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

Outline

☞ 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 ⇔ 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

Characteristics of some financial time series

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

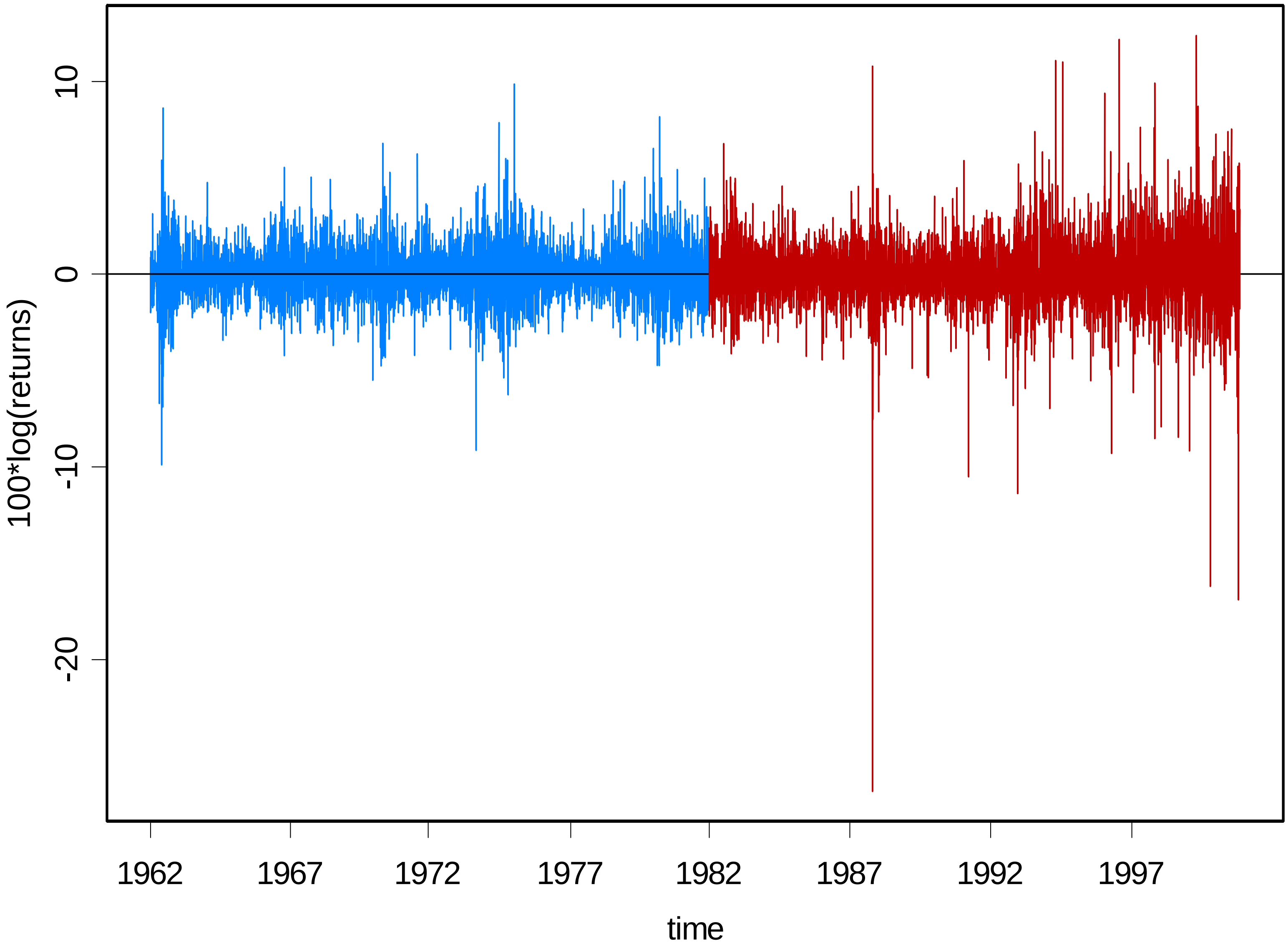
$$\hat{\rho}_X(h) \text{ near } 0 \text{ for all lags } h > 0 \text{ (MGD sequence)}$$

- $|X_t|$ and X_t^2 have slowly decaying autocorrelations

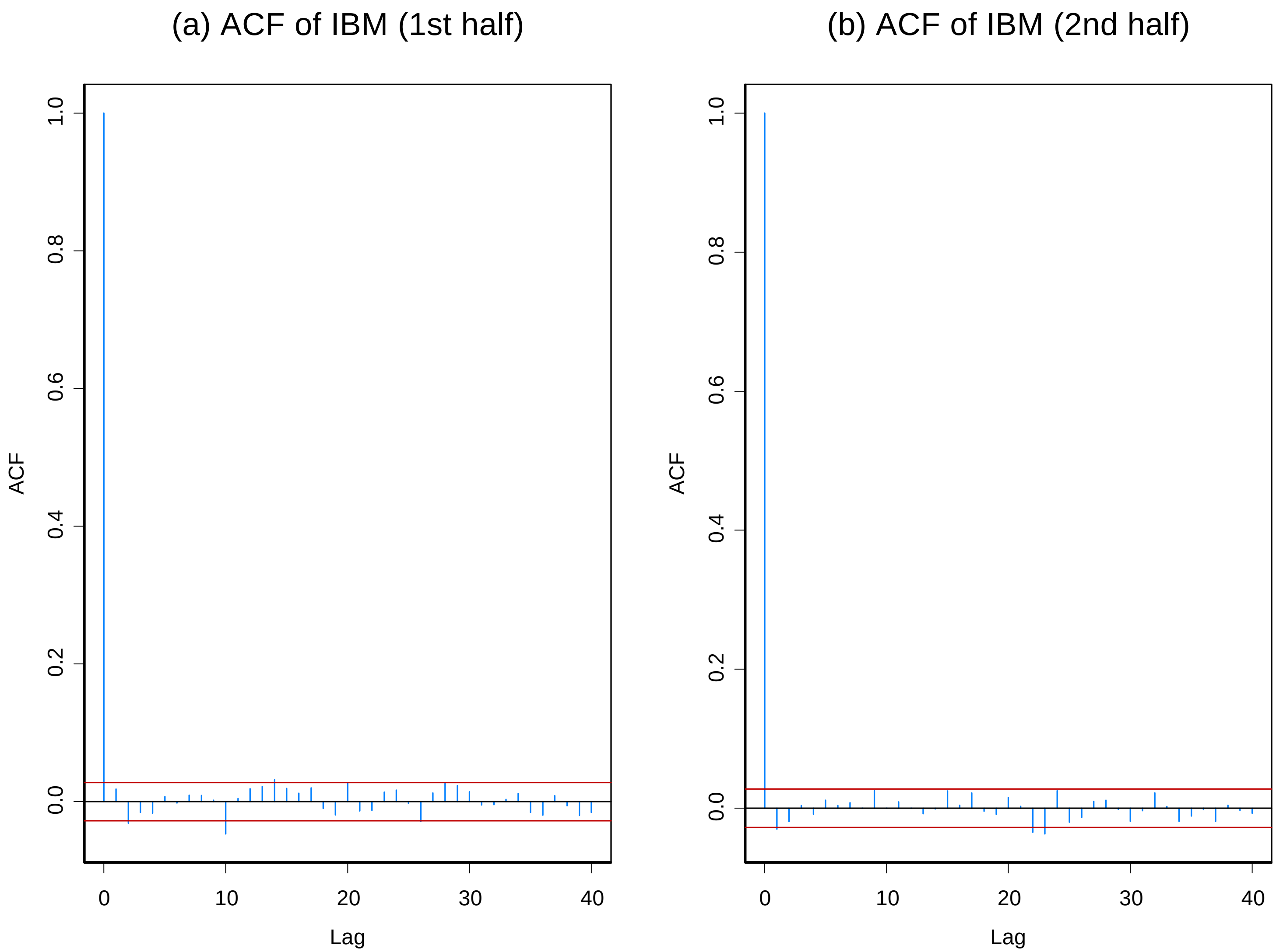
$\hat{\rho}_{|X|}(h)$ and $\hat{\rho}_{X^2}(h)$ converge to 0 *slowly* as h increases.

