On Hurst exponent estimation under heavy-tailed distributions

PhDr. Ladislav Krištoufek
kristoufek@ies-prague.org
PhDr. Jozef Barunik
barunik@utia.cas.cz
Institute of Economic Studies at Charles University in Prague, Czech Republic
Institute of Information Theory and Automation of Academy of Sciences of the Czech Republic

Presentation structure

- Motivation
- Hurst exponent estimators
- Stable distributions
- Monte Carlo simulations methodology
- Results

Long-range dependence (LRD)

- LRD process is asymptotically described by hyperbolically decaying autocorrelation function:

$$\rho(k) \approx Ck^{2H-2}$$

- *H* is called Hurst exponent after river engineer Harold Edwin Hurst
- Estimation of *H* is crucial for LRD detection

Hurst exponent values

- Hurst exponent lies in interval $0 < H < 1$
- $H = 0.5$ → independent or short-range dependent (SRD) process
- $H > 0.5$ → persistent process
- $H < 0.5$ → anti-persistent process

Persistence vs. anti-persistence

- Persistent process: an increment (decrement) is statistically more likely to be followed by an increment (decrement) so that a trend is followed
- Anti-persistent process: an increment (decrement) is statistically more likely to be followed by decrement (increment) so that the process reverts more frequently than random process (also called reverting process)

Why does LRD matter?

- Predictability
- Periodic and non-periodic cycles detection
- Portfolio selection
- Option pricing
- Risk management

Hurst exponent estimation

- Estimating *H* directly from the autocorrelation function is impossible
- Two most popular techniques:
  - Rescaled range analysis (R/S analysis)
  - Detrended fluctuation analysis (DFA)

R/S analysis

- Time series of length *T* is divided into *N* adjacent sub-periods of length *υ*
- For each sub-period, rescaled range is calculated as *R/S*, where *R* is a range of time series profile (cumulative deviations from the mean) and *S* is a standard deviation of returns
- Average rescaled range then scales for different lengths *υ* (a scale) as

$$\left(R/S\right)_{\upsilon} \approx c * \upsilon^{H}$$

DFA

- Instead of rescaled ranges, DFA is based on scaling of deviation from the chosen trend (DFA-0, DFA-1, etc. with respect to degree of polynomial fit)
- Fluctuations from the trend $F_{DFA}(\upsilon) = \sqrt{1/T \sum_{t=1}^{T} Y_{\upsilon,l}^{2}(t)}$ scale as $F_{DFA}(\upsilon) \approx c * \upsilon^{H}$

Main motivation of the paper

- Hurst exponent can only be estimated → estimators testing
- Majority of research papers base Monte Carlo simulations of standard normal distribution for the independent process
- Financial time series are known to be fat-tailed
- Main motivation – to fill this gap!

Stable distributions

- If X, X_1, X_2,…,X_n are iid stable random variables, then for every n

$$X_1 + X_2 + \ldots + X_n \stackrel{d}{=} c_n X + d_n$$

- Class of laws satisfying the above is characterized by parameters *(α, β, γ, δ)* and stand for characteristic parameter, skewness parameter, scale and location parameter, respectively.

Heavy tails of stable distributions

- Asymptotic tail behavior can be described as

$$\lim_{x \to \infty} x^{\alpha} P(X > x) = c_{\alpha} (1 + \beta) \gamma^{\alpha}$$

$$\lim_{x \to \infty} x^{\alpha} P(X < -x) = c_{\alpha} (1 - \beta) \gamma^{\alpha},$$

$$c_{\alpha} = \sin\left(\frac{\pi \alpha}{2}\right) \Gamma(\alpha)$$

Nolan's parameterization

- There are several parameterizations of stable distributions in literature
- We use parameterization of Nolan (2003):

$$\phi(u) = \begin{cases} \exp\left(-\gamma^\alpha |u|^\alpha \left[1 + i\beta \left(\tan \frac{\pi\alpha}{2}\right)(\mathrm{sign}\, u)\left(|\gamma u|^{1-\alpha} - 1\right)\right] + i\delta u\right) & \alpha \neq 1 \\ \exp\left(-\gamma |u| \left[1 + i\beta \frac{2}{\pi} (\mathrm{sign}\, u) \ln(\gamma |u|)\right] + i\delta u\right) & \alpha = 1 \end{cases}$$

Monte Carlo simulations design (1)

