'échantillon T mais plus lentement

$$1/m + m/T \rightarrow 0 \text{ quand } T \rightarrow \infty$$

Estimateur local de Whittle

Défini à partir de l'estimateur de Whittle

$$\hat{d}_{LW} = \arg \min_d G(d, m) := \left\{ \ln \left(\frac{1}{m} \sum_{j=1}^m \frac{I_Y(\lambda_j)}{\lambda_j^{-2}} \right) - \frac{2d}{m} \sum_{j=1}^m \ln(\lambda_j) \right\},$$

- $I_Y(\lambda_j)$ est le périodogramme évalué aux m fréquences de Fourier $\lambda_j = 2\pi j/T, j = 1, \dots, m \ll [T/2]$,
- m est une fenêtre qui tend vers l'infini avec la taille de l'échantillon T mais plus lentement

$$1/m + m/T \rightarrow 0 \text{ quand } T \rightarrow \infty$$

- L'estimateur est Gaussien, mais cette hypothèse n'est pas nécessaire pour le processus lui-même : son spectre est différentiable au voisinage de la fréquence zéro, et il admet une représentation moyenne mobile

Normalité asymptotique de l'estimateur local de Whittle

Théorème

$$\sqrt{m}(\hat{d}_{LW} - d) \xrightarrow{d} N(0, 1/4)$$

Normalité asymptotique de l'estimateur local de Whittle

Théorème

$$\sqrt{m}(\hat{d}_{LW} - d) \xrightarrow{d} N(0, 1/4)$$

- Ce résultat vient de l'hypothèse que le spectre

$$f_Y(\lambda) = C\lambda^{-2d} [1 + E_\beta(2d)\lambda^\beta + o(\lambda^\beta)], \quad 0 < |E_\beta(2d)| < \infty,$$

$\lambda \rightarrow 0^+$, avec $\beta \in (0, 2]$ contrôle le degré de variabilité du spectre au voisinage de la fréquence zéro

Normalité asymptotique de l'estimateur local de Whittle

Théorème

$$\sqrt{m}(\hat{d}_{LW} - d) \xrightarrow{d} N(0, 1/4)$$

- Ce résultat vient de l'hypothèse que le spectre

$$f_Y(\lambda) = C\lambda^{-2d} [1 + E_\beta(2d)\lambda^\beta + o(\lambda^\beta)], \quad 0 < |E_\beta(2d)| < \infty,$$

$\lambda \rightarrow 0^+$, avec $\beta \in (0, 2]$ contrôle le degré de variabilité du spectre au voisinage de la fréquence zéro

- C est une constante strictement positive,

Normalité asymptotique de l'estimateur local de Whittle

Théorème

$$\sqrt{m}(\hat{d}_{LW} - d) \xrightarrow{d} N(0, 1/4)$$

- Ce résultat vient de l'hypothèse que le spectre

$$f_Y(\lambda) = C\lambda^{-2d} [1 + E_\beta(2d)\lambda^\beta + o(\lambda^\beta)], \quad 0 < |E_\beta(2d)| < \infty,$$

$\lambda \rightarrow 0^+$, avec $\beta \in (0, 2]$ contrôle le degré de variabilité du spectre au voisinage de la fréquence zéro

- C est une constante strictement positive,
- la fenêtre m vérifie

$$\frac{1}{m} + \frac{m^{2\beta+1} \log^2 m}{T^{2\beta}} \rightarrow 0, \quad T \rightarrow \infty,$$

(Voir Delgado et Robinson, 1996)

Fenêtre optimale pour l'estimateur local de Whittle

Fenêtre minimisant l'erreur carrée moyenne :

$$m_{LW}^{opt} = \left[\frac{(\beta + 1)^4}{2\beta^3 E_\beta (2d)^2 (2\pi)^{2\beta}} \right]^{1/(1+2\beta)} T^{2\beta/(1+2\beta)}.$$

Nécessite cependant la connaissance de d , β , etc.

Hypothèses supplémentaires pour obtenir une fenêtre faisable :

- $\beta = 2$, c-à-d le cas où le spectre est le plus lisse,
- On suppose la spécification suivante pour le spectre :

$$f_Y(\lambda) = |2 \sin(\lambda/2)|^{-2d} f_*(\lambda),$$

Alors on estime le paramètre de longue portée à l'aide de la procédure itérative :

$$\hat{d}^{(k)} = \arg \min_d G(d, m^{(k)}),$$

$$m_{LW}^{opt(k+1)} = \left(\frac{3T}{4\pi} \right)^{4/5} \left| E_2(d^{(k)}) \right|^{-2/5}, \quad m^{(0)} = T^{4/5}$$

Fenêtre optimale pour l'estimateur local de Whittle

- Sans connaissance additionnelle de E_2 , cette procédure n'est pas faisable.