- process exhibits 'volatility clustering'.

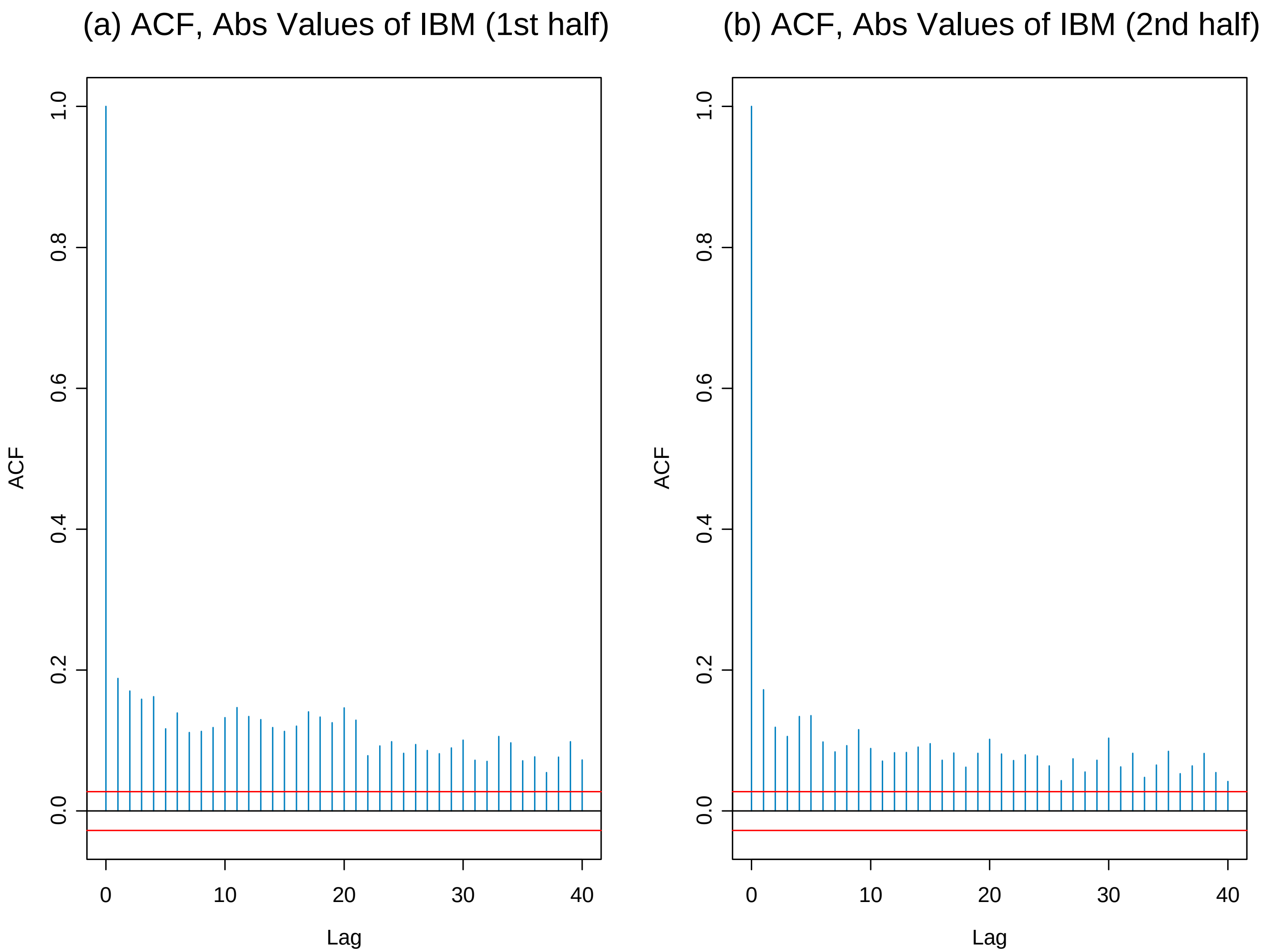
Log returns for IBM 1/3/62-11/3/00 (blue=1961-1981)



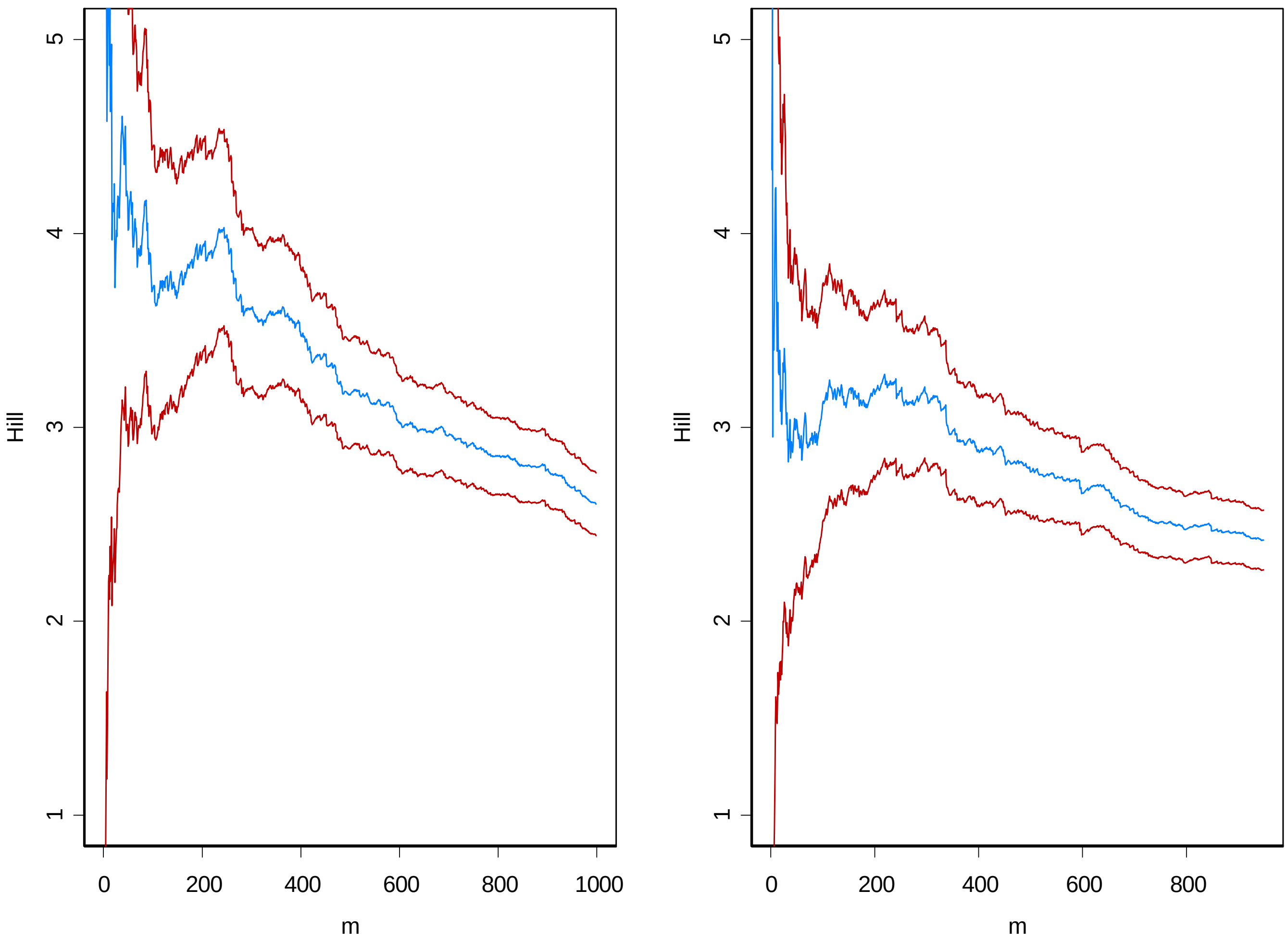
Sample ACF IBM (a) 1962-1981, (b) 1982-2000



Sample ACF of abs values for IBM (a) 1961-1981, (b) 1982-2000



Hill's plot of tail index for IBM (1962-1981, 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 ν df.)
- $\{\sigma_t\}$ is the volatility process
- σ_t and Z_t are independent.