- For all of the tested methods (R/S, DFA-0, DFA-1 and DFA-2), we use time series length T and scales u which are integer power of basis 2
- We use minimum scale of 16 and maximum scale of a quarter of the time series length

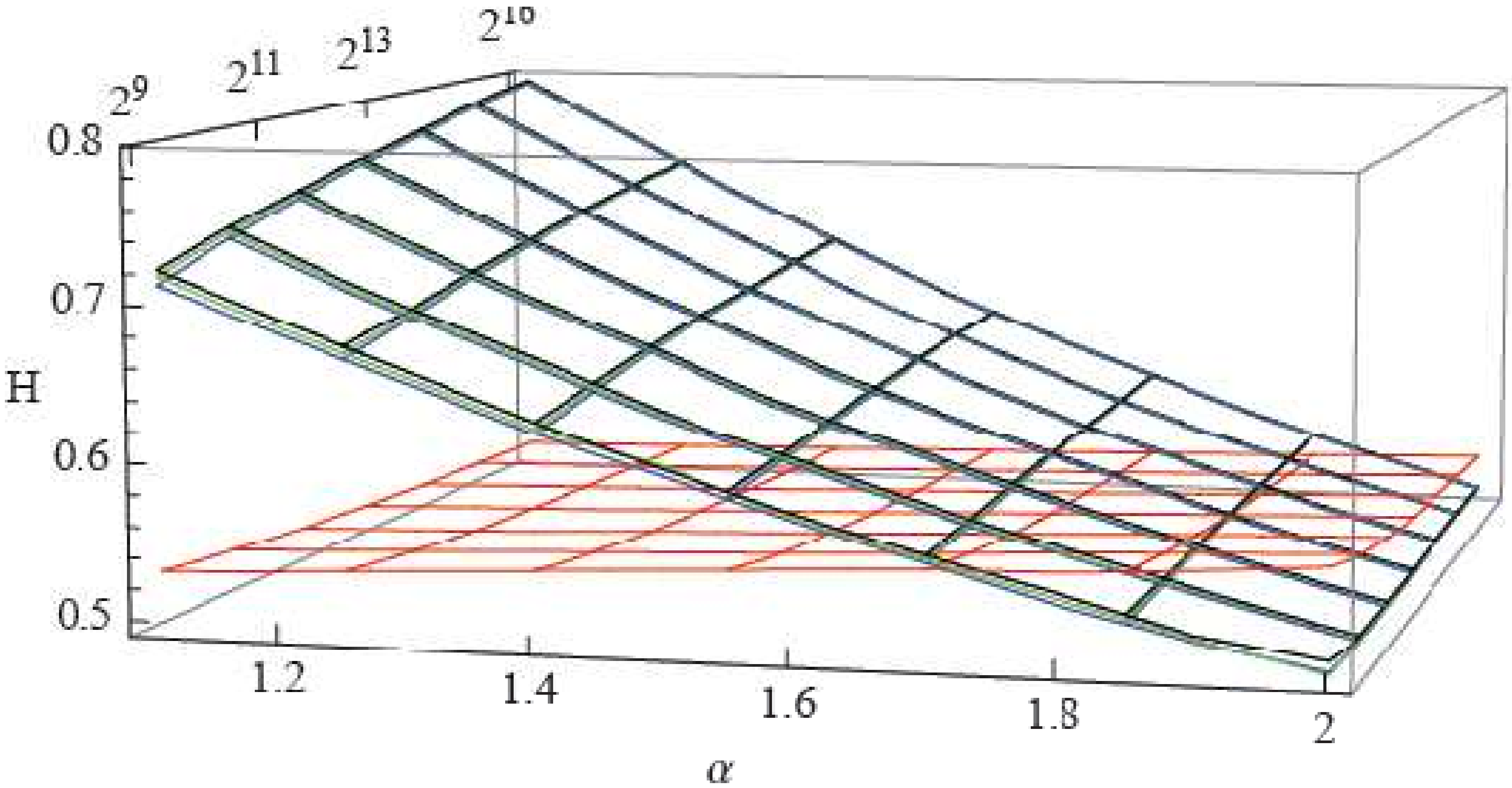
Monte Carlo simulations design (2)

- We set three out of four parameters of stable distribution and get $(\alpha,\beta,\gamma,\delta)=(\alpha,0,2^{0.5}/2,0)$ and change α while $1.1<\alpha<2$ with a step of 0.1
- For $\alpha=1.1$, we have the heaviest tails and for $\alpha=2$, we have standard normal distribution