Fenêtre optimale pour l'estimateur local de Whittle

- Sans connaissance additionnelle de E_2 , cette procédure n'est pas faisable.
- Hypothèses supplémentaires :

Fenêtre optimale pour l'estimateur local de Whittle

- Sans connaissance additionnelle de E_2 , cette procédure n'est pas faisable.
- Hypothèses supplémentaires :
 - Spécification du spectre

$$f_Y(\lambda) = |2 \sin(\lambda/2)|^{-2d} f_*(\lambda)$$

$f_*(\cdot)$ étant deux fois continûment différentiable et positive pour $\lambda = 0$

Fenêtre optimale pour l'estimateur local de Whittle

- Sans connaissance additionnelle de E_2 , cette procédure n'est pas faisable.
- Hypothèses supplémentaires :
 - Spécification du spectre

$$f_Y(\lambda) = |2 \sin(\lambda/2)|^{-2d} f_*(\lambda)$$

$f_*(\cdot)$ étant deux fois continûment différentiable et positive pour $\lambda = 0$

- On suppose de plus que

$$E_2((2d)^{(k)}) = \frac{f_*''(0)}{2f_*(0)} + \frac{2d}{24}$$

Fenêtre optimale pour l'estimateur local de Whittle

- Sans connaissance additionnelle de E_2 , cette procédure n'est pas faisable.
- Hypothèses supplémentaires :
 - Spécification du spectre

$$f_Y(\lambda) = |2 \sin(\lambda/2)|^{-2d} f_*(\lambda)$$

$f_*(\cdot)$ étant deux fois continûment différentiable et positive pour $\lambda = 0$

- On suppose de plus que

$$E_2((2d)^{(k)}) = \frac{f_*''(0)}{2f_*(0)} + \frac{2d}{24}$$

- ou $f_*(0)$ et sa dérivée seconde $f_*''(0)$ sont les premiers et troisièmes coefficients de la régression suivante

$$I(\lambda_j) = G(|1 - \exp(i\lambda)|^{-2d(0)}(1, \lambda_j, \lambda_j^2/2))$$

aux fréquences de Fourier $\lambda_j, j = 1, \dots, m^{(0)}$.

Autre estimateur : l'estimateur du log-périodogramme I

- Cet estimateur est aussi basé sur l'approximation du spectre au voisinage de la fréquence zéro :

$$f(\lambda) \sim c_f \lambda^{-2d}, \quad \rightarrow 0_+,$$

Autre estimateur : l'estimateur du log-périodogramme I

- Cet estimateur est aussi basé sur l'approximation du spectre au voisinage de la fréquence zéro :

$$f(\lambda) \sim c_f \lambda^{-2d}, \quad \rightarrow 0_+,$$

- Donc,

$$I(\lambda_j) = f(\lambda_j) \frac{I(\lambda_j)}{f(\lambda_j)} \sim c_f \lambda_j^{-2d} \frac{I(\lambda_j)}{f(\lambda_j)} \quad (C)$$

Autre estimateur : l'estimateur du log-périodogramme I

- Cet estimateur est aussi basé sur l'approximation du spectre au voisinage de la fréquence zéro :

$$f(\lambda) \sim c_f \lambda^{-2d}, \quad \rightarrow 0_+,$$

- Donc,

$$I(\lambda_j) = f(\lambda_j) \frac{I(\lambda_j)}{f(\lambda_j)} \sim c_f \lambda_j^{-2d} \frac{I(\lambda_j)}{f(\lambda_j)} \quad (C)$$

- En prenant le logarithme on obtient la régression

$$\log I(\lambda_j) = \log c_f - 2d \log \lambda_j + \log \left(\frac{I(\lambda_j)}{f(\lambda_j)} \right)$$

Autre estimateur : l'estimateur du log-périodogramme II

- Sous des hypothèses de Gaussianité et de dépendance à faible portée, les $\log \left(\frac{I(\lambda_j)}{f(\lambda_j)} \right)$ sont asymptotiquement i.i.d.

