Regular Variation and Financial Time Series Models

Richard A. Davis
Colorado State University
www.stat.colostate.edu/~rdavis

Thomas Mikosch
University of Copenhagen

Bojan Basrak
Eurandom
Characteristics of some financial time series

- IBM returns
- Multiplicative models for log-returns (GARCH, SV)

Regular variation

- univariate case
- multivariate case
- new characterization: $X$ is RV $\iff c'X$ is RV?

Applications of regular variation

- Stochastic recurrence equations (GARCH)
- Point process convergence
- Extremes and extremal index
- Limit behavior of sample correlations

Wrap-up
Define \( X_t = \ln (P_t) - \ln (P_{t-1}) \) (log returns)

- heavy tailed

\[
P(|X_1| > x) \sim C x^{-\alpha}, \quad 0 < \alpha < 4.
\]

- uncorrelated

\[
\hat{\rho}_X(h) \text{ near 0 for all lags } h > 0 \text{ (MGD sequence)}
\]

- \( |X_t| \) and \( X_t^2 \) have slowly decaying autocorrelations

\[
\hat{\rho}_{|X|}(h) \text{ and } \hat{\rho}_{X^2}(h) \text{ converge to 0 slowly as } h \text{ increases.}
\]

- process exhibits ‘volatility clustering’.
Sample ACF IBM (a) 1962-1981, (b) 1982-2000

(a) ACF of IBM (1st half)

(b) ACF of IBM (2nd half)
Sample ACF of abs values for IBM (a) 1961-1981, (b) 1982-2000
Multiplicative models for log(returns)

**Basic model**

\[ X_t = \ln (P_t) - \ln (P_{t-1}) \quad \text{(log returns)} \]

\[ = \sigma_t Z_t, \]

where

- \( \{Z_t\} \) is IID with mean 0, variance 1 (if exists). (e.g. \( N(0,1) \) or a \( t \)-distribution with \( \nu \) df.)
- \( \{\sigma_t\} \) is the volatility process
- \( \sigma_t \) and \( Z_t \) are independent.

**Properties:**

- \( E X_t = 0, \text{Cov}(X_t, X_{t+h}) = 0, h>0 \) (uncorrelated if \( \text{Var}(X_t) < \infty \))
- conditional heteroscedastic (condition on \( \sigma_t \)).
Multiplicative models for log(returns)-cont

\[ X_t = \sigma_t Z_t \] (observation eqn in state-space formulation)

Two classes of models for volatility:

(i) GARCH(p,q) process (General AutoRegressive Conditional Heteroscedastic-observation-driven specification)

\[ \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \cdots + \alpha_p X_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_q \sigma_{t-q}^2 . \]

Special case: ARCH(1):

\[ X_t^2 = (\alpha_0 + \alpha_1 X_{t-1}^2) Z_t^2 \]
\[ = \alpha_1 Z_t^2 X_{t-1}^2 + \alpha_0 Z_t^2 \]
\[ = A_t X_{t-1}^2 + B_t \] (stochastic recurrence eqn)

\[ \rho_{X^2}(h) = \alpha_1^h , \text{ if } \alpha_1^2 < 1/3. \]
Multiplicative models for log(returns)-cont

GARCH(2,1): \( X_t = \sigma_t Z_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \alpha_2 X_{t-2}^2 + \beta_1 \sigma_{t-1}^2 \).

