The Frequency-Domain Bootstrap of Dahlhaus & Janas

Dmitriy Izyumin, Eugene Shvarts, Nick Ulle

Final Project for Prof. Aue's STA 237A, Time Series Analysis

June 5, 2014

Bootstrapping and Time Series

- Bootstrapping is based on resampling i.i.d. variables
- How to deal with the dependence structure of time series?
- Need (approximately) i.i.d. variables
  - Time Domain: residuals, innovations
  - Frequency Domain: (Studentized) periodogram ordinates
- Paper describes favorable properties of the latter

Spectral Mean

- $\phi = (\phi^{(1)}, \cdots, \phi^{(d)})$, each $\phi^{(i)}$ of bounded variation
- I_T - tapered periodogram
- Spectral mean $A(\phi, f) = \int_0^\pi \phi(\alpha) f(\alpha) d\alpha$
- Canonical estimate $A(\phi, I_T) = \int_0^\pi \phi(\alpha) I_T(\alpha) d\alpha$

Spectral Mean - Important Examples

- Spectral mean $A(\phi, f) = \int_0^\pi \phi(\alpha) f(\alpha) d\alpha$
- Autocovariance estimate $\hat{\gamma}(u), u \in \mathbb{Z}$
  $\phi(\alpha) = 2\cos(\alpha u)$
- Spectral distribution function estimate $\hat{F}_T(\lambda)$
  $\phi(\alpha) = \mathbb{1}_{[0,\lambda]}(\alpha)$

Ratio Statistics - Definition

- Normalized spectral density: $g(\alpha) = \frac{f(\alpha)}{F(\pi)}$
- Estimate: $J_T(\alpha) = \frac{I_T(\alpha)}{\hat{F}_T(\pi)}$
- Ratio statistic: $A(\phi, J_T) = \frac{\int_0^\pi \phi(\alpha) I_T(\alpha) d\alpha}{\int_0^\pi I_T(\alpha) d\alpha} = \frac{A(\phi, I_T)}{A(1, I_T)}$

Ratio Statistics - Important Examples

- Ratio statistic: $A(\phi, J_T) = \frac{\int_0^\pi \phi(\alpha) I_T(\alpha) d\alpha}{\int_0^\pi I_T(\alpha) d\alpha} = \frac{A(\phi, I_T)}{A(1, I_T)}$
- Autocorrelation estimate $\hat{\rho}_T(u), u \in \mathbb{Z}$
  $\phi(\alpha) = \cos(\alpha u)$
- Normalized sdf estimate $\hat{F}_T(\lambda)/\hat{F}_T(\pi)$
  $\phi(\alpha) = \mathbb{1}_{[0,\lambda]}(\alpha)$

Whittle Estimators

- Parametric family of spectral densities:
  $\mathcal{F} = \{f_\theta : \theta \in \Theta\}, \Theta \in \mathbb{R}$
- Whittle's likelihood: $\mathcal{L}_T(\theta) = \frac{1}{2\pi} \int_0^\pi \left[\log f_\theta(\alpha) + \frac{I_T(\alpha)}{f_\theta(\alpha)}\right] d\alpha$
- $\hat{\theta}$ obtained by minimizing $\mathcal{L}_T(\theta)$
- Spectral mean with $\phi(\alpha) = \nabla \frac{1}{f_\theta}(\alpha)$

Spectral Mean Distribution

If we were to Studentize our periodogram values $\{I_j\}$ with estimated spectral density $\hat{f}$ in order to obtain bootstrap values $\{I_j^*\}$, we could form the bootstrap approximation

$$B(\phi, I_T^*) := \frac{\pi}{n} \sum_{j=1}^{n} \phi_j I_j^* \quad \text{, where} \quad \phi_j := \phi(2\pi j/T) \quad .$$

Spectral Mean Distribution

If we were to Studentize our periodogram values $\{I_j\}$ with estimated spectral density $\hat{f}$ in order to obtain bootstrap values $\{I_j^*\}$, we could form the bootstrap approximation

$$B(\phi, I_T^*) := \frac{\pi}{n} \sum_{j=1}^{n} \phi_j I_j^* \quad \text{, where} \quad \phi_j := \phi(2\pi j/T) \quad .$$

Under appropriate conditions ...