Properties:

- $EX_t = 0$, $\text{Cov}(X_t, X_{t+h}) = 0$, $h>0$ (uncorrelated if $\text{Var}(X_t) < \infty$)
- conditional heteroscedastic (condition on σ_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):

$$\begin{aligned} X_t^2 &= (\alpha_0 + \alpha_1 X_{t-1}^2) Z_t^2 \\ &= \alpha_1 Z_t^2 X_{t-1}^2 + \alpha_0 Z_t^2 \\ &= A_t X_{t-1}^2 + B_t \end{aligned}$$

(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 $\mathbf{Y}_t = (\sigma_t^2, X_{t-1}^2)'$ follows the SRE given by

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

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)

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

$$\rho_{X^2}(h) = Cor(\sigma_t^2, \sigma_{t+h}^2)/EZ_1^4$$

Question:

- Joint distributions of process regularly varying if distr of Z_1 is regularly varying?

Regular variation — univariate case

Def: The random variable *X* is *regularly varying with index* α if

$$P(|X|> t\,x)/P(|X|>t) \to x^{-\alpha} \text{ and } P(X> t)/P(|X|>t) \to p,$$

or, equivalently, if

$$P(X> t\,x)/P(|X|>t) \to px^{-\alpha} \text{ and } P(X< -t\,x)/P(|X|>t) \to qx^{-\alpha},$$

where $0 \le p \le 1$ and $p+q=1$.

Equivalence:

X is RV(α) *if and only if* $P(X \in t \bullet) /P(|X|>t) \to_v \mu(\bullet)$

($\to_v$ vague convergence of measures on $\mathbb{R}\backslash\{0\}$). 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.

Regular variation — univariate case (cont)

Another formulation (polar coordinates):

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

Then

X is RV(α) *if and only if*

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

or

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

($\to_v$ 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

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

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

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

Equivalence:

$$\frac{P(\mathbf{X} \in t\bullet)}{P(|\mathbf{X}| > t)} \to_v \mu(\bullet)$$

μ is a measure on R^m which satisfies for $x > 0$ and A bounded away from 0,

$$\mu(xB) = x^{-\alpha}\mu(xA).$$

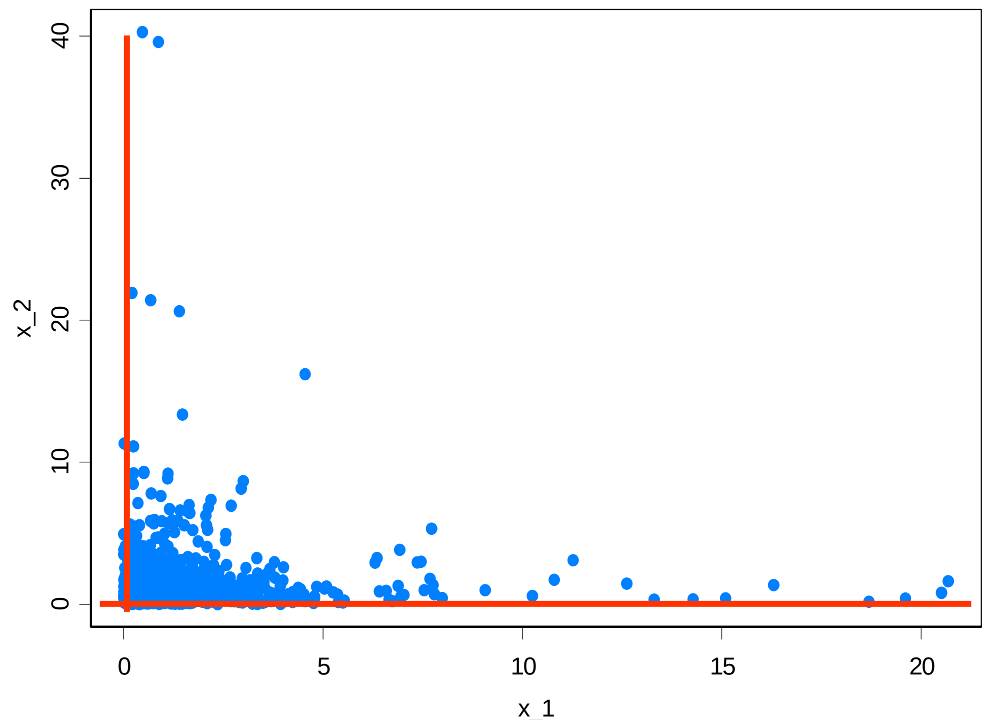
Examples:

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

$$P(\theta = (0,1)) = P(\theta = (1,0)) = .5 \quad \text{(mass on axes).}$$

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.



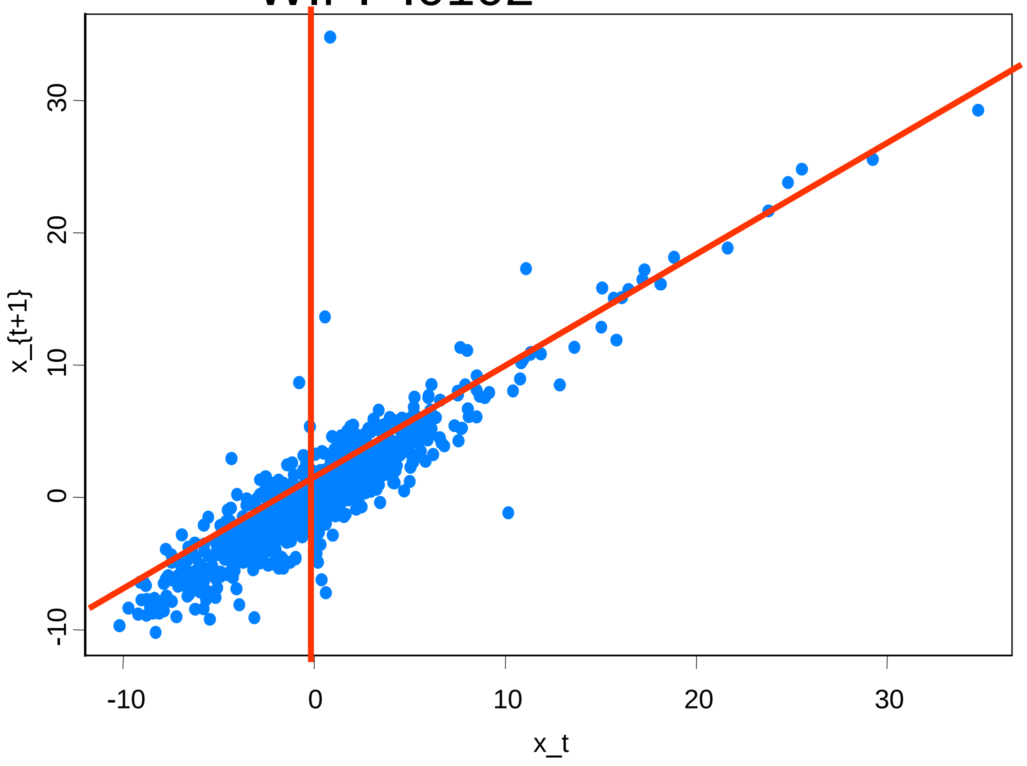
2. If $X_1 = X_2 > 0$, then $\mathbf{X} = (X_1, X_2)$ is multivariate regularly varying with index α and *spectral distribution*

