Second Order Properties of Distribution Tails and Estimation of Tail Exponents in Random Difference Equations

Changryong Baek
University of North Carolina

Joint work with
Vladas Pipiras, Herwig Wendt and Patrice Abry

June 23, 2009

Graybill VIII - 6th International Conference on Extreme Value Analysis

Outline

1. Introduction
   - RDEs (Random Difference Equations)
   - Statement of problem
2. Theory
   - 2RV(2nd order Regular Variation)-like properties of RDEs
   - Extensions and open questions
3. Estimation
   - Methods
   - Comparison in simulations

Random Difference Equations (RDEs)

In one dimension, RDE is given by

$$X_n = A_n X_{n-1} + B_n, \quad X_0 = x, \quad n \geq 1,$$

where (A_n, B_n) are typically assumed to be i.i.d. vectors. Its stationary solution, also called RDE, satisfies

$$X \stackrel{d}{=} AX + B$$

with $(A, B) \stackrel{d}{=} (A_1, B_1)$ independent of X. Focus on $A, B, X > 0$.

Theorem (Kesten (1973))

Under mild assumptions on A and B, it is known that

$$P(X > x) \sim cx^{-\alpha},$$

for $\alpha > 0$ satisfying $EA^{\alpha} = 1$.

Examples of RDEs

- Beta prime distribution $\beta(a,b)$: The only known RDEs X with explicit power-law tail (tail exponent is b),

$$\beta(a,b) \stackrel{d}{=} \frac{1}{Beta(b,a)} - 1.$$

- (Squares of) ARCH(1) model in Finance

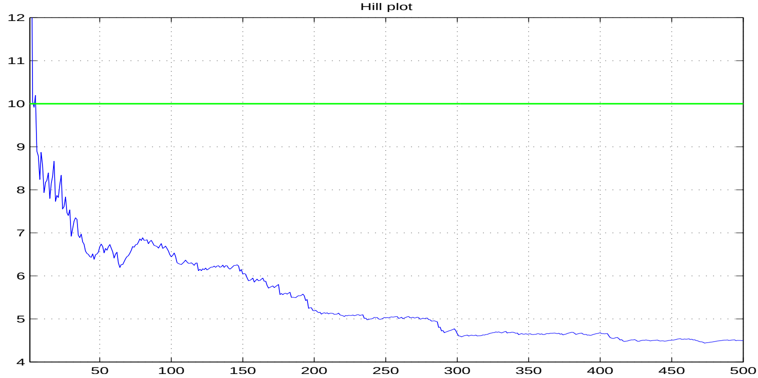
$$\xi_n^2 = \lambda\epsilon_n^2\xi_{n-1}^2 + \beta\epsilon_n^2.$$

- MC (Multiplicative Cascades) in Physics with larger tail exponent is expected in practice.

Statement of problem

Huge bias from Hill plots for larger tail exponent.

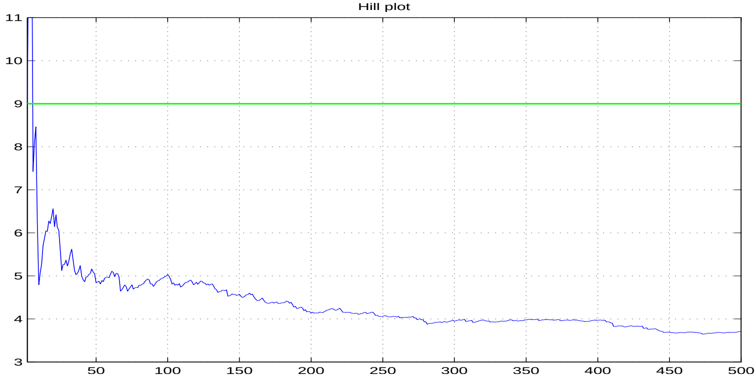
Hill plot for 5000 **independent** realizations of ARCH(1) model with $\alpha = 10$.



Statement of problem

Huge bias from Hill plots for larger tail exponent.