Then \( Y_t = (\sigma_t^2, X_{t-1}^2) \)' follows the SRE given by

\[
\begin{bmatrix}
\sigma_t^2 \\
X_{t-1}^2
\end{bmatrix}
= \begin{bmatrix}
\alpha_1 & \beta_1 \\
\alpha_2 & 0
\end{bmatrix}
\begin{bmatrix}
\sigma_{t-1}^2 \\
X_{t-2}^2
\end{bmatrix}
+ \begin{bmatrix}
\alpha_0 \\
0
\end{bmatrix}
\]

Questions:

• Existence of a unique stationary solution to the SRE?
• Regular variation of the joint distributions?
Multiplicative models for log(returns)-cont

\[ X_t = \sigma_t Z_t \]  (observation eqn in state-space formulation)

(ii) stochastic volatility process (parameter-driven specification)

\[
\log \sigma_t^2 = \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j}, \quad \sum_{j=-\infty}^{\infty} \psi_j^2 < \infty, \{\varepsilon_t\} \sim \text{IIDN}(0, \sigma^2)
\]

\[
\rho_{\chi^2}(h) = \text{Cor}(\sigma_t^2, \sigma_{t+h}^2) / EZ_1^4
\]

Question:

• Joint distributions of process regularly varying if distr of \( Z_1 \) is regularly varying?
**Def:** The random variable \( X \) is *regularly varying with index* \( \alpha \) if

\[
P(\lvert X \rvert > tx)/P(\lvert X \rvert > t) \to x^{-\alpha} \quad \text{and} \quad P(X > t)/P(\lvert X \rvert > t) \to p,
\]

or, equivalently, if

\[
P(X > tx)/P(\lvert X \rvert > t) \to px^{-\alpha} \quad \text{and} \quad P(X < -t x)/P(\lvert X \rvert > t) \to qx^{-\alpha},
\]

where \( 0 \leq p \leq 1 \) and \( p+q=1 \).

**Equivalence:**

\( X \) is \( \text{RV(}\alpha\text{)} \) if and only if \( P(X \in t \bullet) /P(\lvert X \rvert > t) \to_v \mu(\bullet) \)

(\( \to_v \) vague convergence of measures on \( \mathbb{R}\setminus\{0\} \)). In this case,

\[
\mu(dx) = \left(p\alpha x^{-\alpha-1} I(x>0) + q\alpha (-x)^{-\alpha-1} I(x<0)\right) dx
\]

**Note:** \( \mu(tA) = t^{-\alpha} \mu(A) \) for every \( t \) and \( A \) bounded away from 0.
Another formulation (polar coordinates):

Define the $\pm 1$ valued rv $\theta$, $P(\theta = 1) = p$, $P(\theta = -1) = 1 - p = q$.

Then

$X$ is $RV(\alpha)$ if and only if

$$\frac{P(|X| > t, X/|X| \in S)}{P(|X| > t)} \rightarrow x^{-\alpha} P(\theta \in S)$$

or

$$\frac{P(|X| > t, X/|X| \in \bullet)}{P(|X| > t)} \rightarrow_{v} x^{-\alpha} P(\theta \in \bullet)$$

($\rightarrow_{v}$ vague convergence of measures on $S^0 = \{-1, 1\}$).
Regular variation — multivariate case

Multivariate regular variation of \( X=(X_1, \ldots, X_m) \): There exists a random vector \( \theta \in S^{m-1} \) such that

\[
P(|X|> t \, x, \ X/|X| \in \bullet )/P(|X|>t) \to_v x^{-\alpha} \, P( \theta \in \bullet )
\]

\( \to_v \) vague convergence on \( S^{m-1} \), unit sphere in \( R^m \).

- \( P( \theta \in \bullet ) \) is called the spectral measure
- \( \alpha \) is the index of \( X \).

Equivalence:

\[
\frac{P( X \in t \bullet )}{P(|X|> t)} \to_v \mu(\bullet)
\]

\( \mu \) is a measure on \( R^m \) which satisfies for \( x > 0 \) and \( A \) bounded away from 0,

\[
\mu(xB) = x^{-\alpha} \mu(xA).
\]
Examples:

1. If $X_1 > 0$ and $X_2 > 0$ are iid RV($\alpha$), then $X = (X_1, X_2)$ is multivariate regularly varying with index $\alpha$ and *spectral distribution*

   $$P(\theta = (0,1)) = P(\theta = (1,0)) = .5$$ (mass on axes).