$$P(\theta = (1/\sqrt{2}, 1/\sqrt{2})) = 1.$$

3. AR(1): $X_t = .9 X_{t-1} + Z_t$, $\{Z_t\}$~IID symmetric stable (1.8)

$$\text{Distr of } \theta: \begin{cases} \pm(1,.9)/\text{sqrt}(1.81), & \text{W.P. } .9898 \\ \pm(0,1), & \text{W.P. } .0102 \end{cases}$$

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

$$a_n^{-1}\sum_{t=1}^{n}\mathbf{X}_t$$

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

$$a_n^{-1}\bigvee_{t=1}^{n}\mathbf{X}_t$$

- Weak convergence of *point processes* with iid points.
- Solution to *stochastic recurrence equations*, $\mathbf{Y}_t = \mathbf{A}_t\,\mathbf{Y}_{t-1} + \mathbf{B}_t$
- Weak convergence of *sample autocovariances*.

Operations on regularly varying vectors — products

Products (Breiman 1965). Suppose X, Y > 0 are independent with

X~RV(α) and $EY^{\alpha+\varepsilon} < \infty$ for some $\varepsilon > 0$. Then XY ~ RV(α) 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

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

Then

AX is regularly varying with index α.

Applications of multivariate regular variation (cont)

Linear combinations:

$\mathbf{X} \sim RV(\alpha) \Rightarrow$ all linear combinations of $\mathbf{X}$ are regularly varying

i.e., there exist α and slowly varying fcn $L(.)$, s.t.

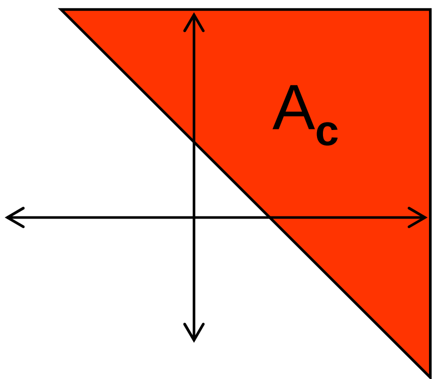
$$P(\mathbf{c}^T\mathbf{X} > t)/(t^{-\alpha}L(t)) \to w(\mathbf{c}), \text{ exists for all real-valued } \mathbf{c},$$

where

$$w(t\mathbf{c}) = t^{-\alpha}w(\mathbf{c}).$$

Use vague convergence with $A_{\mathbf{c}}=\{\mathbf{y}: \mathbf{c}^T\mathbf{y} > 1\}$, i.e.,

$$\frac{P(\mathbf{X} \in tA_{\mathbf{c}})}{t^{-\alpha}L(t)} = \frac{P(\mathbf{c}^T\mathbf{X} > t)}{P(|\mathbf{X}| > t)} \to \mu(A_{\mathbf{c}}) =: w(\mathbf{c}),$$



where $t^{-\alpha}L(t) = P(|\mathbf{X}| > t)$.

Applications of multivariate regular variation (cont)

Converse?

X ~RV(α) $\Leftarrow$ all linear combinations of **X** are regularly varying?

There exist α and slowly varying fcn $L(.)$, s.t.

(LC) $P(\mathbf{c}^T\mathbf{X} > t)/(t^{-\alpha}L(t)) \to w(\mathbf{c})$, exists for all real-valued **c.**

Theorem (Basrak, Davis, Mikosch, `02). Let **X** be a random vector.

1. If **X** satisfies (LC) with α non-integer, then **X** is RV(α).
2. If **X** > 0 satisfies (LC) for non-negative **c** and α is non-integer, then **X** is RV(α).
3. If **X** > 0 satisfies (LC) with α an odd integer, then **X** is RV(α).

Applications of multivariate regular variation (cont)

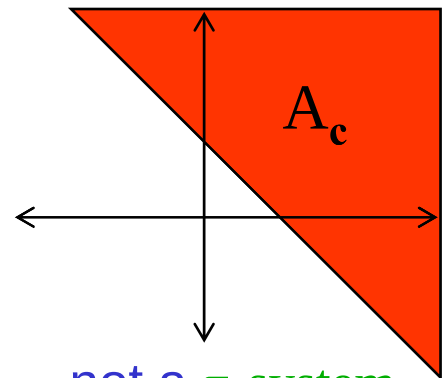
Idea of argument: Define the measures

$$m_t(\bullet) = P(\mathbf{X} \in t\bullet) / (t^{-\alpha} L(t))$$

- By assumption we know that for fixed $\mathbf{c}$, $m_t(A_\mathbf{c}) \to \mu(A_\mathbf{c})$.
- $\{m_t\}$ is tight: For B bded away from 0, $\sup_t m_t(B) < \infty$.
- Do subsequential limits of $\{m_t\}$ coincide?

If $m_{t'} \to_v \mu_1$ and $m_{t''} \to_v \mu_2$, then

$\mu_1(A_c) = \mu_2(A_c)$ for all $\mathbf{c} \neq \mathbf{0}$.



Problem: Need $\mu_1 = \mu_2$ but only have equality on $A_\mathbf{c}$, not a π-system.

In general, equality need not hold (see Ex 6.1.35 in Meerschaert & Scheffler (2001)).

Applications of theorem

1. Kesten (1973). Under general conditions, (LC) holds with L(t)=1 for stochastic recurrence equations of the form

$$\mathbf{Y}_t = \mathbf{A}_t \mathbf{Y}_{t-1} + \mathbf{B}_t, \qquad (\mathbf{A}_t, \mathbf{B}_t) \sim \text{IID},$$

$\mathbf{A}_t$ $d \times d$ random matrices, $\mathbf{B}_t$ random d-vectors.

It follows that the distributions of $\mathbf{Y}_t$, and in fact all of the finite dim'l distrs of $\mathbf{Y}_t$ are regularly varying (if α is non-even).

2. GARCH processes. Since squares of a GARCH process can be embedded in a SRE, the *finite dimensional distributions* of a *GARCH* are regularly varying.

Examples

Example of ARCH(1): $X_t=(\alpha_0+\alpha_1 X^2_{t-1})^{1/2}Z_t$, $\{Z_t\}\sim$IID.