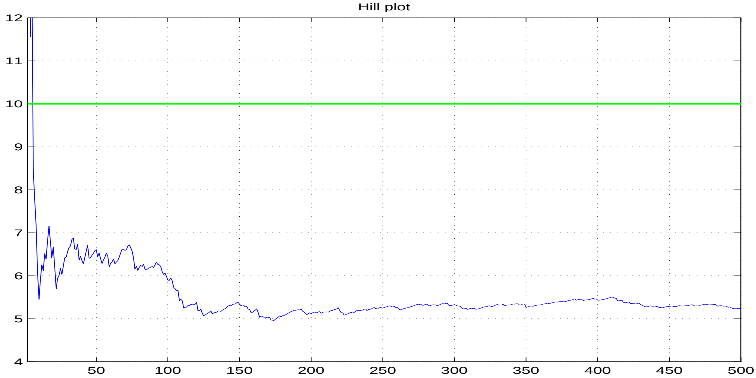
Hill plot for 5000 independent $\beta(5, 9)$ distribution with $\alpha = 9$.



Statement of problem

Huge bias from Hill plots for larger tail exponent.

Hill plot for 5000 independent realizations of MC with $\alpha = 10$.



Possible explanations

Because dependence is removed two explanations are possible:

- X is replaced by X_N in practice and the approximation is poor. This, however, is not the case.
- The distribution tail of X deviates from exact power-law tail.

Second order regular variation (2RV)

2RV is a natural framework to examine bias.
The distribution tail $P(X > x)$ is 2RV with first order parameter $\alpha > 0$ and second order parameter $\rho < 0$, if for suitable $g(x)$ and any $a > 0$,

$$\lim_{x \to \infty} \frac{x^{\alpha} P(X > x) - (ax)^{\alpha} P(X > ax)}{g(x)} = \frac{a^{\rho} - 1}{\rho}.$$

In practice, this is replaced by

$$P(X > x) = c_1 x^{-\alpha} + c_2 x^{-\alpha+\rho} + o(x^{-\alpha+\rho}).$$

2RV-like properties for RDEs

2RV for RDEs appears an open and difficult problem.
Nevertheless, one can show:

Theorem 1 (weaker form of 2RV for RDEs)

Under mild assumptions on A, B and X, one has

$$\int_x^\infty (P(X>u) - P(AX>u))\, du \sim Cx^{-\alpha} \qquad (\bigstar)$$

Remark. This result suggests that

$$P(X>x) - P(AX>x) \sim C\alpha x^{-\alpha-1}$$

or

$$\lim_{x\to\infty} \frac{x^\alpha P(X>x) - x^\alpha (EA^\alpha)^{-1} P(X>A^{-1}x)}{x^{-1}} = C\alpha,$$

which can be viewed as 2RV at random $a = 1/A$.

2RV-like properties for RDEs

Remark. Theorem 1 is proved just by using Kesten's result and the form of RDEs.

On a more practical side, one has:

Theorem 2 (important relation to practical 2RV)

Under mild assumptions on X, A and B, the result $(\bigstar)$ in Theorem 1 is consistent with

$$P(X > x) = c_1 x^{-\alpha} + c_2 x^{-\alpha+\rho} + o(x^{-\alpha+\rho}),$$

only when

$$\boxed{\rho = -1.}$$

Extensions and open questions

Extensions:

- RDEs with real-valued coefficients.
- Multidimensional RDEs with nonnegative matrix coefficients. Examples include GARCH model.

Open questions:

- Exact 2RV.
- Relevance in practice.

Estimation with known ρ

The results above suggest that for RDE X

$$P(X > x) = c_1 x^{-\alpha} + c_2 x^{-\alpha+\rho} + o(x^{-\alpha+\rho})$$

with $\rho = -1$.

- Large literature on estimation when ρ is *unknown*. Feureverger and Hall (1999), Beirlant, Dierckx, Goegebeur and Matthys (1999), Beirlant, Dierckx, Guillou and Stărică (2002), Gomes and Martins (2002) to name a few.
- For known ρ, what would be the best estimator?

