Modeling multivariate extremes via regular variation: an application to high-frequency financial returns

Grant Weller

NCSU Statistics Seminar
March 3, 2015

Joint work with Mingyu Tang, Carnegie Mellon University

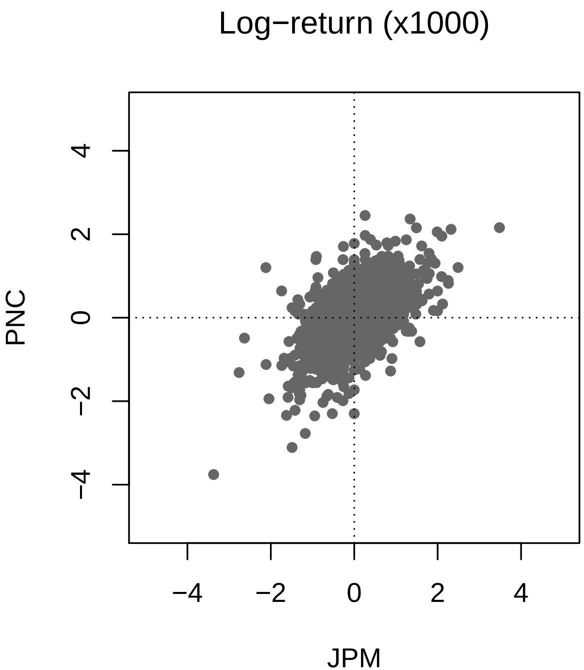
Support from NSF grant DMS-1043903

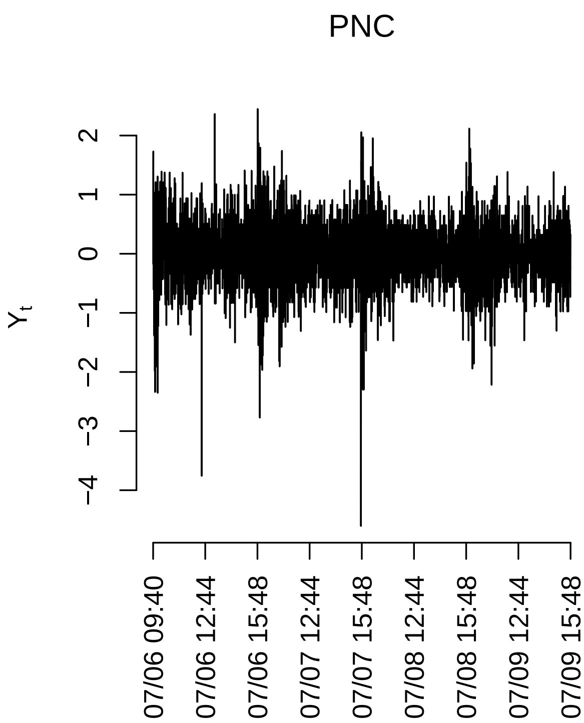
High-frequency trading

- More than 70% of total volume in 2010
- “Flash crash” (May 2010)
- Controversial – market manipulation?

Why study high-frequency data?

- Statistical arbitrage for flash trading
- Protection against risk / identifying trading opportunities





“HFT isn’t going away anytime soon; the best we can do is try to understand it.” – Rene Carmona

Statistical properties of high-frequency returns

Some of "the usuals"...

- Temporal volatility clustering
- *Heavy tails*
- Correlation and tail dependence

... and additionally, microstructure properties:

- Intraday volatility patterns
- Spikes / jumps

Tail properties have been found to vary with time scales considered (Müller et al, 1999).

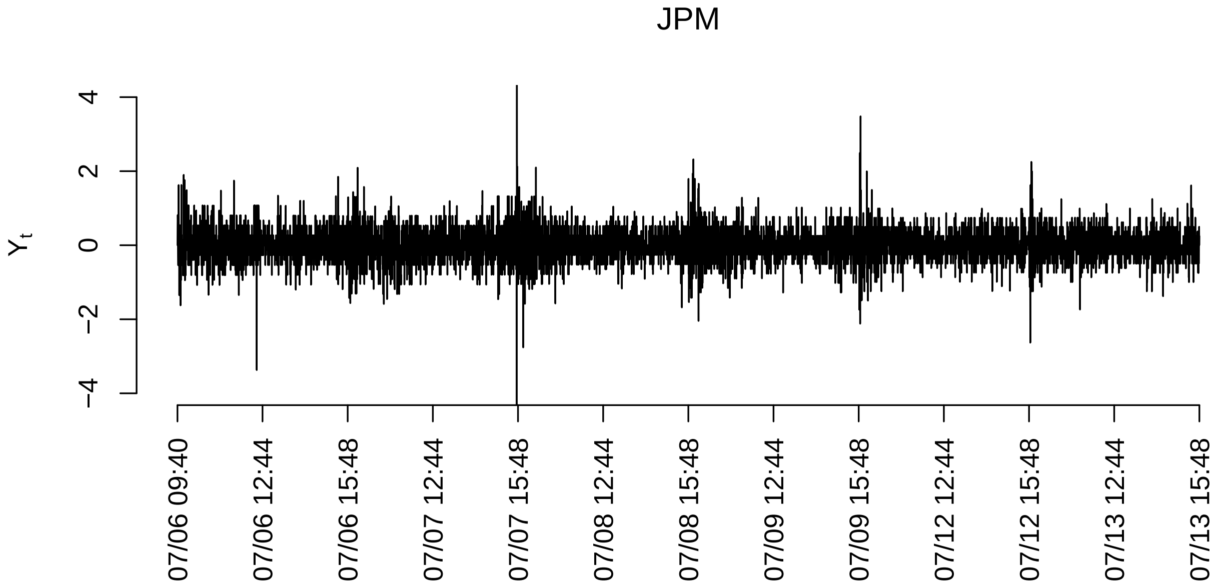
Data

We obtained eight days of 15-second returns on four banking sector stocks.

As is common practice, we work with *log-returns*:

$$Y_{it} = \log(p_{i,t}/p_{i,t-1})$$

for stock i at time increment t.



Split the data into training (4 days) and test (4 days) sets.