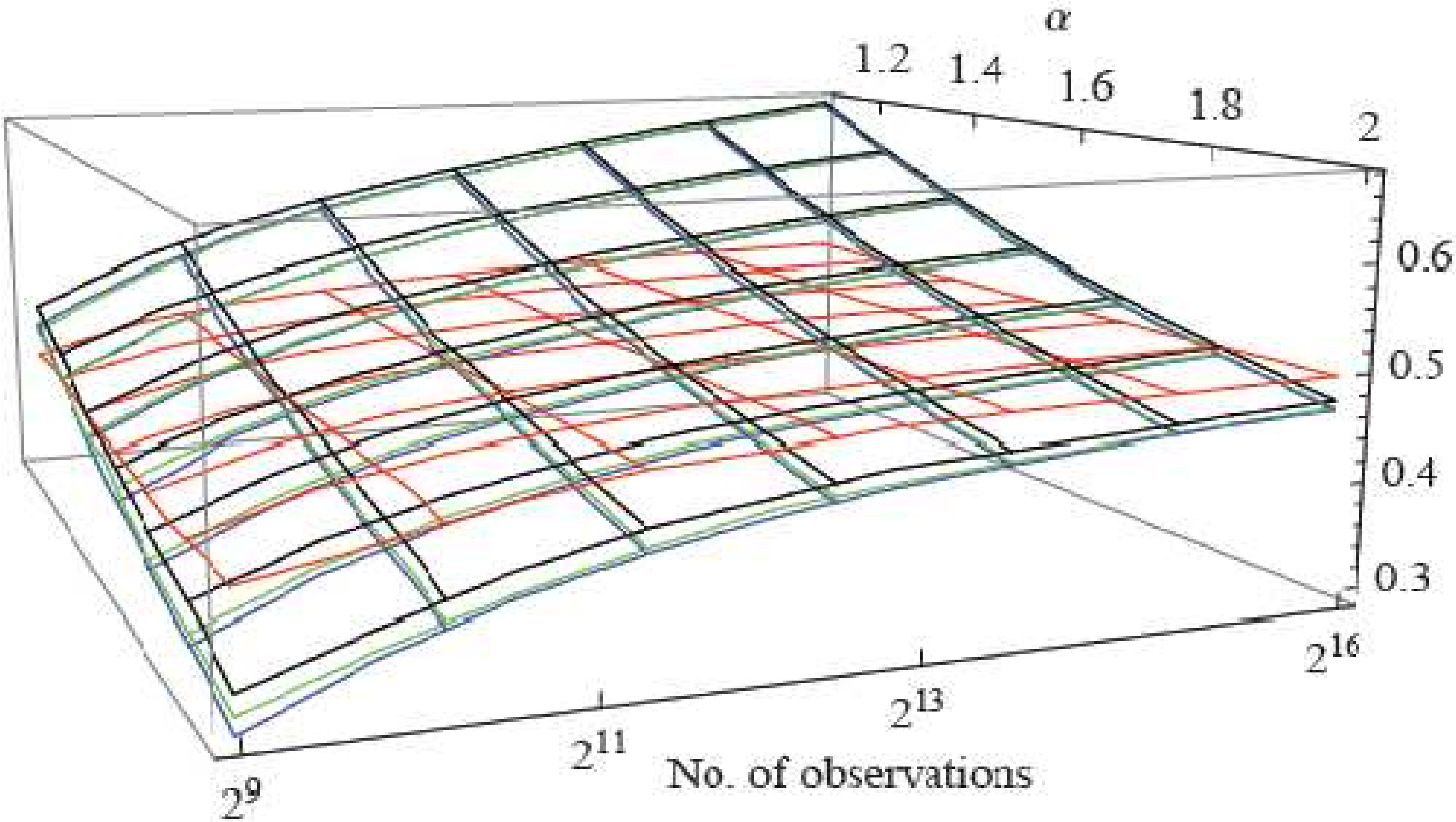
Monte Carlo simulations design (3)

- For each method, we use time series length from 2^9 to 2^{16}
- For each method, each time series length and each parameter α, we simulate 10 000 time series and estimate Hurst exponent for each