Number of possible estimators

RK2 Rank-based least squares estimator based on

$$\log P(X > x) \approx \log c_1 - \alpha \log x + c_2/c_1 x^{-1}$$

Replacing x by order statistic $X_{(n-i+1)}$ gives

$$\log(i/n) \approx \log c_1 - \alpha \log X_{(n-i+1)} + c_2/c_1 X_{(n-i+1)}^{-1}.$$

QQ2 Reversing the role of $\log P(X > x)$ and $\log x$.

JK, P Generalized jackknife estimators adapting Gomes, Martins and Neves (2000) and Peng (1998).

FH Feuerverger and Hall (1999) estimator with known ρ.

ML Maximum likelihood estimator based on the exact tail $P(X > x) = c_1 x^{-\alpha} + c_2 x^{-\alpha-1}$.

Properties of proposed estimators

Theorem 3 (asymptotic normality)

Under suitable assumptions,

$$\sqrt{k}\,(\widehat{\alpha} - \alpha) \xrightarrow{d} \mathcal{N}(0, \sigma^2),$$

where k is the number of upper order statistics used and asymptotic variance

$$\sigma^2_{RK2} = \sigma^2_{QQ2} = 2\alpha^2(\alpha + 1)^2,$$

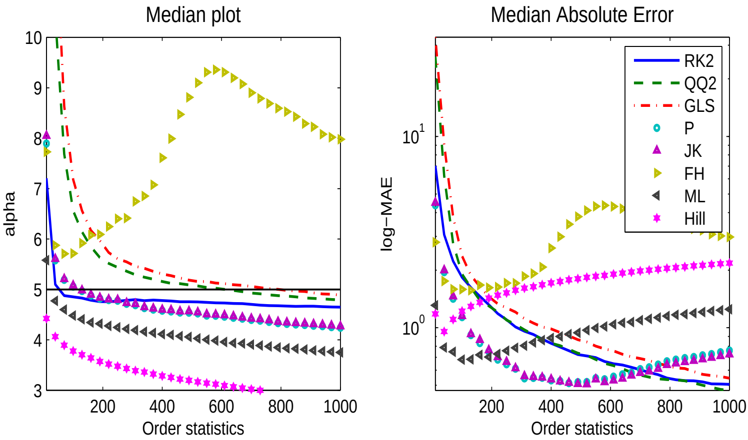
$$\sigma^2_{JK} = \sigma^2_{P} = \alpha^2((\alpha + 1)^2 + \alpha^2),$$

$$\sigma^2_{FH} = \sigma^2_{ML} = \alpha^2(\alpha + 1)^2.$$

In simulations: proposed estimators successfully provide bias correction. RK2 estimator appears to perform best in practice.

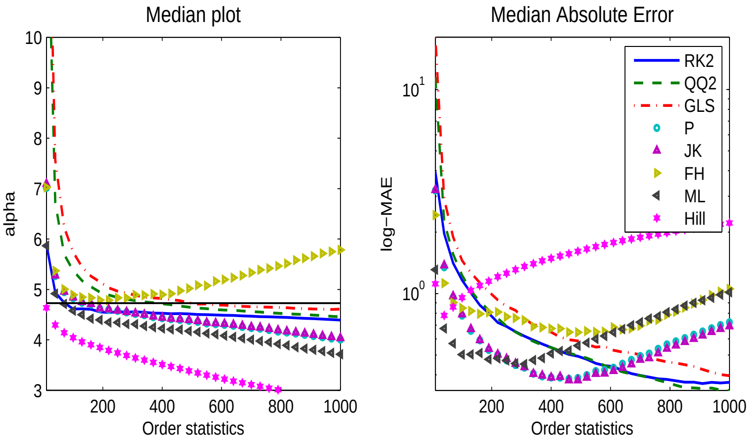
Comparison in simulations

Beta prime distribution with $\alpha = 5,\ T = 5,000$.