Outline

- Modeling marginal temporal features
- Studying tail dependence
  - What is tail dependence?
  - Modeling via regular variation
- Tail dependence in banking stock returns
- Portfolio VaR forecasting
- Extensions

Extracting marginal temporal structure

Let Y_{it} be the log-return of stock i at time t. We use the model

$$Y_{it} = \phi_i(t)\sigma_{i,t}Z_{i,t},$$

where

$$\begin{aligned}\phi_i(t) &= \phi_i(t + 24\ hrs)\\ \sigma^2_{i,t} &= \gamma + \alpha Z^2_{i,t-1} + \beta\sigma^2_{i,t-1}\\ Z_{i,t} &\overset{iid}{\sim} (0,1)\end{aligned}$$

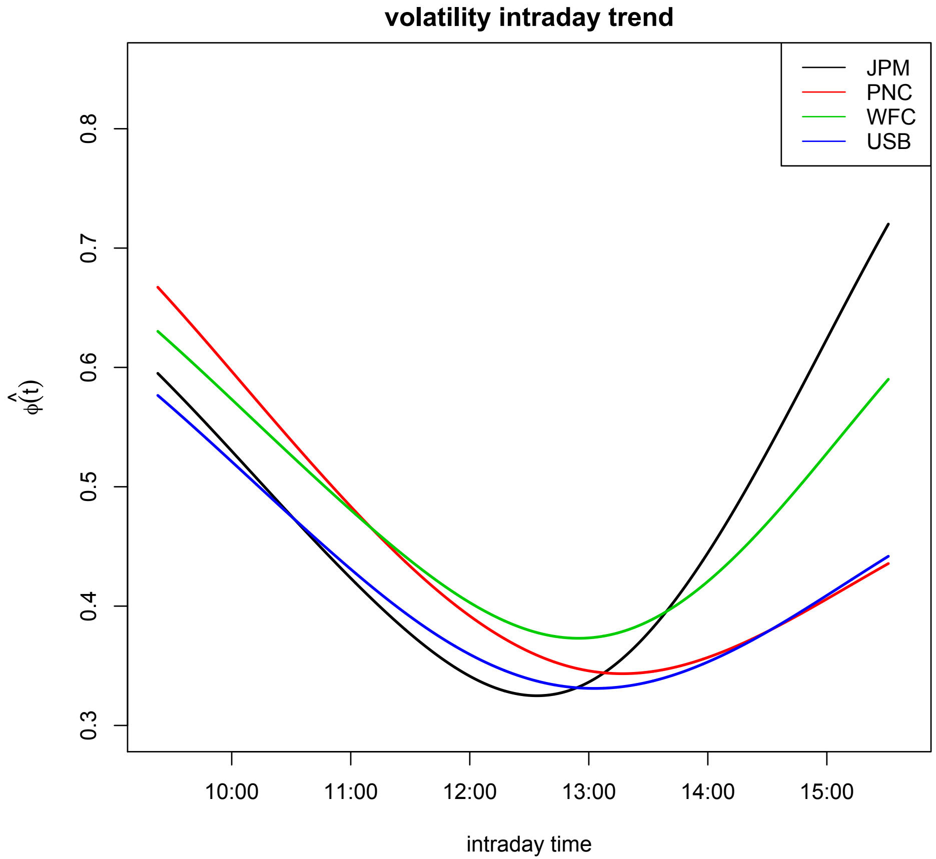
GARCH(1,1) structure with intraday seasonal volatility.

Modeling:

- Remove intraday seasonality
- Estimate GARCH model

Intraday volatility

We fit a smoothing spline to the intraday volatility pattern.

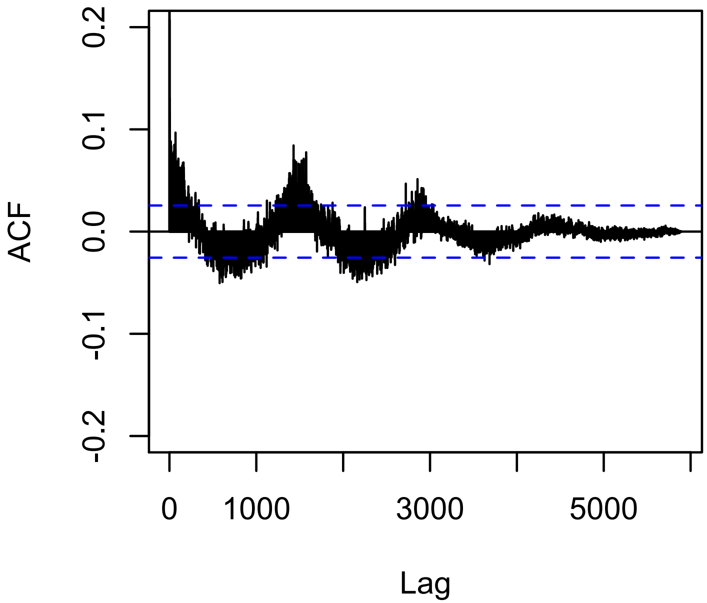


Typical ‘U-shaped’ volatility curves

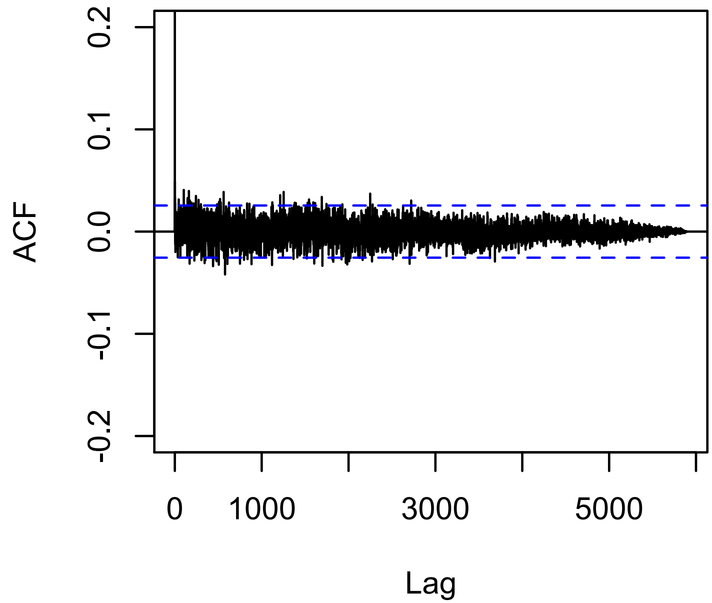
GARCH model

$$\sigma_{i,t}^2 = \gamma + \alpha Z_{i,t-1}^2 + \beta \sigma_{i,t-1}^2$$

Volatility sequences:



$|Y_t|$



$|\hat{Z}_t|$

Marginal extremal behavior

We fit generalized Pareto distributions to exceedances of upper and lower tails of estimated residual sequences $Z_{j,t}$.

$$\mathbb{P}(Z > z \mid Z > u) = \left(1 + \xi \frac{z-u}{\sigma}\right)^{-1/\xi}$$

Retain 7.5% of the data (in each tail) for GPD fitting.

Estimates of ξ:

Stock	JPM	PNC	WFC	USB
Upper	−0.10 (0.08)	0.11 (0.08)	0.23 (0.13)	−0.01 (0.09)
Lower	0.26 (0.08)	0.22 (0.08)	0.34 (0.10)	0.20 (0.08)

Lower tail appears heavier than upper tail.

Outline

- Modeling marginal temporal features
- Studying tail dependence
  - What is tail dependence?
  - Modeling via regular variation
- Tail dependence in banking stock returns
- Portfolio VaR forecasting
- Extensions

Multivariate extremes and tail dependence

You have already learned about univariate EVT. In the multivariate setting, we need a description of 'tail dependence'.