Interpretation: Unlikely that $X_1$ and $X_2$ are very large at the same time.

**Figure:** plot of $(X_{t1}, X_{t2})$ for realization of 10,000.
2. If $X_1 = X_2 > 0$, then $X = (X_1, X_2)$ is multivariate regularly varying with index $\alpha$ and *spectral distribution*
\[ P(\theta = (1/\sqrt{2}, 1/\sqrt{2}) ) = 1. \]

3. AR(1): $X_t = 0.9 X_{t-1} + Z_t$, $\{Z_t\}$~IID symmetric stable (1.8)

Distr of $\theta$: \[
\begin{align*}
\pm(1,0.9)/\sqrt{1.81} &\text{, W.P. } 0.9898 \\
\pm(0,1) &\text{, W.P. } 0.0102
\end{align*}
\]

**Figure:** plot of $(X_t, X_{t+1})$ for realization of 10,000.
Applications of multivariate regular variation

• Domain of attraction for *sums of iid random vectors* (Rvaceva, 1962). That is, when does the partial sum

\[ a_n^{-1} \sum_{t=1}^{n} X_t \]

converge for some constants \( a_n \)?

• *Spectral measure* of multivariate stable vectors.

• *Domain of attraction* for componentwise maxima of iid random vectors (Resnick, 1987). Limit behavior of

\[ a_n^{-1} \vee_{t=1}^{n} X_t \]

• Weak convergence of *point processes* with iid points.

• Solution to *stochastic recurrence equations*, \( Y_t = A_t Y_{t-1} + B_t \)

• Weak convergence of *sample autocovariances*. 

**Products** (Breiman 1965). Suppose $X, Y > 0$ are independent with $X \sim RV(\alpha)$ and $EY^{\alpha+\varepsilon} < \infty$ for some $\varepsilon > 0$. Then $XY \sim RV(\alpha)$ with

$$P(XY > x) \sim EY^{\alpha} P(X > x).$$

**Multivariate version.** Suppose the random vector $X$ is regularly varying and $A$ is a matrix independent of $X$ with

$$0 < E\|A\|^{\alpha+\varepsilon} < \infty.$$

Then

$AX$ is regularly varying with index $\alpha$. 
Linear combinations:

\( X \sim RV(\alpha) \Rightarrow \) all linear combinations of \( X \) are regularly varying

i.e., there exist \( \alpha \) and slowly varying fcn \( L(.) \), s.t.

\[
P(c^T X > t) / (t^{-\alpha} L(t)) \rightarrow w(c), \text{ exists for all real-valued } c,
\]

where

\[
w(tc) = t^{-\alpha} w(c).
\]

Use vague convergence with \( A_c = \{ y : c^T y > 1 \} \), i.e.,

\[
\frac{P(X \in tA_c)}{t^{-\alpha} L(t)} = \frac{P(c^T X > t)}{P(|X| > t)} \rightarrow \mu(A_c) = w(c),
\]

where \( t^{-\alpha} L(t) = P(|X| > t) \).
Applications of multivariate regular variation (cont)

Converse?

\( \mathbf{X} \sim \text{RV}(\alpha) \iff \text{all linear combinations of } \mathbf{X} \text{ are regularly varying?} \)

\[ \text{There exist } \alpha \text{ and slowly varying fcn } L(.), \text{ s.t.} \]
\[ (\text{LC}) \quad P(\mathbf{c}^\top \mathbf{X} > t)/(t^{-\alpha} L(t)) \to w(\mathbf{c}), \text{ exists for all real-valued } \mathbf{c}. \]

Theorem (Basrak, Davis, Mikosch, `02). Let \( \mathbf{X} \) be a random vector.

1. If \( \mathbf{X} \) satisfies (LC) with \( \alpha \) non-integer, then \( \mathbf{X} \) is \( \text{RV}(\alpha) \).

2. If \( \mathbf{X} > 0 \) satisfies (LC) for non-negative \( \mathbf{c} \) and \( \alpha \) is non-integer, then \( \mathbf{X} \) is \( \text{RV}(\alpha) \).