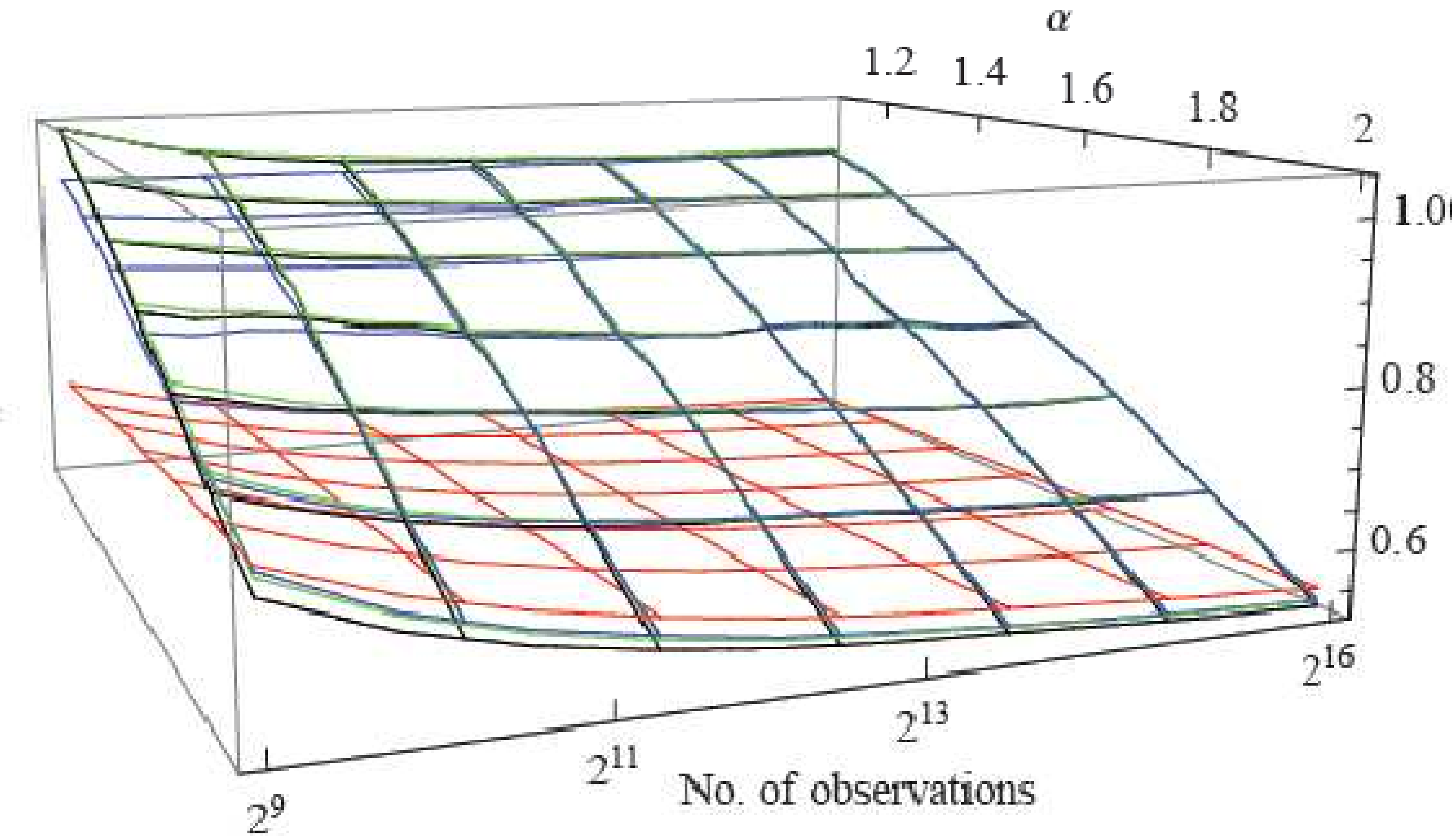
Results – mean values of *H*



Results – 2.5% quantiles



Results – 97.5% quantiles



Summary of results

- For normal distribution, DFA is the superior method
- For fat tails, DFA expected value diverges from *0.5* with decreasing α
- Also, for fat tails, DFA expected value diverges for increasing T
- R/S is robust against fat tails and the expected value even converges to *0.5* with fatter tails

Implications

- R/S is robust to fat tails and thus usable for financial time series
- DFA is not robust to fat tails and thus one must be cautious while using it for financial data
- The findings go against usual notion of strict superiority of DFA to R/S

Are there any questions?

Thank you.