Some important questions:

- How do we talk about dependence for extremes?
- What do we actually mean by a 'multivariate extreme'?
- What modeling frameworks are available for inference?

Here's the 20-minute version...

Dependence for extremes

Multivariate extremes typically focuses on quantities like

$$\mathbb{P}(\max_{i=1,\ldots,n} X_i \le x, \max_{i=1,\ldots,n} Y_i \le y) := G(x,y)$$

or the conditional probability

$$\mathbb{P}(Y > y \mid X > x)$$

for large values of x and y.

Loosely, the question we want to answer is "How do the extreme values of Y depend on the extreme values of X?"

Examples:

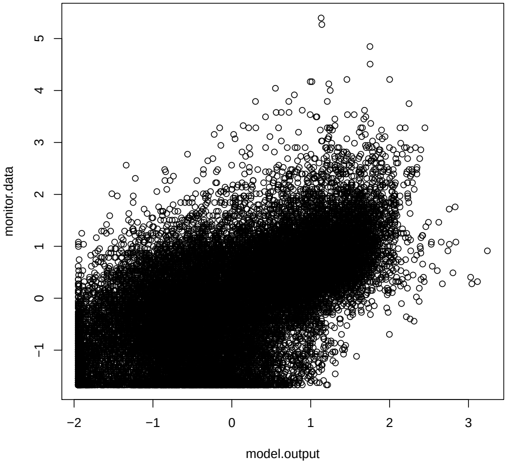
- consecutive days of excessive heat
- monitoring dangerous pollutant levels
- extreme precipitation events affecting multiple locations
- combinations of extreme values which lead to failure

Tail dependence

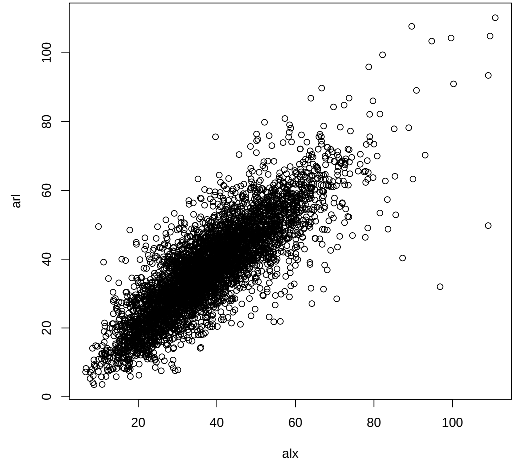
A full characterization of tail dependence is crucial for many risk assessments. But, this is a challenging problem.

'Usual' way of describing dependence:

$$\rho = \frac{\mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)]}{\sqrt{\mathbb{E}[(X - \mathbb{E}X)^2]\mathbb{E}[(Y - \mathbb{E}Y)^2]}}$$



$\widehat{\rho} = 0.59$



$\widehat{\rho} = 0.83$

Correlation focuses on 'spread from center' and typically does not capture tail dependence.

Measuring tail dependence

Many summary measures exist for tail dependence. Most can be written as functions of each other.

Assume X and Y have cdfs F. An intuitive measure is

$$\begin{aligned}\chi(q) &= \mathbb{P}(F_Y(Y) > q \mid F_X(X) > q) \\ &= \mathbb{P}(Y > y_q \mid X > x_q),\end{aligned}$$

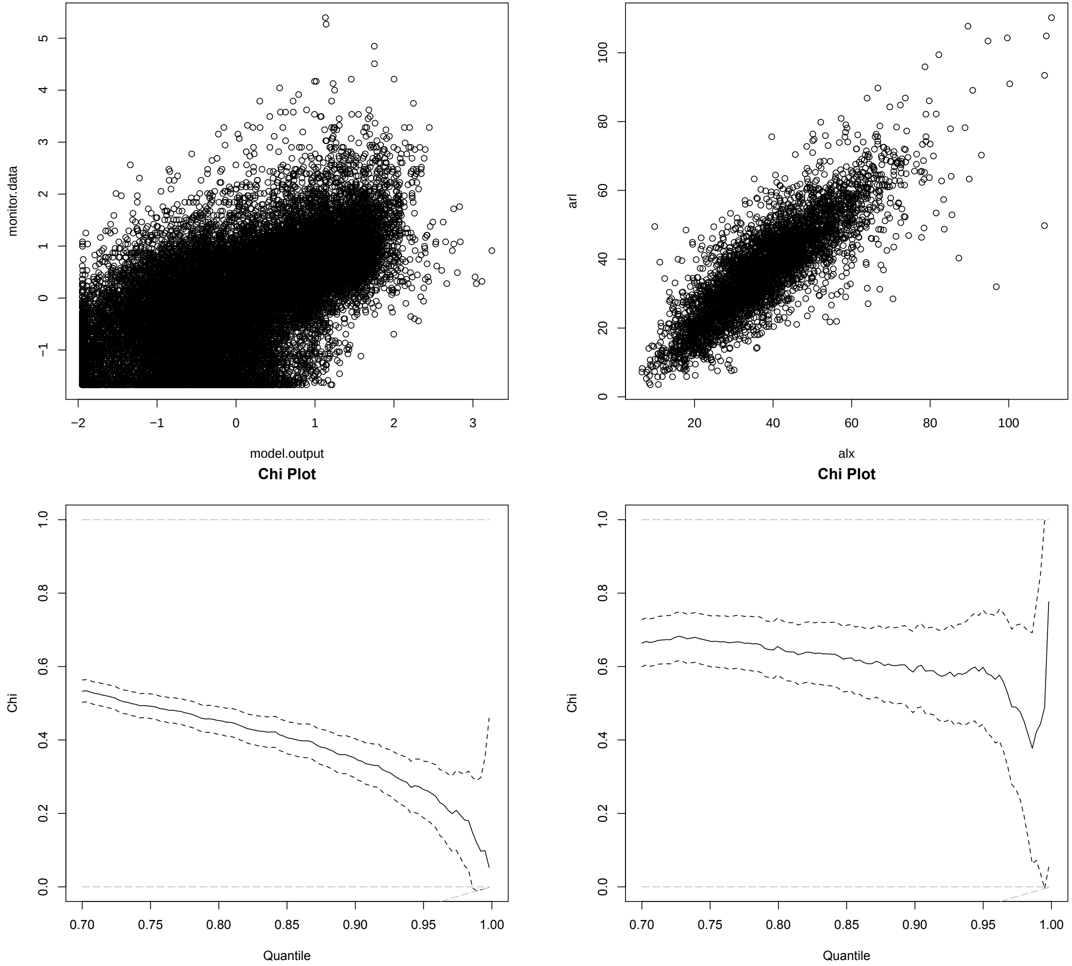
where x_q, y_q are the q quantile levels of X and Y.

If $\lim_{q \to 1} \chi(q) = 0$, X and Y are said to be *asymptotically independent*. Otherwise they are asymptotically dependent.