3. If \( \mathbf{X} > 0 \) satisfies (LC) with \( \alpha \) an odd integer, then \( \mathbf{X} \) is \( \text{RV}(\alpha) \).
Idea of argument: Define the measures

\[ m_t(\cdot) = \frac{P(X \in t \cdot)}{(t^\alpha L(t))} \]

• By assumption we know that for fixed \( c \), \( m_t(A_c) \to \mu(A_c) \).

• \( \{m_t\} \) is tight: For \( B \) bded away from 0, \( \sup_t m_t(B) < \infty \).

• Do subsequential limits of \( \{m_t\} \) coincide?

If \( m_t' \to \nu \mu_1 \) and \( m_t'' \to \nu \mu_2 \), then

\[ \mu_1(A_c) = \mu_2(A_c) \text{ for all } c \neq 0. \]

Problem: Need \( \mu_1 = \mu_2 \) but only have equality on \( A_c \), not a \( \pi \)-system.

In general, equality need not hold (see Ex 6.1.35 in Meerschaert & Scheffler (2001)).
1. Kesten (1973). Under general conditions, (LC) holds with \( L(t) = 1 \) for stochastic recurrence equations of the form

\[ Y_t = A_t Y_{t-1} + B_t, \quad (A_t, B_t) \sim \text{IID}, \]

\[ A_t \text{ } d \times d \text{ random matrices, } B_t \text{ random } d\text{-vectors.} \]

It follows that the distributions of \( Y_t \), and in fact all of the finite dim’l distrs of \( Y_t \) are regularly varying (if \( \alpha \) is non-even).

2. GARCH processes. Since squares of a GARCH process can be embedded in a SRE, the \textit{finite dimensional distributions} of a \textit{GARCH} are regularly varying.
Example of ARCH(1): \( X_t = (\alpha_0 + \alpha_1 X_{t-1}^2)^{1/2} Z_t, \{Z_t\} \sim \text{IID} \).

\( \alpha \) found by solving \( E|\alpha_1 Z^2|^{\alpha/2} = 1 \).

\[
\begin{array}{c|cccc}
\alpha_1 & .312 & .577 & 1.00 & 1.57 \\
\alpha & 8.00 & 4.00 & 2.00 & 1.00 \\
\end{array}
\]

Distr of \( \theta \):

\[
P(\theta \in \bullet) = E\{||B,Z||^{\alpha} I(\text{arg}((B,Z)) \in \bullet)\} / E||B,Z||^{\alpha}
\]

where

\[
P(B = 1) = P(B = -1) = .5
\]
Example of ARCH(1): $\alpha_0=1, \alpha_1=1, \alpha=2$, $X_t=(\alpha_0+\alpha_1 X_{t-1}^2)^{1/2}Z_t$, $\{Z_t\} \sim \text{IID}$

Figures: plots of $(X_t, X_{t+1})$ and estimated distribution of $\theta$ for realization of 10,000.
Example: SV model $X_t = \sigma_t Z_t$

Suppose $Z_t \sim \text{RV}(\alpha)$ and

$$
\log \sigma_t^2 = \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j}, \quad \sum_{j=-\infty}^{\infty} \psi_j^2 < \infty, \{\varepsilon_t\} \sim \text{IIDN}(0,\sigma^2).
$$

Then $Z_n=(Z_1,\ldots,Z_n)'$ is regularly varying with index $\alpha$ and so is

$$
X_n = (X_1,\ldots,X_n)' = \text{diag}(\sigma_1,\ldots,\sigma_n) Z_n
$$

with spectral distribution concentrated on $(\pm 1,0), (0, \pm 1)$.