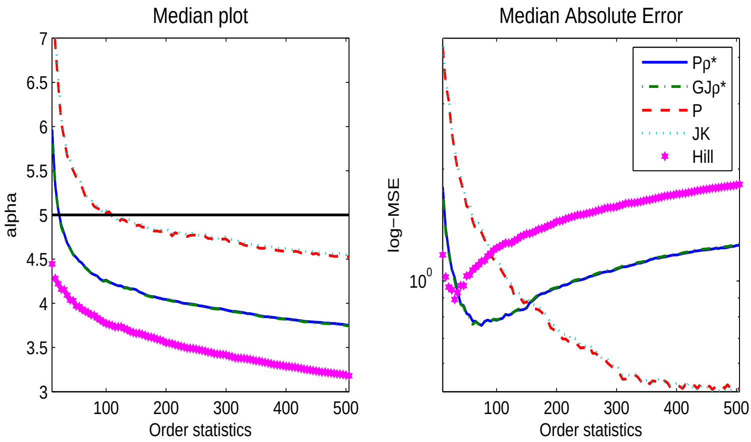
Comparison in simulations

ARCH(1) model with $\alpha = 4.73$, $T = 5,000$.



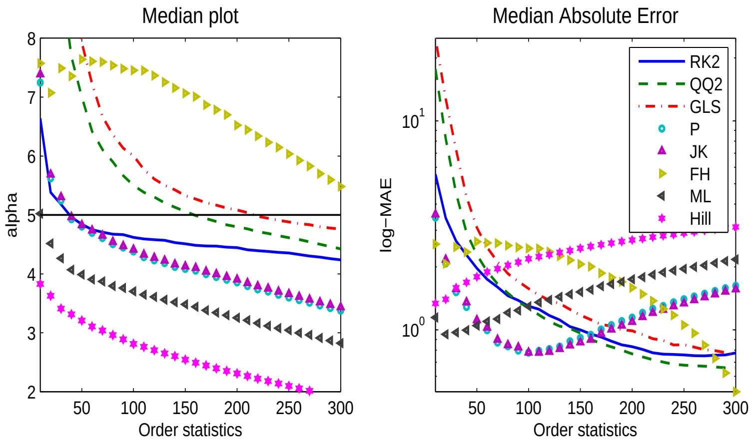
Comparison in simulations

Beta prime distribution with $\alpha = 5,\ T = 5,000$.



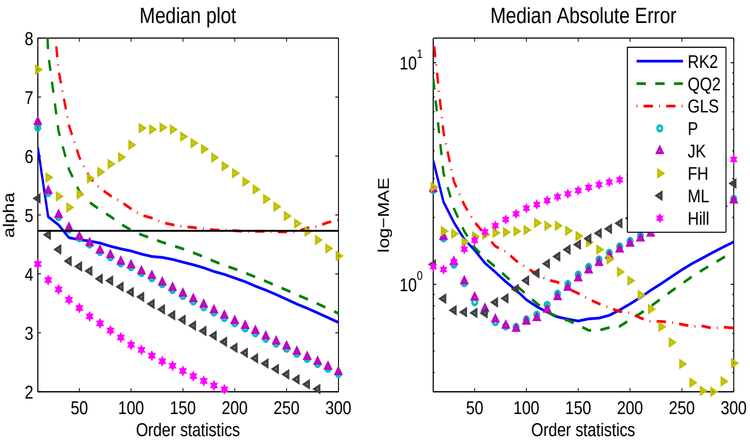
Comparison in simulations

Beta prime distribution with $\alpha = 5$, $T = 500$.



Comparison in simulations

ARCH(1) model with $\alpha = 4.73$, $T = 500$.



References

1. Baek, C., Pipiras, V., Wendt, H. & Abry, P. (2009), ‘Second order properties of distribution tails and estimation of tail exponents in random difference equations’, To appear in Extremes.
2. Baek, C. & Pipiras, V. (2009), ‘Estimation of parameters in heavy-tailed distribution when its second order tail parameter is known’, Preprint.

Thank you!