- $\sqrt{T}\left(A(\phi, I_T) - A(\phi, f)\right)$ is asymptotically normal with variance $= \quad 2\pi \int \phi^2 f^2 + (\kappa_4/\sigma^4)\left(\int \phi f\right)^2 \quad .$
- $\sqrt{T}\left(B(\phi, I_T^*) - B(\phi, \hat{f})\right)$ is asymptotically normal with variance $\propto \quad 2\pi \int \phi^2 f^2 \quad .$

Spectral Mean Distribution

If we were to Studentize our periodogram values $\{I_j\}$ with estimated spectral density $\hat{f}$ in order to obtain bootstrap values $\{I_j^*\}$, we could form the bootstrap approximation

$$B(\phi, I_T^*) := \frac{\pi}{n} \sum_{j=1}^{n} \phi_j I_j^* \quad \text{, where} \quad \phi_j := \phi(2\pi j/T) \quad .$$

Under appropriate conditions ...

- $\sqrt{T}\left(A(\phi, I_T) - A(\phi, f)\right)$ is asymptotically normal with variance $= \quad 2\pi \int \phi^2 f^2 + (\kappa_4/\sigma^4)\left(\int \phi f\right)^2 \;$.
- $\sqrt{T}\left(B(\phi, I_T^*) - B(\phi, \hat{f})\right)$ is asymptotically normal with variance $\propto \quad 2\pi \int \phi^2 f^2 \;$.

For ratio statistics, second term vanishes!

Algorithm

Let $n = \lceil T/2 \rceil$ and $I_j := I_T(2\pi j/T)$.

1. Obtain periodogram sample values $\{I_j\}$, $j = 1 : n$.
2. Obtain estimate $\hat{f}$ of spectral density f (e.g., kernel estimate).

Algorithm

Let $n = \lceil T/2 \rceil$ and $I_j := I_T(2\pi j/T)$.

1. Obtain periodogram sample values $\{I_j\},\ j = 1 : n$.
2. Obtain estimate $\hat{f}$ of spectral density f (e.g., kernel estimate).
3. Obtain Studentized periodogram samples $\hat{\epsilon}_j := I_j/\hat{f}_j$.
4. Rescale the $\hat{\epsilon}_j$'s by their mean, so that $\tilde{\epsilon}_j := \hat{\epsilon}_j/\{(1/n)\sum_j \hat{\epsilon}_j\}$.

Algorithm

Let $n = \lceil T/2 \rceil$ and $I_j := I_T(2\pi j/T)$.

1. Obtain periodogram sample values $\{I_j\}$, $j = 1 : n$.
2. Obtain estimate $\hat{f}$ of spectral density f (e.g., kernel estimate).
3. Obtain Studentized periodogram samples $\hat{\epsilon}_j := I_j/\hat{f}_j$.
4. Rescale the $\hat{\epsilon}_j$'s by their mean, so that $\tilde{\epsilon}_j := \hat{\epsilon}_j/\{(1/n)\sum_j \hat{\epsilon}_j\}$.
5. Sample the empirical distribution of $\{\tilde{\epsilon}_j\}$ to obtain a bootstrap sample $\{\epsilon_j^*\}$ (whose size may differ from n).
6. Define bootstrap periodogram values via $I_j^* := \hat{f}_j \epsilon_j^*$.

Some Assumptions

Performance guarantee of the bootstrap approximation requires ...

- $\{X_t : t \in \mathbb{Z}\}$ is a real-valued linear process: $X_t = \sum_{u \in \mathbb{Z}} a_u \xi_{t-u}$
  The $\{\xi_t\}$ satisfy $\mathbb{E}\xi_1 = 0$, $\mathbb{E}\xi_1^2 = 1$, $\mathbb{E}\xi_1^8 < \infty$, $\mathbb{E}\xi_t^3 = 0$.
  The $\{a_t\}$ satisfy $\sum_t t^2 |a_t| < \infty$.
- $\sup_{\alpha \in [0,\pi]} |\hat{f}(\alpha) - f(\alpha)| \to 0$ almost surely.
- ϕ is a d-dimensional vector of BV functions $\phi^{(r)} : [0,\pi] \to \mathbb{R}$, each with even periodic extension to all of $\mathbb{R}$.
- The taper h_t has the form $h_t := h(t/T)$, with $h : \mathbb{R} \to [0,1]$ a BV function with support in $(0,1]$.

More Technical Assumptions

Edgeworth expansions are used to show that we outperform the normal approximation. For these we need ...