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

α_1	.312	.577	1.00	1.57
α	8.00	4.00	2.00	1.00

Distr of θ:

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

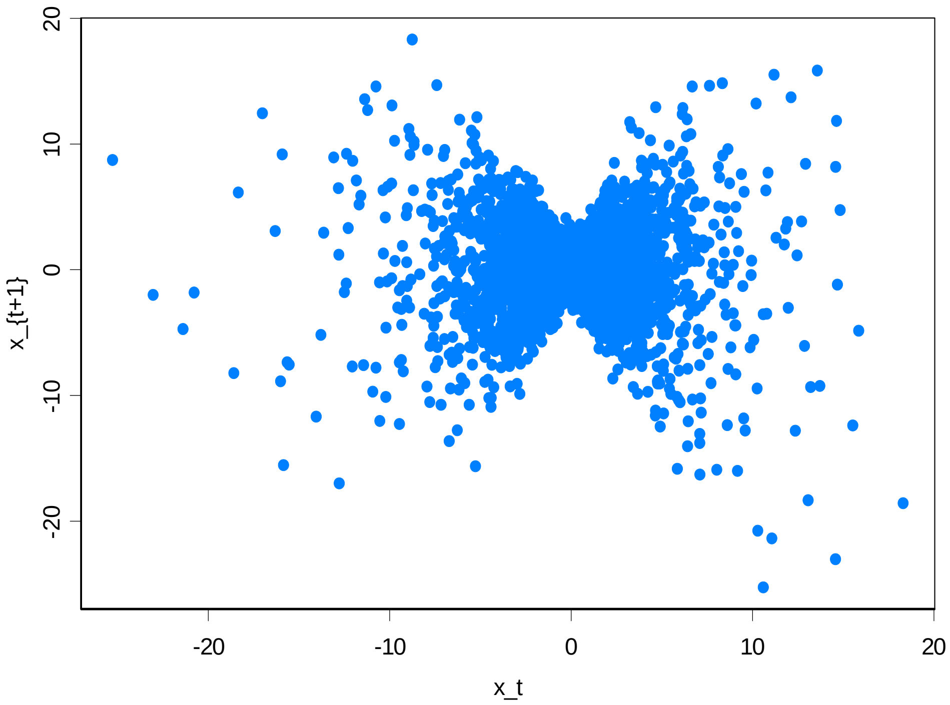
where

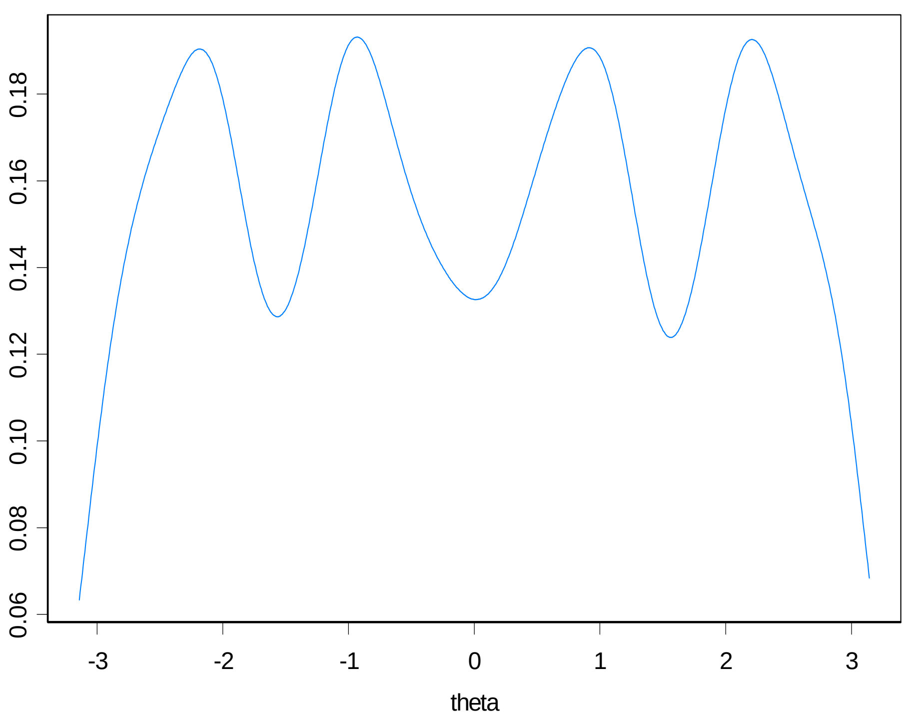
$$P(B = 1) = P(B = -1) = .5$$

Examples (cont)

Example of ARCH(1): $\alpha_0=1$, $\alpha_1=1$, $\alpha=2$, $X_t=(\alpha_0+\alpha_1 X^2_{t-1})^{1/2}Z_t$, $\{Z_t\}$~IID

Figures: plots of (X_t, X_{t+1}) and estimated distribution of θ for realization of 10,000.





Applications of theorem (cont)

Example: SV model $X_t = \sigma_t Z_t$

Suppose $Z_t \sim RV(\alpha)$ and

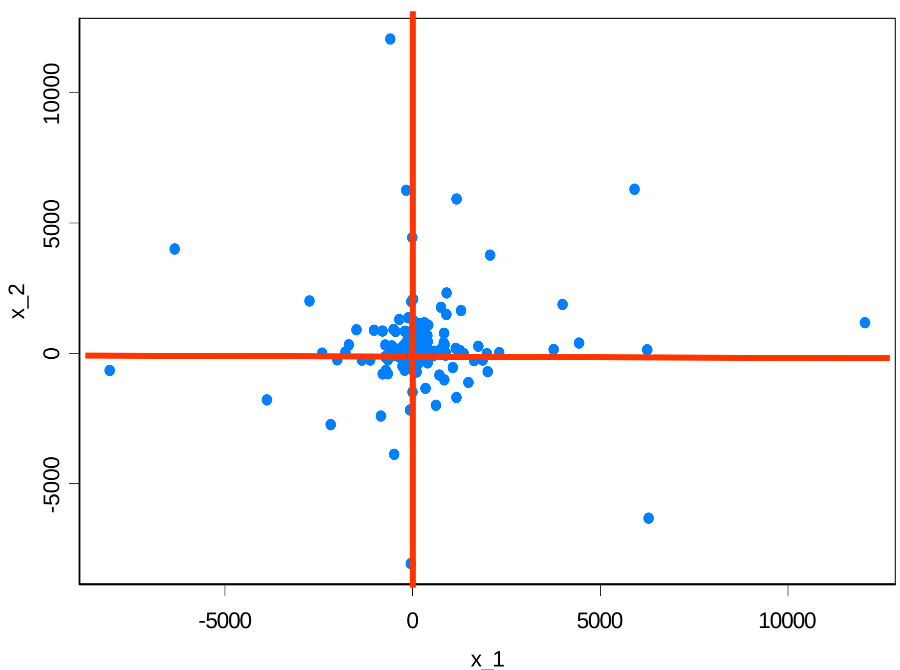
$$\log\sigma_t^2 = \sum_{j=-\infty}^{\infty}\psi_j\varepsilon_{t-j}, \quad \sum_{j=-\infty}^{\infty}\psi_j^2 < \infty, \{\varepsilon_t\} \sim \text{IID N}(0,\sigma^2).$$

Then $\mathbf{Z}_n=(Z_1,\ldots,Z_n)'$ is regulary varying with index α and so is

$$\mathbf{X}_n = (X_1,\ldots,X_n)' = \text{diag}(\sigma_1,\ldots,\sigma_n)\,\mathbf{Z}_n$$

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

Figure: plot of (X_t,X_{t+1}) for realization of 10,000.



Point process application

Theorem Let $\{\mathbf{X}_t\}$ be an iid sequence of random vectors satisfying 1 of the 3 conditions in the theorem. Then

$$N_n := \sum_{t=1}^{n} \varepsilon_{\mathbf{X}_t / a_n} \xrightarrow{d} N := \sum_{j=1}^{\infty} \varepsilon_{P_i \boldsymbol{\theta}_i},$$

if and only if for every $\mathbf{c} \neq \mathbf{0}$

$$N_{n,\mathbf{c}} := \sum_{t=1}^{n} \varepsilon_{\mathbf{c}'\mathbf{X}_t / a_n} \xrightarrow{d} N_{\mathbf{c}} := \sum_{j=1}^{\infty} \varepsilon_{\mathbf{c}' P_i \boldsymbol{\theta}_i},$$

where $\{a_n\}$ satisfies $nP(|\mathbf{X}_t| > a_n) \to 1$, and N is a Poisson process with intensity measure μ.