This is quite different from correlation: e.g., if $(X, Y)^T$ follow a bivariate normal distribution with *any correlation less than one*, X and Y are asymptotically independent.

Estimating tail dependence from data

$\hat{\chi}$ is an empirical measure of asymptotic dependence.



Modeling tail dependence requires more.

Defining a ‘multivariate extreme’

How should we think of multivariate extremes? Which portion of the data do we keep for estimation purposes?

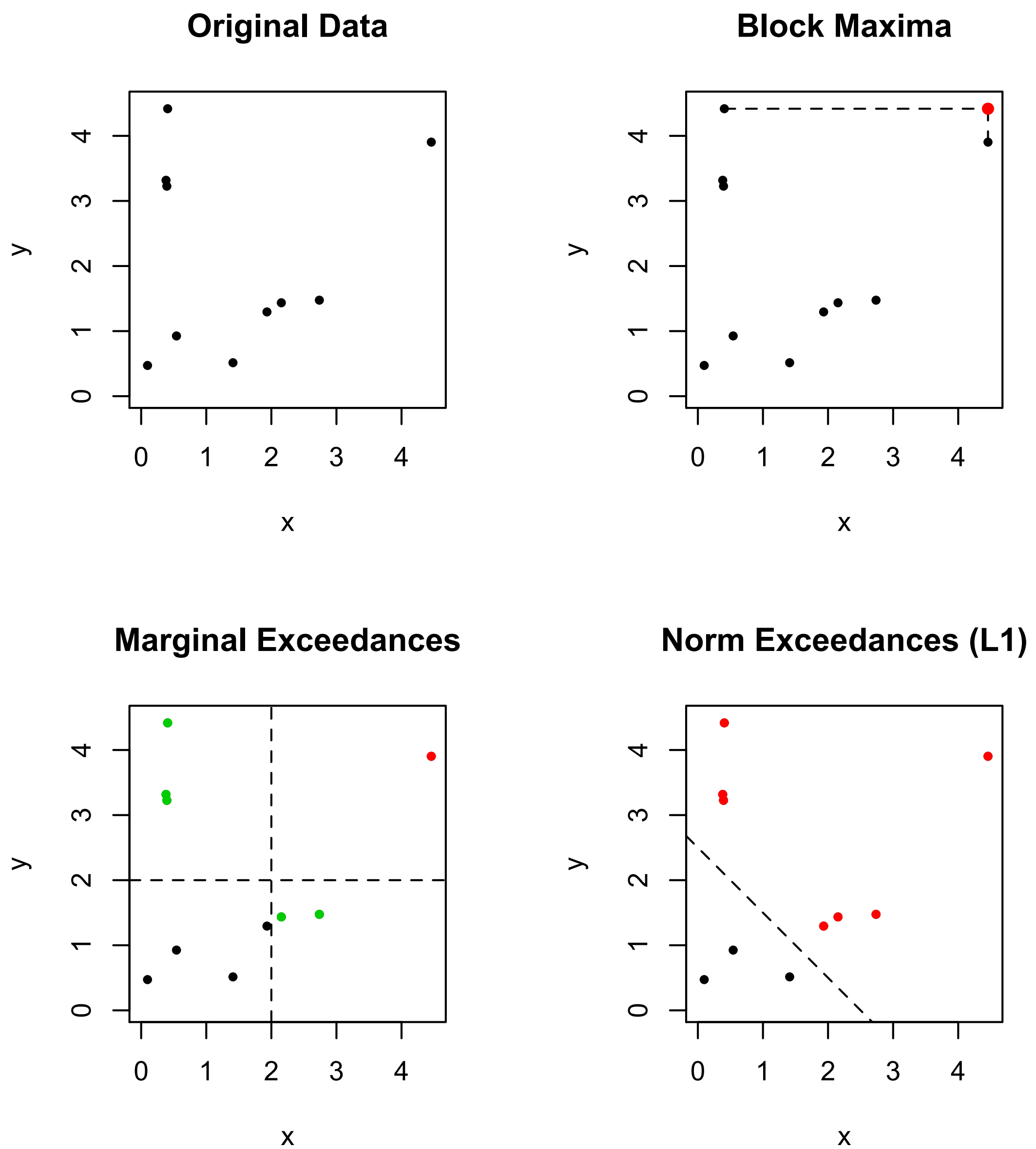
Let $\{\mathbf{Z}_i = (Z_{i,1}, ..., Z_{i,d})^T\}$, $i = 1, ..., n$ be a sequence of observations of a d-dimensional random vector.

Block maxima definition: Construct multivariate block maxima $\mathbf{M}_n = (\vee_{i=1}^n X_{i,1}, ..., \vee_{i=1}^n X_{i,d})^T$. Leads to multivariate max-stable distributions.

Marginal exceedances: For each margin $j = 1, ..., d$, select a threshold u_j and retain data such that $Z_{i,j} > u_j$. Leads to multivariate GPD and censored likelihood methods.

Norm exceedances: (in the heavy-tail case) Retain data such that $\|\mathbf{Z}_i\| > u$ for some norm $\|\cdot\|$. Leads to multivariate regular variation framework.

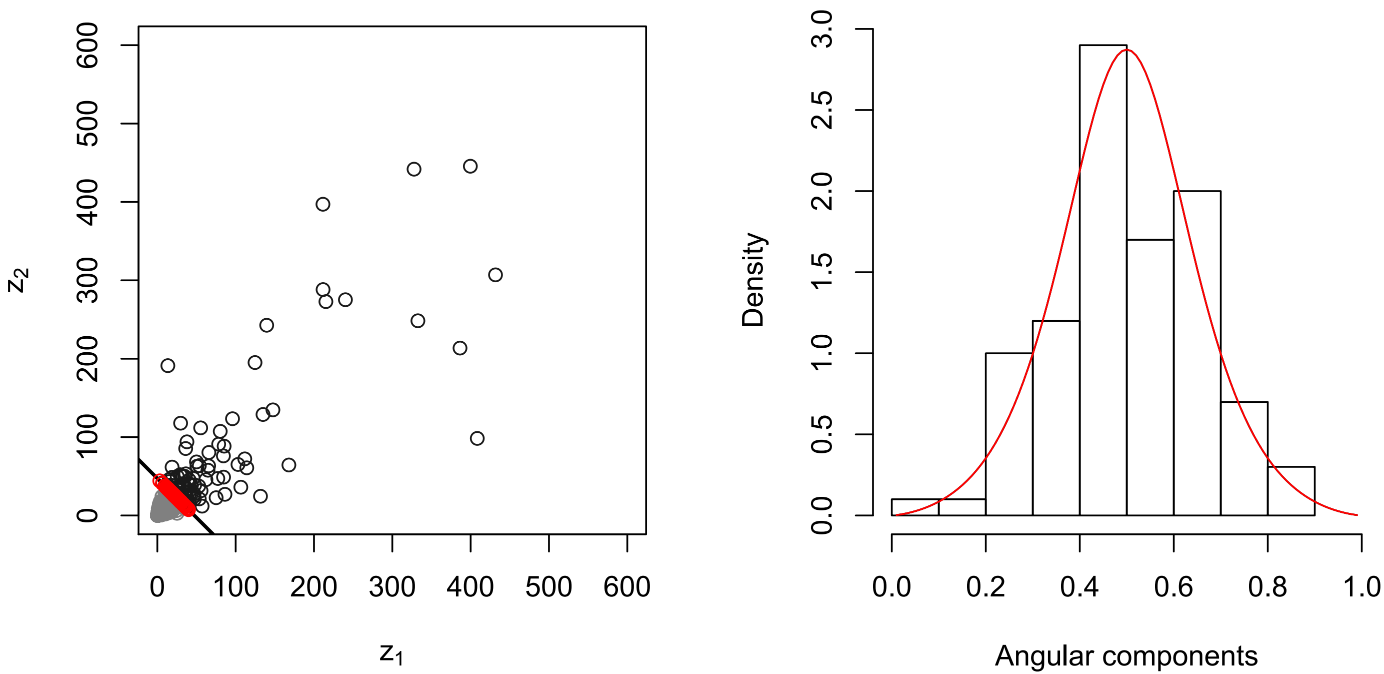
See Huser et al. (2015) for a comparative study