- The $\{a_t\}$ and Fourier coefficients $\{\hat{\phi}(\omega)\}$ vanish exponentially. That is, for a fixed $\tau \in (0,1)$, $|a_t| \leq \tau^{|t|}$ and $\|\hat{\phi}(\omega)\| \leq \tau^{|\omega|}$.
- several other detailed technical assumptions on h, ξ, and ϕ (more than five minutes allow! See the paper).

Convergence Rate

- To bootstrap the distribution of $A(\phi, J_T) - A(\phi, g)$, use the statistic $B(\phi, J_T^*) - B(\phi, \hat{g})$, where $\hat{g}_j := \hat{f}_j / \{(\pi/n) \sum_k \hat{f}_k\}$.
- Let $D_T^2 = V_T^{-1}$, where V_T is the covariance matrix of $\sqrt{T} A(\phi, J_T)$. Define $\hat{D}_T$ analogously via $\sqrt{T} B(\phi, J_T^*)$.
- Let P^* denote the conditional distribution given the data.

Theorem

Suppose the assumptions stated in the paper hold. Then for almost all samples $\{I_j\}$,

$$\sup_C \Big| P(\sqrt{T} D_T(A(\phi, J_T) - A(\phi, g)) \in C) -$$

$$P^*(\sqrt{T} \hat{D}_T(B(\phi, J_T^*) - B(\phi, \hat{g})) \in C) \Big| = o(T^{-1/2}) \quad .$$

The sup is taken over convex measurable subsets of $\mathbb{R}^d$.

Convergence Rate

Theorem

Suppose the assumptions stated in the paper hold. Then for almost all samples $\{I_j\}$,

$$\sup_C \Big| P(\sqrt{T} D_T(A(\phi, J_T) - A(\phi, g)) \in C) -$$

$$P^*(\sqrt{T} \hat{D}_T(B(\phi, J_T^*) - B(\phi, \hat{g})) \in C) \Big| = o(T^{-1/2}) \quad .$$

The sup is taken over convex measurable subsets of $\mathbb{R}^d$.

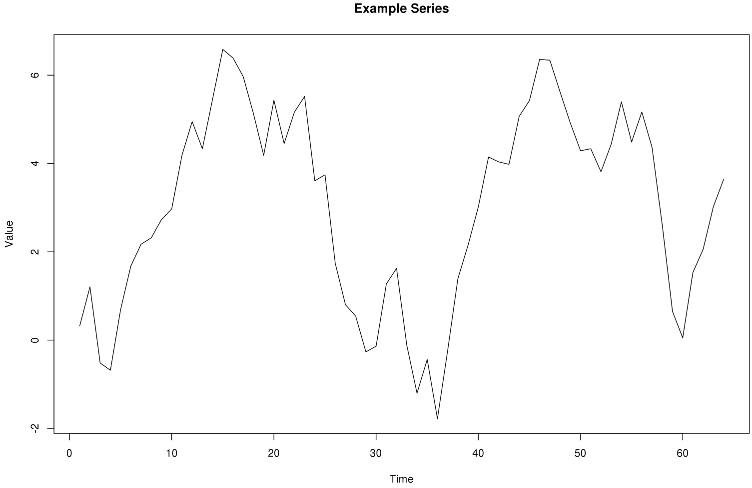
In short, estimating the distribution of the standardized spectral density by resampling from the standardized periodogram is consistent, and the convergence rate is asymptotically better than that achieved by the normal approximation.

Setup

Simulate the $\mathrm{AR}(1)$ process $X_t = 0.9X_{t-1} + Z_t$.

- Innovations $\{Z_t\} \overset{\text{iid}}{\sim} \mathsf{U}(-\sqrt{3}, \sqrt{3})$ are mean 0, variance 1 white noise.
- Draw 64 observations.
- Recall that $\hat{\rho}(1)$ is the Yule-Walker estimate for the AR parameter 0.9.

Series Plot



Practical Details

- Computed the periodogram with a 10% Tukey-Hanning taper.
- Smoothed the log-periodogram using the Epanechnikov kernel

$$K(x) = \frac{3}{4}\pi\left[1 - \left(\frac{x}{\pi}\right)^2\right].$$

and bandwidth 0.1 to obtain an estimate of the spectral density.