- $\{P_i\}$ 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

Point process convergence

<u>Theorem</u> (Davis & Hsing `95, Davis & Mikosch `97). Let $\{\mathbf{X}_t\}$ 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) $\lim_{k\to\infty}\limsup_{n\to\infty} P(\vee_{k\le|t|\le r_n} |\mathbf{X}_t| > a_n y \,\big|\, |\mathbf{X}_0| > a_n y) = 0.$

Then

$$\gamma = \lim_{k\to\infty} E(|\theta_0^{(k)}|^\alpha - \vee_{j=1}^{k} |\theta_j^{(k)}|)_+ / E|\theta_0^{(k)}|^\alpha \qquad \text{(extremal index)}$$

exists. If $\gamma > 0$, then

$$N_n := \sum_{t=1}^{n} \varepsilon_{\mathbf{X}_t/a_n} \xrightarrow{d} N := \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i \mathbf{Q}_{ij}},$$

Point process convergence(cont)

- (P_i) are points of a Poisson process on $(0,\infty)$ with intensity function

$$\nu(dy)=\gamma\alpha y^{-\alpha-1}dy.$$

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

$$\lim_{k\to\infty} E(|\theta_0^{(k)}|^\alpha - \bigvee_{j=1}^{k}|\theta_j^{(k)}|)_+ \, I_\bullet(\sum_{|t|\le k}\varepsilon_{\theta_t^{(k)}}) / E(|\theta_0^{(k)}|^\alpha - \bigvee_{j=1}^{k}|\theta_j^{(k)}|)_+$$

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

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

Then

$$P^n(b_n^{-1} X_1 \le x) \rightarrow \exp\{-x^{-\alpha}\}.$$

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

(i) GARCH:

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

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

(ii) SV model:

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

extremal index $\gamma = 1$ no clustering.

Extremes for GARCH and SV processes (cont)

(i) GARCH: $P(b_n^{-1}M_n \le x) \to \exp\{-\gamma x^{-\alpha}\}$

(ii) SV model: $P(b_n^{-1}M_n \le x) \to \exp\{-x^{-\alpha}\}$

Remarks about extremal index.

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

(ii) Numerical example. Suppose c is a threshold such that

$$P^n(b_n^{-1}X_1 \le c) \sim .95$$

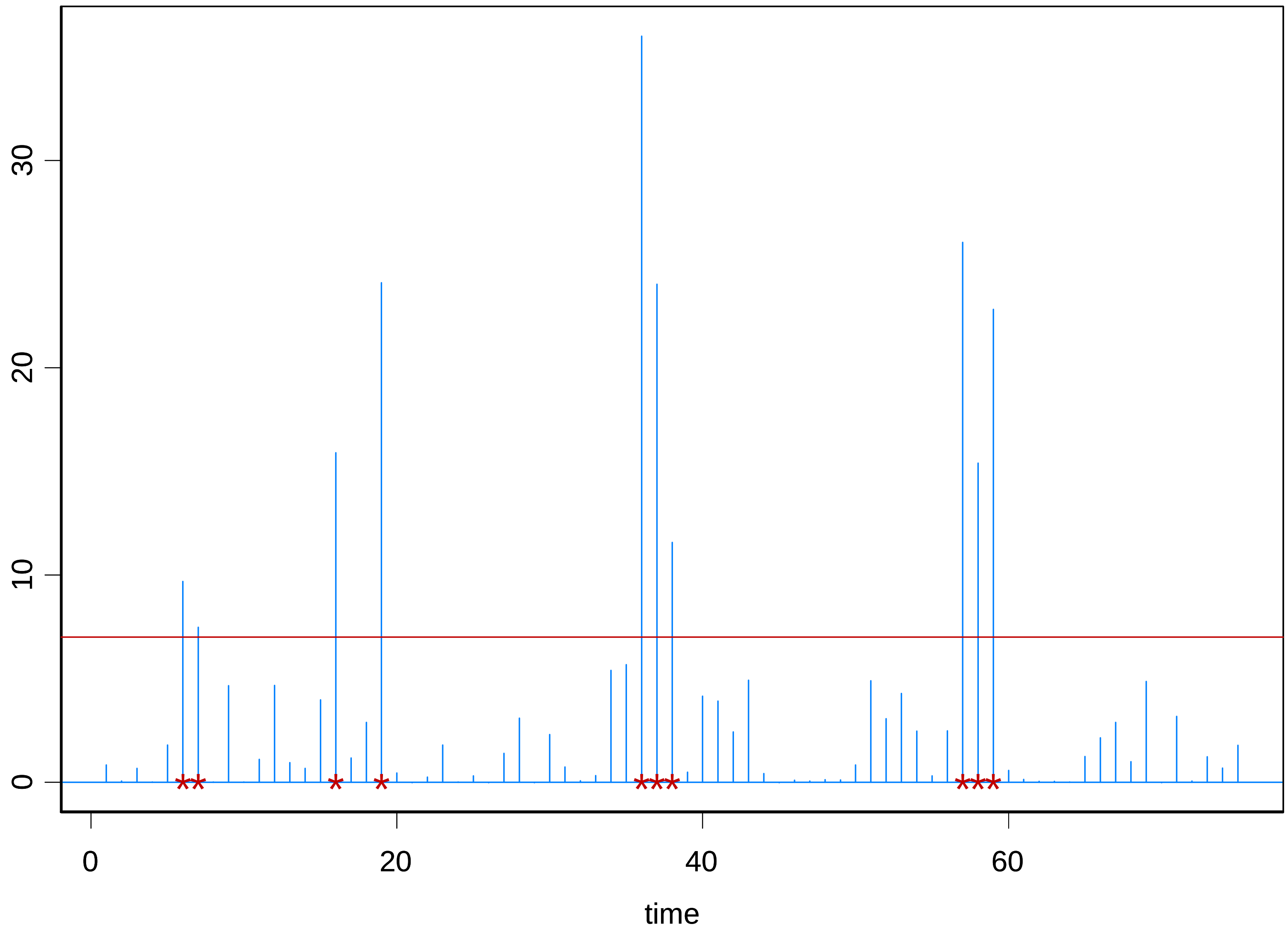
Then, if $\gamma = .5$, $P(b_n^{-1}M_n \le c) \sim (.95)^{.5} = .975$

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

(iv) Use γ to *discriminate* between GARCH and SV models.