Original Data
Block Maxima
Marginal Exceedances
Norm Exceedances (L1)
x
y
0
1
2
3
4

Multivariate regular variation

Intuitive description: joint tail decay like a power function.

- Spectral decomposition
- Tail dependence described by an *angular measure*



MRV definition 1 (Resnick, 2007)

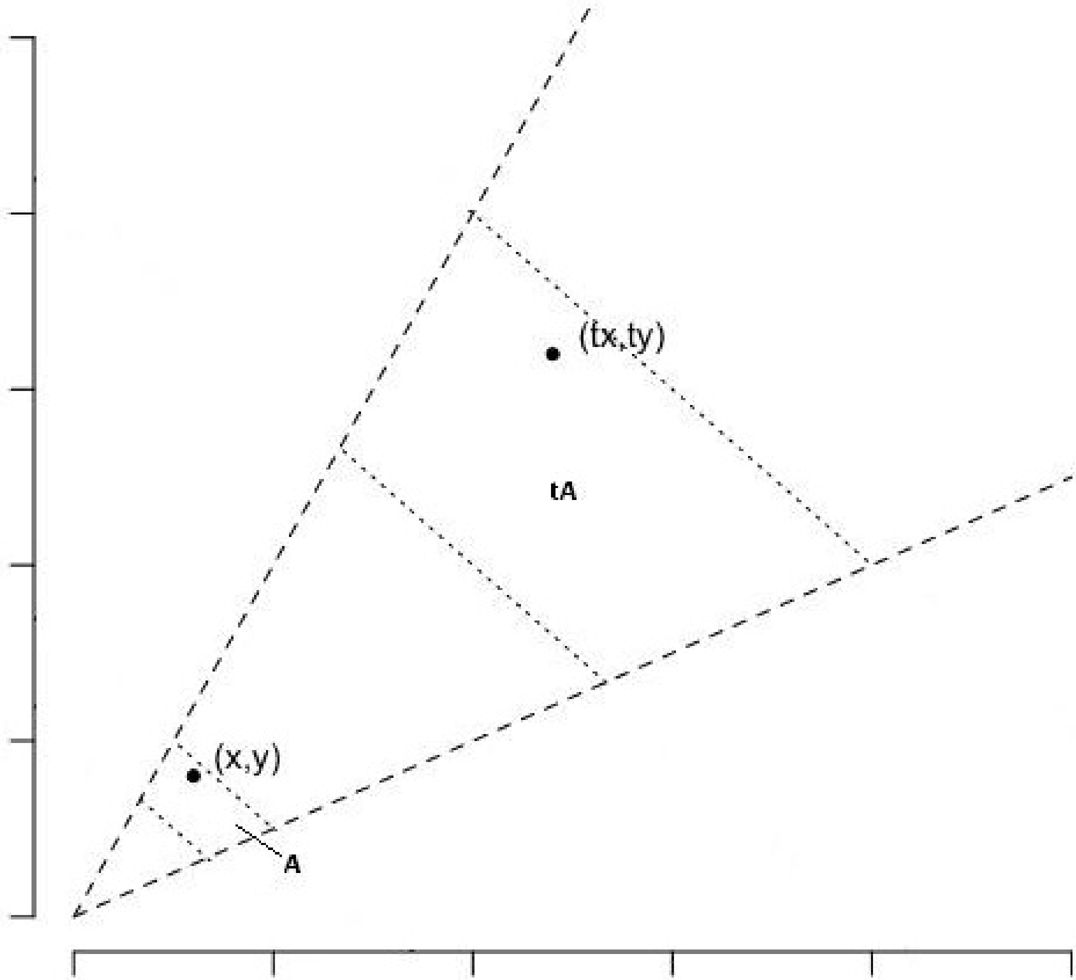
A random vector $\mathbf{Z}$ is regular varying on the cone $\mathfrak{C} = [\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$ if there exists a normalizing sequence $\{a_n\}_{n=1}^{\infty}$ with $a_n \to \infty$ such that

$$n\mathbb{P}\left(\frac{\mathbf{Z}}{a_n} \in \cdot\right) \xrightarrow{v} \nu(\cdot)$$

as $n \to \infty$ in $M_+(\mathfrak{C})$, where $\xrightarrow{v}$ denotes vague convergence of measures.

- Scaling property: $\nu(tA) = t^{-\alpha}\nu(A)$ for $t > 0$, where α is called the *tail index*
- Extremes of the multivariate sample occur according to the limiting measure ν
- a_n is regular varying of index $1/\alpha$ (i.e. $a_n \sim n^{1/\alpha}$)

If marginals are unit Fréchet, $\alpha = 1$.



$$\nu(tA) = t^{-\alpha}\nu(A)$$

MRV definition 2: the angular measure

Define 'radial' and 'angular' components $R = \|\mathbf{Z}\|$, $\mathbf{W} = \mathbf{Z}\|\mathbf{Z}\|^{-1}$, where $\|\cdot\|$ is any norm on $\mathfrak{C}$.