Autre estimateur : l'estimateur du log-périodogramme II

- Sous des hypothèses de Gaussianité et de dépendance à faible portée, les $\log \left(\frac{I(\lambda_j)}{f(\lambda_j)} \right)$ sont asymptotiquement i.i.d.
- Bien que cette propriété ne soit pas vraie dans le cas de la longue portée, cette relation (C) est utilisée pour estimer d avec la régression suivante :

$$\log I(\lambda_j) = \log c_f - 2d \log \lambda_j + \varepsilon_j, \quad j = 1, \dots, m., \quad \varepsilon_j = \log \left(\frac{I(\lambda_j)}{f(\lambda_j)} \right) - \eta$$

η est la constante d'Euler.

Normalité asymptotique et fenêtre optimale

Cet estimateur est obtenu par moindres carrés,

Théorème

$$\sqrt{m}(\hat{d}_{LP} - d) \xrightarrow{d} N\left(0, \frac{\pi^2}{24}\right).$$

Normalité asymptotique et fenêtre optimale

Cet estimateur est obtenu par moindres carrés,

Théorème

$$\sqrt{m}(\hat{d}_{LP} - d) \xrightarrow{d} N\left(0, \frac{\pi^2}{24}\right).$$

- La fenêtre m vérifie les mêmes conditions que dans le cas de l'estimateur local de Whittle.

Normalité asymptotique et fenêtre optimale

Cet estimateur est obtenu par moindres carrés,

Théorème

$$\sqrt{m}(\hat{d}_{LP} - d) \xrightarrow{d} N\left(0, \frac{\pi^2}{24}\right).$$

- La fenêtre m vérifie les mêmes conditions que dans le cas de l'estimateur local de Whittle.
- Une fenêtre optimale m_{LP}^{opt} minimisant l'erreur moyenne quadratique

$$m_{LP}^{opt} = \left[\frac{27(2f_*(0))^2}{512\pi^2(f_*''(0))^2} \right]^{1/5} T^{4/5}.$$

Normalité asymptotique et fenêtre optimale

Cet estimateur est obtenu par moindres carrés,

Théorème

$$\sqrt{m}(\hat{d}_{LP} - d) \xrightarrow{d} N\left(0, \frac{\pi^2}{24}\right).$$

- La fenêtre m vérifie les mêmes conditions que dans le cas de l'estimateur local de Whittle.
- Une fenêtre optimale m_{LP}^{opt} minimisant l'erreur moyenne quadratique

$$m_{LP}^{opt} = \left[\frac{27(2f_*(0))^2}{512\pi^2(f_*''(0))^2} \right]^{1/5} T^{4/5}.$$

- Une fenêtre faisable est obtenue en estimant $f_*(0)$ et $f_*''(0)$ comme pour la fenêtre optimale de l'estimateur local de Whittle.

Estimateurs du type "Pox-plot"

- Ce sont des estimateurs basés sur les statistiques R/S et V/S , (avec $q = 0$).

Estimateurs du type "Pox-plot"

- Ce sont des estimateurs basés sur les statistiques R/S et V/S , (avec $q = 0$).
- Sous l'hypothèse de dépendance forte, la statistique R/S peut être utilisée pour estimer le paramètre d . (Mandelbrot et Taqqu, 1979)

Estimateurs du type "Pox-plot"

- Ce sont des estimateurs basés sur les statistiques R/S et V/S , (avec $q = 0$).
- Sous l'hypothèse de dépendance forte, la statistique R/S peut être utilisée pour estimer le paramètre d . (Mandelbrot et Taqqu, 1979)
- Définissons $\hat{s}_T^2 = \hat{s}_T^2(0)$, alors $\hat{s}_T^2 \rightarrow \text{Var } X_0$.

Estimateurs du type "Pox-plot"

- Ce sont des estimateurs basés sur les statistiques R/S et V/S , (avec $q = 0$).
- Sous l'hypothèse de dépendance forte, la statistique R/S peut être utilisée pour estimer le paramètre d . (Mandelbrot et Taqqu, 1979)
- Définissons $\hat{\sigma}_T^2 = \hat{\sigma}_T^2(0)$, alors $\hat{\sigma}_T^2 \rightarrow \text{Var } X_0$.
- Puisque

$$S_k = \sum_{j=1}^k (X_j - EX_j) - \frac{k}{T} \sum_{j=1}^T (X_j - EX_j),$$

$$\frac{1}{T^{1/2+d}} \sum_{j=1}^{[Tt]} (X_j - EX_j) \xrightarrow{D[0,1]} C W_{1/2+d}(t), \quad C \text{ constante positive.}$$

Estimateurs du type "Pox-plot"