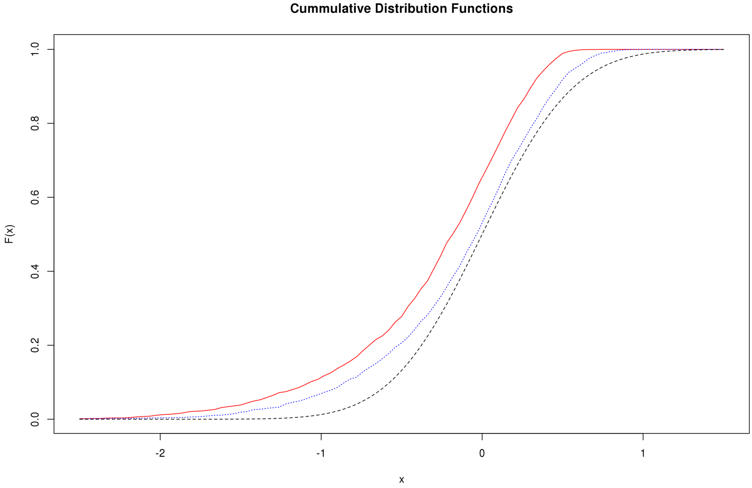
Error Distribution

- Since $\hat{\rho}(1)$ is the Yule-Walker estimate for the AR parameter, then as $n \to \infty$,

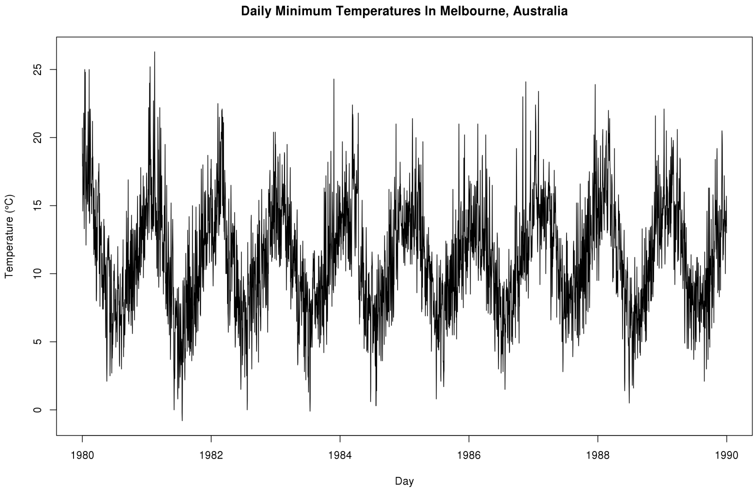
$$\sqrt{T}\left(\frac{\hat{\rho}(1) - 0.9}{\sqrt{c}}\right) \xrightarrow{\mathrm{d}} \mathrm{N}(0, 1 - 0.9^2),$$

  where c is a correction for the taper.
- 2000 bootstrapped estimates (from one simulation) produce a bootstrapped distribution.
- 2000 simulations approximate the true distribution.

Distributions Plot

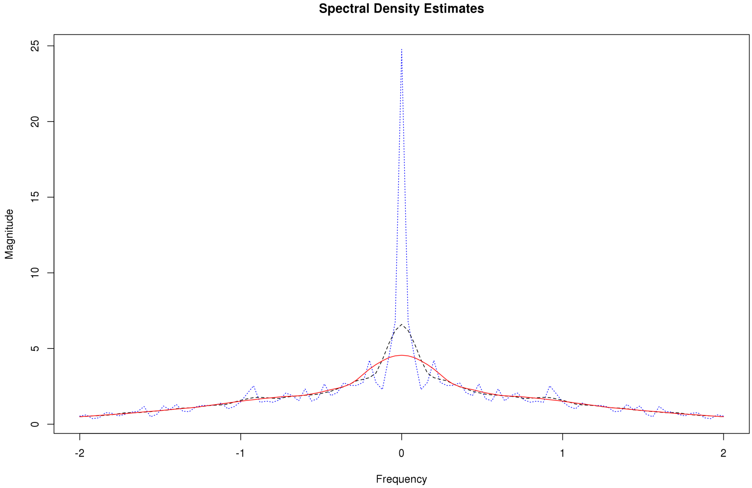


Series Plot



Daily Minimum Temperatures In Melbourne, Australia
Temperature (°C)
0
5
10
15
20
25
1980
1982
1984
1986
1988
1990
Day

Spectral Density Estimate



Autocorrelation Estimates

- Lag-1 autocorrelation confidence interval $(0.5234, 0.5787)$.
- Lag-2 autocorrelation confidence interval $(0.2041, 0.2917)$.