The regular variation condition can then be written

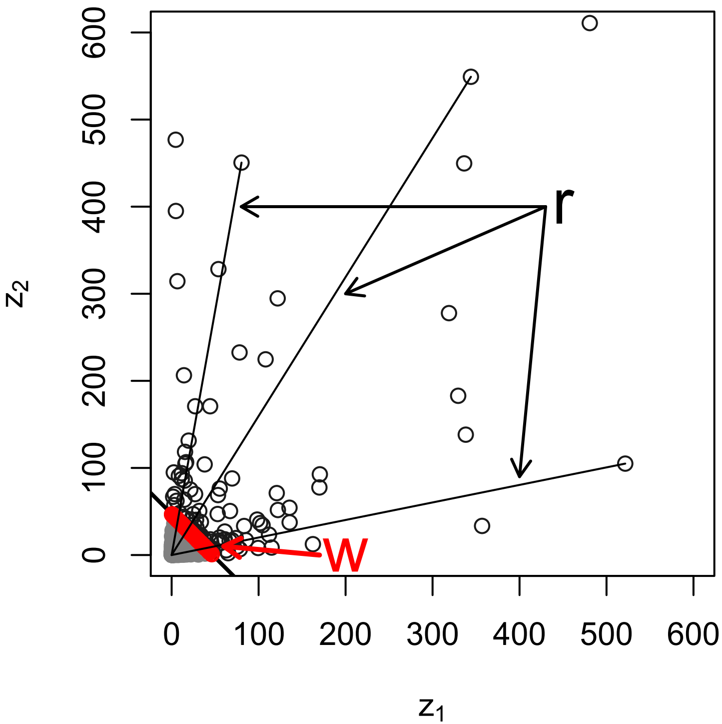
$$n\mathbb{P}\left(\frac{R}{a_n} > r, \mathbf{W} \in B\right) \xrightarrow{v} r^{-\alpha} H(B)$$

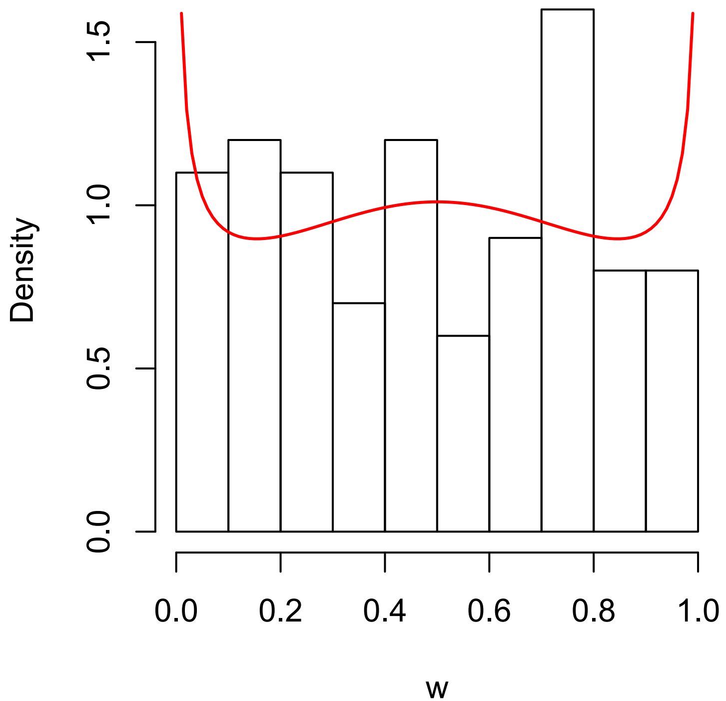
for any Borel set $B \in \mathcal{N} = \{\mathbf{z} \in \mathfrak{C} : \|\mathbf{z}\| = 1\}$.

- H is a measure on $\mathcal{N}$ which characterizes tail dependence
- By choice of a_n, H can be made to be a probability measure
- Equivalent convergence of Poisson point process

Loosely, $(R, \mathbf{W}) \sim r^{-1-\alpha} H(d\mathbf{w})$ for large r.

Radial and angular components





Note: common marginal distributions assumed.

Likelihood inference

For a fixed sample of size n, assume

$$n\mathbb{P}\left(\frac{R}{a_n} > r, \mathbf{W} \in \cdot\right) \cong r^{-\alpha} H(\cdot)$$

for $r > r_0$ (large).

Likelihood:

$$\mathcal{L}(\boldsymbol{\theta} \mid \mathbf{z}_1, \ldots, \mathbf{z}_n) \propto \exp\{-r_0^{-\alpha}\} \prod_{i=1}^{N_0} r_i^{-(1+\alpha)} h(\mathbf{w}_i; \boldsymbol{\theta})$$

where $r_i = \|\mathbf{z}_i\|$ and $\mathbf{w}_i = \mathbf{z}_i \|\mathbf{z}_i\|^{-1}$, for the points $\mathbf{z}_1, \ldots, \mathbf{z}_{N_0}$ with $r_i > r_0$.

MLE can be computed numerically.

Models for H

No finite parameterization encompassing all possible extremal dependence structures.

- Non-parametric approaches (usually in $d = 2$)
- Some parametric subfamilies developed

Simplest parametric model is Gumbel's logistic: in $d = 2$, let $r = z_1 + z_2$ and $w = z_1/r$.

$$h(w, \beta) = \frac{1}{2}(\beta^{-1} - 1)\{w(1-w)\}^{-1-1/\beta}\{w^{-1/\beta} + (1-w)^{-1/\beta}\}^{\beta-2}$$

- $\beta \to 1$ corresponds to asymptotic independence
- $\beta \to 0$ in the case of perfect dependence