---

**Figure:** plot of $(X_t, X_{t+1})$ for realization of 10,000.
**Theorem** Let \( \{X_i\} \) be an iid sequence of random vectors satisfying 1 of the 3 conditions in the theorem. Then

\[
N_n := \sum_{t=1}^{n} \mathcal{E}_{X_t/a_n} \xrightarrow{d} N := \sum_{j=1}^{\infty} \mathcal{E}_{P_{i\theta_i}},
\]

if and only if for every \( c \neq 0 \)

\[
N_{n,c} := \sum_{t=1}^{n} \mathcal{E}_{cX_t/a_n} \xrightarrow{d} N_c := \sum_{j=1}^{\infty} \mathcal{E}_{cP_{i\theta_i}},
\]

where \( \{a_n\} \) satisfies \( nP(|X_i| > a_n) \to 1 \), and \( N \) is a Poisson process with intensity measure \( \mu \).

- \( \{P_i\} \) are Poisson pts corresponding to the radial part, i.e., has intensity measure \( \alpha x^{-\alpha-1} \, (dx) \).
- \( \{\theta_i\} \) are iid with the spectral distribution given by the RV
Theorem (Davis & Hsing `95, Davis & Mikosch `97). Let \( \{X_t\} \) be a stationary sequence of random \( m \)-vectors. Suppose

(i) finite dimensional distributions are jointly regularly varying (let \((\theta_k, \ldots, \theta_k)\) be the vector in \( S^{(2k+1)m-1} \) in the definition).

(ii) mixing condition \( A(a_n) \) or strong mixing.

(iii) \( \lim \limsup_{k \to \infty} P\left( \bigvee_{k \leq |i| \leq n} \left| X_i \right| > a_n y \left| X_0 \right| > a_n y \right) = 0. \)

Then

\[
\gamma = \lim_{k \to \infty} E\left( \left| \theta_0^{(k)} \right|^\alpha - \bigvee_{j=1}^k \left| \theta_j^{(k)} \right| \right)_+ / E \left| \theta_0^{(k)} \right|^\alpha
\]

(extramal index)

exists. If \( \gamma > 0 \), then

\[
N_n := \sum_{t=1}^n \varepsilon X_t / a_n \xrightarrow{d} N := \sum_{i=1}^\infty \sum_{j=1}^\infty \varepsilon_{PQ_{ij}},
\]
Point process convergence (cont)

- \( (\mathcal{P}_i) \) are points of a Poisson process on \((0, \infty)\) with intensity function
  \[ \nu(dy) = \gamma \alpha y^{-\alpha - 1} dy. \]

- \( \sum_{j=1}^{\infty} \varepsilon_{Q_j}, \ i \geq 1, \) are iid point process with distribution \( Q, \) and \( Q \) is the weak limit of

\[
\lim_{k \to \infty} E\left( \left| \theta^{(k)}_0 \right|^{\alpha} - \sum_{j=1}^{k} \left| \theta^{(k)}_j \right| \right) + I_\ast \left( \sum_{|j| \leq k} \varepsilon_{\theta^{(k)}_j} \right) / E\left( \left| \theta^{(k)}_0 \right|^{\alpha} - \sum_{j=1}^{k} \left| \theta^{(k)}_j \right| \right)_{+}.
\]

Remarks:

1. GARCH and SV processes satisfy the conditions of the theorem.

2. Limit distribution for sample extremes and sample ACF follows from this theorem.
Extremes for GARCH and SV processes

Setup

- $X_t = \sigma_t Z_t$, \quad \{Z_t\} \sim \text{IID (0,1)}$
- $X_t$ is RV $(\alpha)$
- Choose $\{b_n\}$ s.t. $nP(X_t > b_n) \to 1$

Then

$$nP\left( b_n^{-1} X_1 \leq x \right) \to \exp \{-x^{-\alpha}\}.$$ 

Then, with $M_n = \max\{X_1, \ldots, X_n\}$,

(i) GARCH:

$$P(b_n^{-1} M_n \leq x) \to \exp \{-\gamma x^{-\alpha}\},$$

$\gamma$ is extremal index (0 < $\gamma$ < 1).