- Ce sont des estimateurs basés sur les statistiques R/S et V/S , (avec $q = 0$).
- Sous l'hypothèse de dépendance forte, la statistique R/S peut être utilisée pour estimer le paramètre d . (Mandelbrot et Taqqu, 1979)
- Définissons $\hat{\sigma}_T^2 = \hat{\sigma}_T^2(0)$, alors $\hat{\sigma}_T^2 \rightarrow \text{Var } X_0$.
- Puisque

$$S_k = \sum_{j=1}^k (X_j - EX_j) - \frac{k}{T} \sum_{j=1}^T (X_j - EX_j),$$

$$\frac{1}{T^{1/2+d}} \sum_{j=1}^{[Tt]} (X_j - EX_j) \xrightarrow{D[0,1]} C W_{1/2+d}(t), \quad C \text{ constante positive.}$$

- Donc

$$\frac{\hat{R}_T}{T^{1/2+d}} \xrightarrow{d} C \left[\max_{0 \leq t \leq 1} W_{1/2+d}^0(t) - \min_{0 \leq t \leq 1} W_{1/2+d}^0(t) \right],$$

Estimateurs du type "Pox-plot"

- Ce sont des estimateurs basés sur les statistiques R/S et V/S , (avec $q = 0$).
- Sous l'hypothèse de dépendance forte, la statistique R/S peut être utilisée pour estimer le paramètre d . (Mandelbrot et Taqqu, 1979)
- Définissons $\hat{\sigma}_T^2 = \hat{\sigma}_T^2(0)$, alors $\hat{\sigma}_T^2 \rightarrow \text{Var } X_0$.
- Puisque

$$S_k = \sum_{j=1}^k (X_j - EX_j) - \frac{k}{T} \sum_{j=1}^T (X_j - EX_j),$$

$$\frac{1}{T^{1/2+d}} \sum_{j=1}^{[Tt]} (X_j - EX_j) \xrightarrow{D[0,1]} C W_{1/2+d}(t), \quad C \text{ constante positive.}$$

- Donc

$$\frac{\hat{R}_T}{T^{1/2+d}} \xrightarrow{d} C \left[\max_{0 \leq t \leq 1} W_{1/2+d}^0(t) - \min_{0 \leq t \leq 1} W_{1/2+d}^0(t) \right],$$

- $W_{1/2+d}^0(t)$: pont Brownien fractionnaire.

Estimateurs du type "Pox-plot" : estimation

- Donc

$$\frac{1}{T^{1/2+d}} \frac{\hat{R}_T}{\hat{s}_T} \xrightarrow{d} \frac{C \left[\max_{0 \leq t \leq 1} W_{1/2+d}^0(t) - \min_{0 \leq t \leq 1} W_{1/2+d}^0(t) \right]}{\text{Var}(X_0)^{1/2}}.$$

- Cette équation est la base de l'estimateur R/S :
- En prenant le log des deux côtés de cette expression, on obtient :

$$\log \left(\frac{\hat{R}_T}{\hat{s}_T} \right) \approx \left(\frac{1}{2} + d \right) \log T + C, \quad T \rightarrow \infty,$$

- $1/2 + d$ est la pente de la régression de $\log(\hat{R}_T/\hat{s}_T)$ sur $\log T$.

Estimateurs du type "Pox-plot" : extension

- Ce principe est étendu aux statistiques KPSS et V/S .

Estimateurs du type "Pox-plot" : extension

- Ce principe est étendu aux statistiques KPSS et V/S .
- La pente de la régression de $\log \hat{T}_T$ sur $2 \log T$ donne d , avec

$$\hat{d}_{KPSS} - d = O_P(1/\log T)$$

Estimateurs du type "Pox-plot" : extension

- Ce principe est étendu aux statistiques KPSS et V/S .
- La pente de la régression de $\log \hat{T}_T$ sur $2 \log T$ donne d , avec

$$\hat{d}_{KPSS} - d = O_P(1/\log T)$$

- De même, la pente de la régression de $\log \hat{M}_T$ sur $2 \log T$ donne d , avec

$$\hat{d}_{V/S} - d = O_P(1/\log T)$$

Estimateurs du type "Pox-plot" : extension

- Ce principe est étendu aux statistiques KPSS et V/S .
- La pente de la régression de $\log \hat{T}_T$ sur $2 \log T$ donne d , avec

$$\hat{d}_{KPSS} - d = O_P(1/\log T)$$

- De même, la pente de la régression de $\log \hat{M}_T$ sur $2 \log T$ donne d , avec

$$\hat{d}_{V/S} - d = O_P(1/\log T)$$

Estimateurs du type "Pox-plot" : extension

- Ce principe est étendu aux statistiques KPSS et V/S .
- La pente de la régression de $\log \hat{T}_T$ sur $2 \log T$ donne d , avec

$$\hat{d}_{KPSS} - d = O_P(1/\log T)$$

- De même, la pente de la régression de $\log \hat{M}_T$ sur $2 \log T$ donne d , avec

$$\hat{d}_{V/S} - d = O_P(1/\log T)$$

Ces estimateurs ont quelques inconvénients :