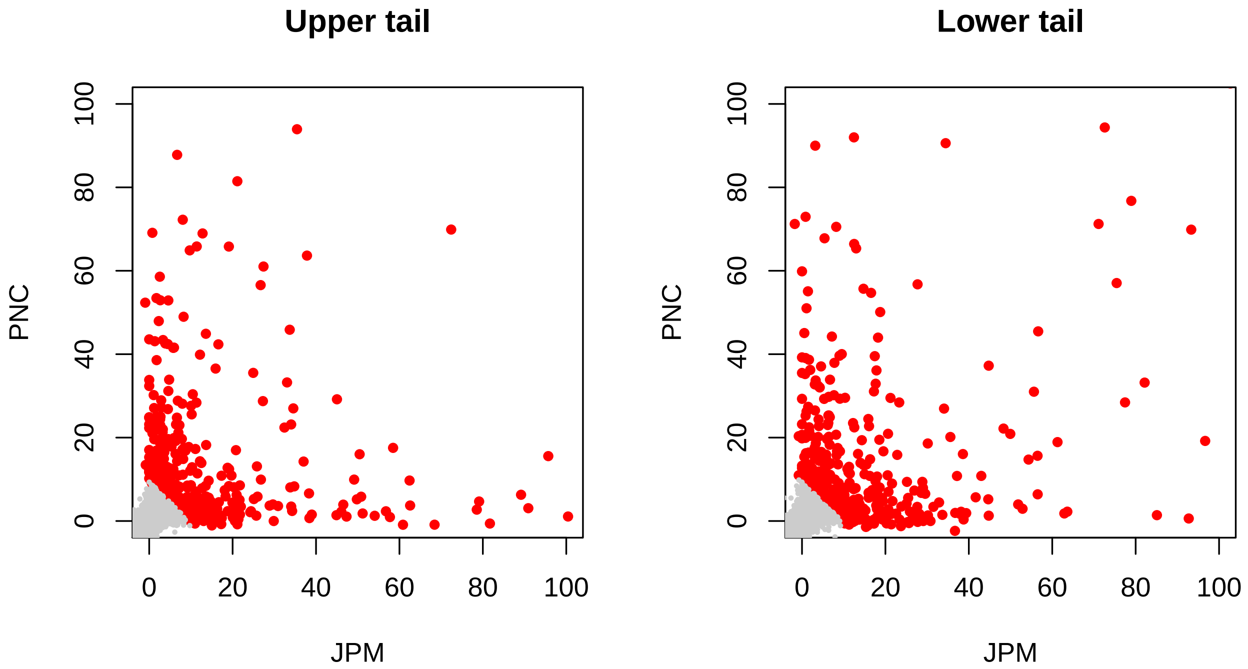
Outline

- Modeling marginal temporal features
- Studying tail dependence
  - What is tail dependence?
  - Modeling via regular variation
- Tail dependence in banking stock returns
- Portfolio VaR prediction
- Extensions

HFT data in tail modeling framework

We examine the pairwise joint upper and lower tail behavior of 15-second returns.

- Transform marginals
- Fit bivariate logistic model



Logistic model parameter estimates

MLE of β for both upper and lower tails.

	PNC	WFC	USB
JPM	0.68 (0.17) 0.65 (0.18)	0.69 (0.17) 0.66 (0.18)	0.69 (0.17) 0.64 (0.19)
PNC		0.66 (0.18) 0.64 (0.19)	0.66 (0.18) 0.62 (0.20)
WFC			0.65 (0.18) 0.66 (0.18)

Moderate levels of asymptotic dependence here across both upper and lower tails.

Outline

- Modeling marginal temporal features
- Studying tail dependence
  - What is tail dependence?
  - Modeling via regular variation
- Tail dependence in banking stock returns
- Portfolio VaR prediction
- Extensions

Portfolio Construction

Consider an equally-weighted (in dollar value) portfolio of two stocks i and j with value

$$X_{ij,t} = w_i p_{i,t} + w_j p_{j,t}$$

at time t, where $w_i p_{i,0} = w_j p_{j,0}$.

What is the behavior of the portfolio log-return sequence $\log(X_{ij,t}/X_{ij,t-1})$?

We consider one-step forecasting of value-at-risk:

$$\mathrm{VaR}_\alpha(Y) = \sup\{z : \mathbb{P}(Y_{t+1} \leq z \mid \mathcal{F}_t) \leq \alpha\},$$

where $\mathcal{F}_t$ represents the information available at time t.

Consider both lower tail (losses) and upper tail (gains).

One-step forecasting

Recall the model

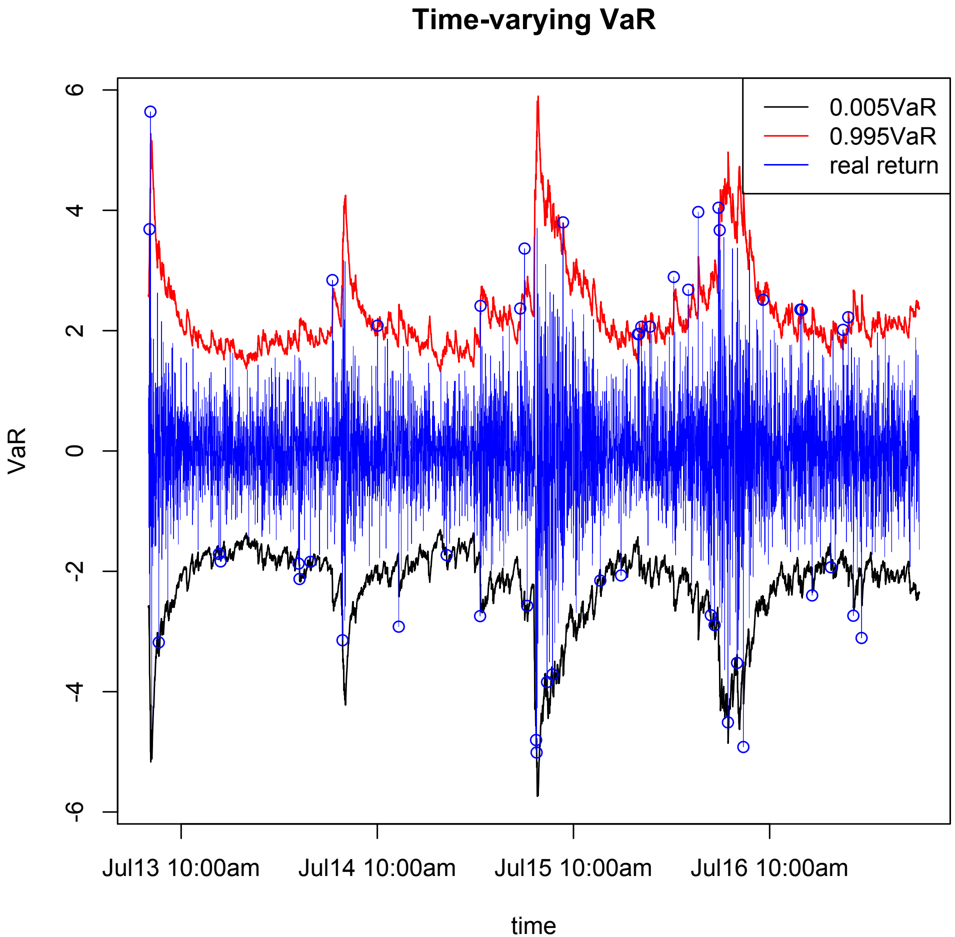
$$Y_{jt} = \phi_j(t)\sigma_{j,t}Z_{j,t}.$$

Forecast the portfolio log-return $X_{ij,t+1}$ via simulation:

1. Generate $(Z_i, Z_j)_{t+1}$ semiparametrically:
   - Use bivariate logistic model above thresholds
   - Empirical distribution below
2. Forecast $\widehat{\sigma}_{i,t+1} = \widehat{\gamma}_i + \widehat{\alpha}_i Z_{i,t}^2 + \widehat{\beta}_i \sigma_{i,t}^2$, also for j.
3. Use $(\widehat{Y}_{i,t+1}, \widehat{Y}_{j,t+1})$ to forecast $X_{ij,t+1}$.

Example

Four days of one-step forecast distributions for an equally-weighted portfolio of JPM and PNC stock.



Comparison with Gaussian approach

Consider the alternative of modeling the residual sequences (Z_i, Z_j) as Gaussian.