(v) Even for the light-tailed SV model (i.e., $\{Z_t\}$ ~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)

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

$$(\hat{\rho}_X(h))_{h=1,\ldots,m} \xrightarrow{d} (V_h / V_0)_{h=1,\ldots,m},$$

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

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

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

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

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

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

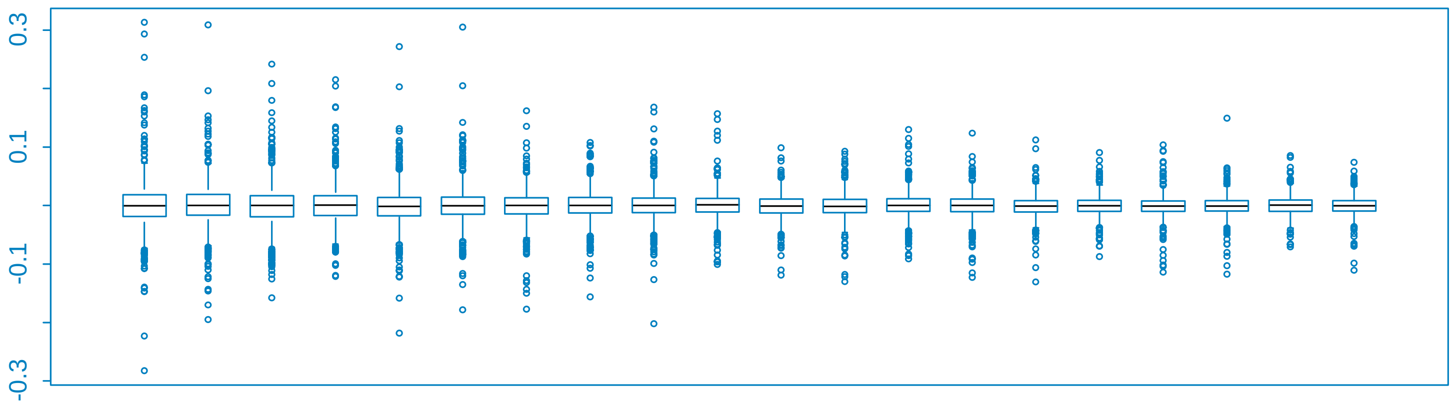
$$(n/\ln n)^{1/\alpha}\hat{\rho}_X(h) \xrightarrow{d} \frac{\|\sigma_1\sigma_{h+1}\|_\alpha}{\|\sigma_1\|_\alpha^2}\frac{S_h}{S_0}.$$

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

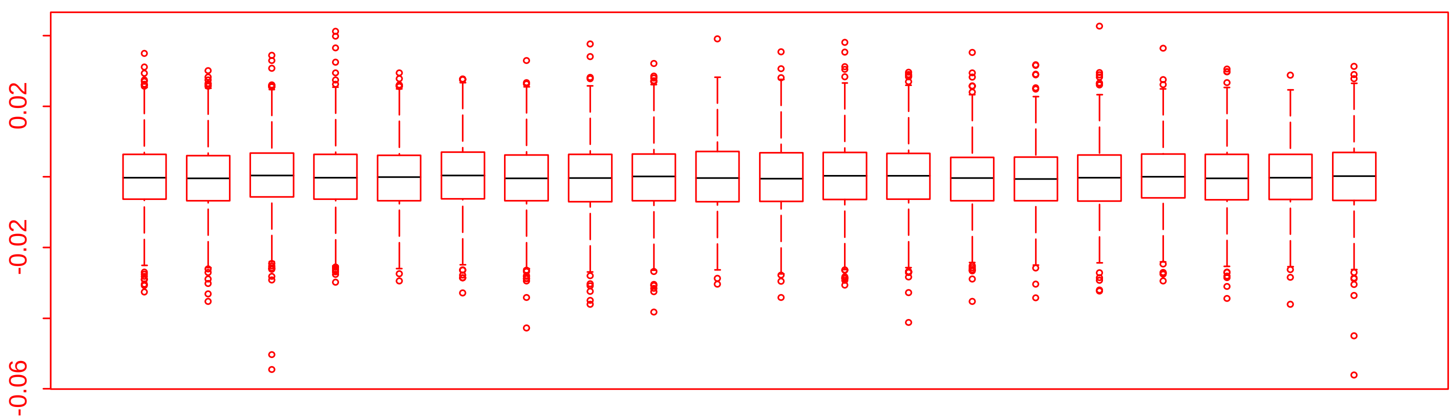
$$\left(n^{1/2}\hat{\rho}_X(h)\right)_{h=1,\ldots,m} \xrightarrow{d} \gamma_X^{-1}(0)\left(G_h\right)_{h=1,\ldots,m}.$$

Sample ACF for GARCH and SV Models (1000 reps)

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

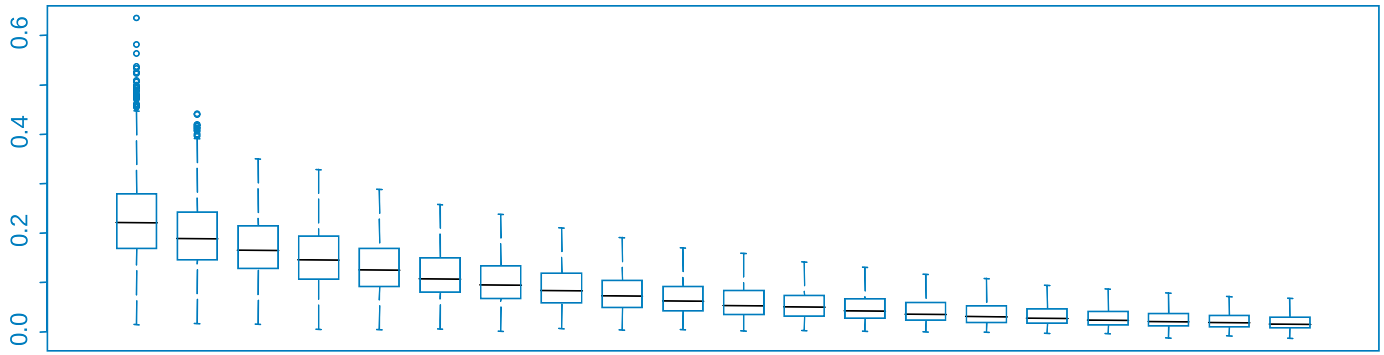


(b) SV Model, n=10000

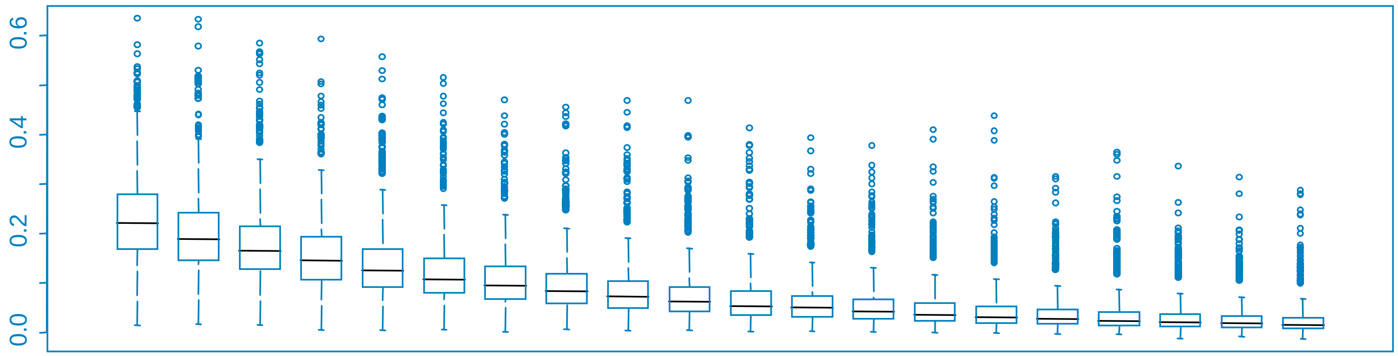


Sample ACF for Squares of GARCH (1000 reps)