- Pas de théorie asymptotique (intervalles de confiance doivent être évalués par simulations)

Estimateurs du type "Pox-plot" : extension

- Ce principe est étendu aux statistiques KPSS et V/S .
- La pente de la régression de $\log \hat{T}_T$ sur $2 \log T$ donne d , avec

$$\hat{d}_{KPSS} - d = O_P(1/\log T)$$

- De même, la pente de la régression de $\log \hat{M}_T$ sur $2 \log T$ donne d , avec

$$\hat{d}_{V/S} - d = O_P(1/\log T)$$

Ces estimateurs ont quelques inconvénients :

- Pas de théorie asymptotique (intervalles de confiance doivent être évalués par simulations)
- Ils ont la faible vitesse de convergence $\log T$.

Estimateurs robustes aux tendances

Un processus avec tendance est défini comme suit :

Définition

$$Y_t = g(t) + \varepsilon_t,$$

- La dépendance de longue portée peut être confondue avec une tendance (voir Bhattacharya *et al.*, 1983).
- Les tests et estimateurs présentés dans cette leçon peuvent être étendus au cas avec tendance en utilisant le périodogramme “effilé”
 - Avec l'estimateur local de Whittle
 - Avec l'estimateur du log periodogramme

Références I



Bhattacharya, R.N., Gupta, V.K., Waymire, E. (1983). The Hurst effect under trends. *Journal of Applied Probability* **20**, 649–662.



Delgado, M., Robinson, P.M. (1996). Optimal spectral bandwidth for long memory. *Statistica Sinica*, **67**, 97–112.



Beran, J. (1994). *Statistics for Long Memory Processes*. Chapman & Hall.



Giraitis, L., Kokoszka, P., Leipus, R., Teyssière, G. (2000). Semiparametric estimation of the intensity of long-memory in conditional heteroskedasticity. *Statistical Inference for Stochastic Processes*, **3**, 113–128.



Giraitis, L., Kokoszka, P.S., Leipus, R. (2001). Testing for long memory in the presence of a general trend *Journal of Applied Probability*, **38**, 1033–1054.



Giraitis, L., Kokoszka, P.S., Leipus, R., Teyssière, G. (2003a). Rescaled variance and related tests for long memory in volatility and levels. *Journal of Econometrics*, **112**, 265–294.



Giraitis, L., Kokoszka, P.S., Leipus, R., Teyssière, G. (2003b). On the power of R/S-type tests under contiguous and semi long-memory alternatives. *Acta Applicandae Mathematicae* **78**, 285–299.



Giraitis, L., Leipus, R., Surgailis, D. (2007). Recent advances in ARCH modelling, dans *Long Memory in Economics*, G. Teyssière et al. éditeurs, Springer Verlag.

Références II



Giraitis, L., Leipus, R., Philippe, A. (2006). A test for stationarity versus trends and unit roots for a wide class of dependent errors. *Econometric Theory*, **22**, 989–1029.



Giraitis, L., Koul, H., Surgailis, D. (2010). *Large Sample Inference for Long–Memory Processes*, World Scientific Publishing, à paraître.



Mandelbrot, B.B., Taqqu, M.S. (1979). Robust R/S analysis of long-run serial correlation. *Proceedings of the 42nd Session of the International Statistical Institute, Manila (1979)*. *Bulletin of the I.S.I.*, Vol **48**, Book 2, 69–104.



Robinson, P.M. (1994). Efficient tests of nonstationary hypotheses. *Journal of American Statistical Association*, **89**, 1420–1437.



Robinson, P.M. (1995a). Gaussian semiparametric estimation of long–range dependence. *The Annals of Statistics*, **23**, 1630–1661.



Robinson, P.M. (1995b). Log periodogram regression of time series with long range dependence. *The Annals of Statistics*, **23**, 1048–1072.



Wilk, M.B., Gnanadesikan, R. (1968). Probability plotting methods for the analysis of data. *Biometrika*, **55**, 1–17.