(ii) SV model:

$$P(b_n^{-1} M_n \leq x) \to \exp \{-x^{-\alpha}\},$$

extremal index $\gamma = 1$ no clustering.
Extremes for GARCH and SV processes (cont)

(i) GARCH: \( P(b_n^{-1}M_n \leq x) \rightarrow \exp\{-\gamma x^{-\alpha}\} \)

(ii) SV model: \( P(b_n^{-1}M_n \leq x) \rightarrow \exp\{-x^{-\alpha}\} \)

Remarks about extremal index.

(i) \( \gamma < 1 \) implies clustering of exceedances

(ii) Numerical example. Suppose \( c \) is a threshold such that \( P^n(b_n^{-1}X_1 \leq c) \sim .95 \)

Then, if \( \gamma = .5 \), \( P(b_n^{-1}M_n \leq c) \sim (.95)^5 = .975 \)

(iii) \( 1/\gamma \) is the mean cluster size of exceedances.

(iv) Use \( \gamma \) to discriminate between GARCH and SV models.

(v) Even for the light-tailed SV model (i.e., \( \{Z_t\} \sim\text{IID N}(0,1) \), the extremal index is 1 (see Breidt and Davis `98)
Extremes for GARCH and SV processes (cont)
Summary of results for ACF of GARCH(p,q) and SV models

**GARCH(p,q)**

\(\alpha \in (0,2)\):

\[
(\hat{\rho}_x(h))_{h=1,\ldots,m} \xrightarrow{d} (V_h/V_0)_{h=1,\ldots,m},
\]

\(\alpha \in (2,4)\):

\[
\left(n^{1-2/\alpha} \hat{\rho}_x(h)\right)_{h=1,\ldots,m} \xrightarrow{d} \gamma^{-1}_X(0)(V_h)_{h=1,\ldots,m}.
\]

\(\alpha \in (4,\infty)\):

\[
\left(n^{1/2} \hat{\rho}_x(h)\right)_{h=1,\ldots,m} \xrightarrow{d} \gamma^{-1}_X(0)(G_h)_{h=1,\ldots,m}.
\]

**Remark:** Similar results hold for the sample ACF based on \(|X_t|\) and \(X_t^2\).
Summary of results for ACF of GARCH(p,q) and SV models (cont)

**SV Model**

$\alpha \in (0, 2)$:

$$\left( \frac{n}{\ln n} \right)^{1/\alpha} \hat{\rho}_X(h) \xrightarrow{d} \frac{\|\sigma_1 \sigma_{h+1}\|_\alpha}{\|\sigma_1\|_\alpha^2} \frac{S_h}{S_0}.$$  

$\alpha \in (2, \infty)$:

$$\left( n^{1/2} \hat{\rho}_X(h) \right)_{h=1, \ldots, m} \xrightarrow{d} \gamma_X^{-1}(0)(G_h)_{h=1, \ldots, m}.$$
Sample ACF for GARCH and SV Models (1000 reps)

(a) GARCH(1,1) Model, n=10000

(b) SV Model, n=10000
Sample ACF for Squares of GARCH (1000 reps)

(a) GARCH(1,1) Model, n=10000

(b) GARCH(1,1) Model, n=100000
Sample ACF for Squares of SV (1000 reps)

(c) SV Model, n=10000

(d) SV Model, n=100000
Wrap-up

• *Regular variation* is a flexible tool for modeling both *dependence* and *tail heaviness*.

• Useful for establishing *point process convergence* of heavy-tailed time series.

• *Extremal index* $\gamma < 1$ for GARCH and $\gamma = 1$ for SV.

Unresolved issues related to $RV \Leftrightarrow (LC)$

• $\alpha = 2n$?

• there is an example for which $X_1, X_2 > 0$, and $(c, X_1)$ and $(c, X_2)$ have the same limits for all $c > 0$.

• $\alpha = 2n - 1$ and $X \not > 0$ (not true in general).