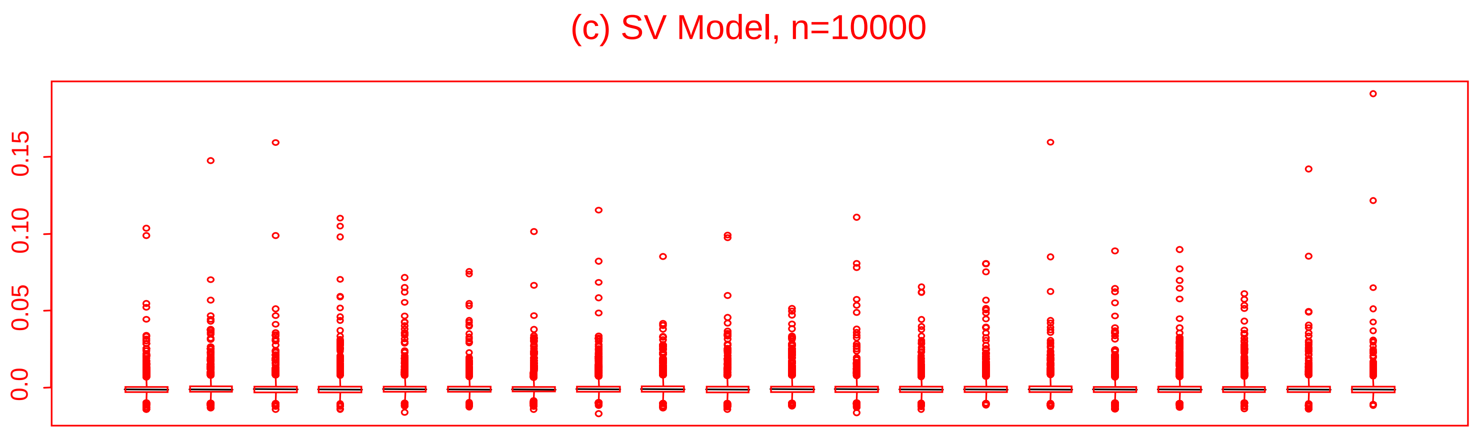
(a) GARCH(1,1) Model, n=10000

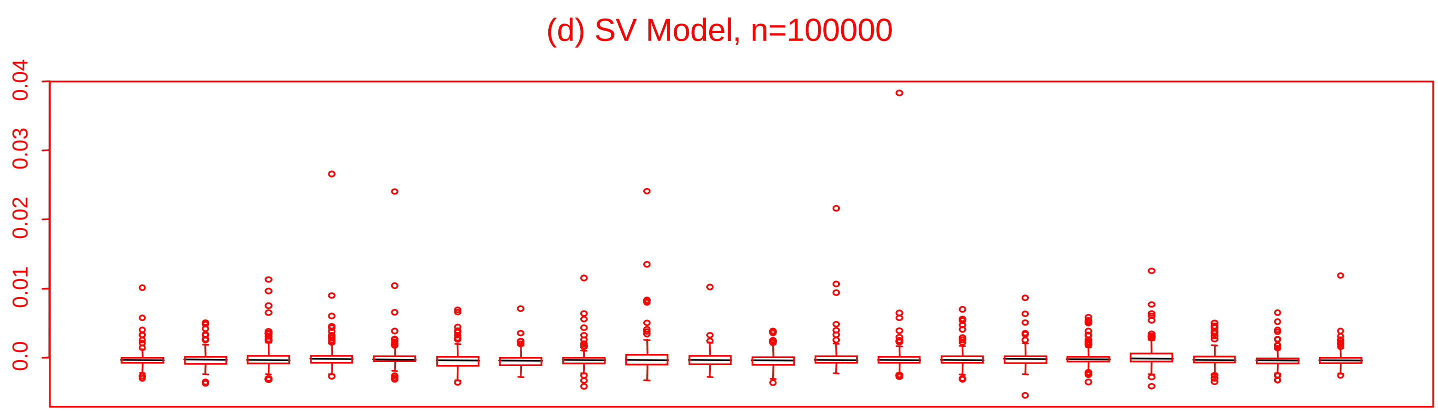


b) GARCH(1,1) Model, n=100000

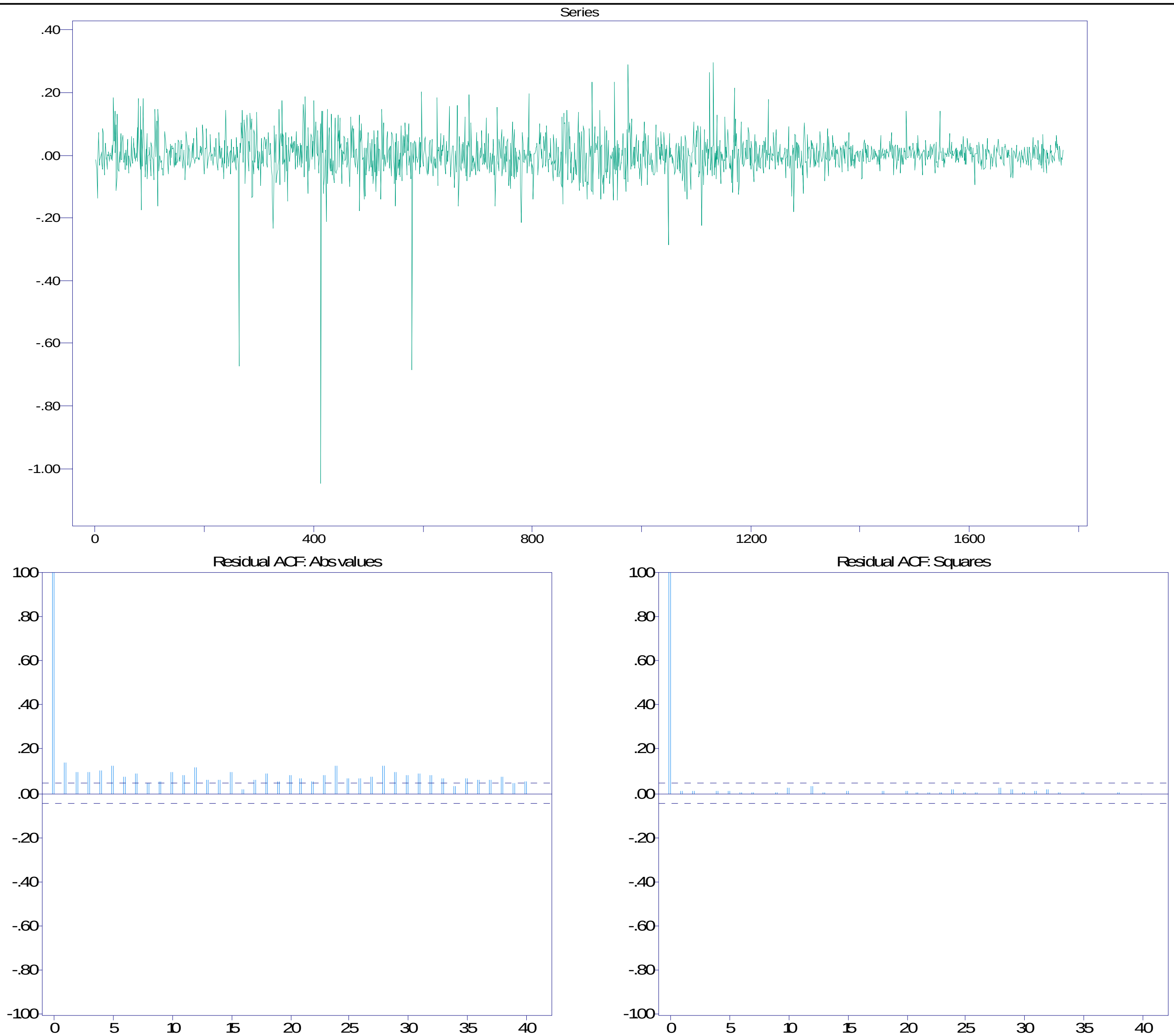


Sample ACF for Squares of SV (1000 reps)





Amazon returns May 16, 1997 to June 16, 2004.



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 $\mathbf{X}_1, \mathbf{X}_2 > 0$, and $(\mathbf{c}, \mathbf{X}_1)$ and $(\mathbf{c}, \mathbf{X}_2)$ have the same limits for all $\mathbf{c} > \mathbf{0}$.

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