- Assumes dependence is completely captured by correlation

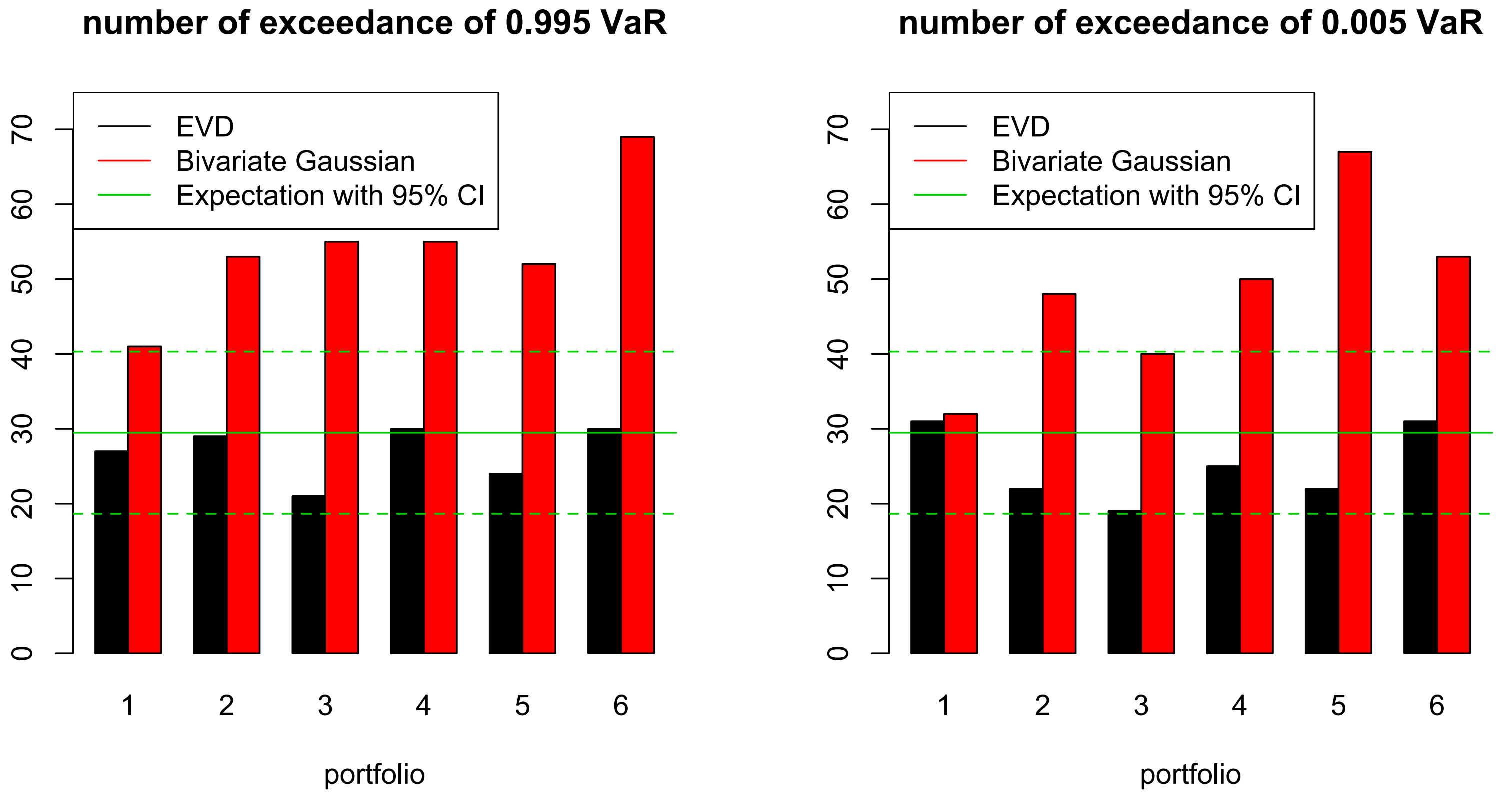
$$\Pr[T_A < 1, T_B < 1] = \Phi_2(\Phi^{-1}(F_A(1)), \Phi^{-1}(F_B(1)), \gamma)$$

Wired magazine: "Recipe for Disaster: The Formula that Killed Wall Street". February 23, 2009.

What does this imply about extremes of portfolio returns?

Comparison with Gaussian approach

Exceedances of estimated VaR from each model:



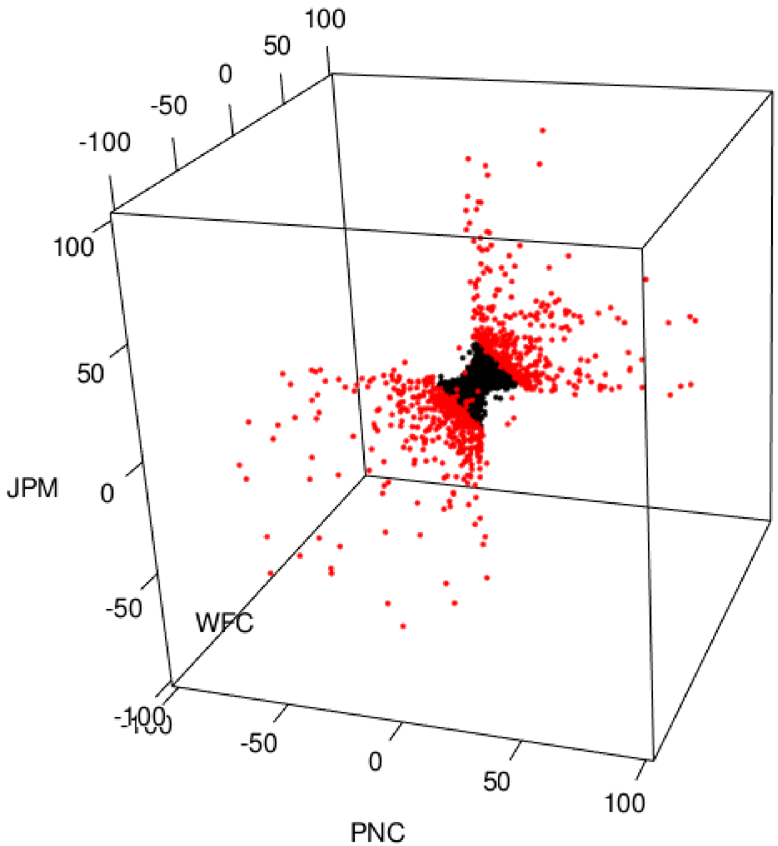
Gaussian underestimates risk by as much as a factor of two.

Outline

- Modeling marginal temporal features
- Studying tail dependence
  - What is tail dependence?
  - Modeling via regular variation
- Tail dependence in banking stock returns
- Portfolio VaR forecasting
- Extensions

“Beyond the positive orthant” (Resnick, 2007)

It seems more elegant to transform to, e.g., Cauchy marginals and model norm exceedances on the cone $[-\infty, \infty] \setminus \{\mathbf{0}\}$.



- Easy extension of regular variation theory
- Difficult in practice
- Extension to $d > 2$ challenging

Appropriate models for high-frequency data

Our findings suggest nearly non-stationary volatility sequences, as the $\widehat{\alpha}_j + \widehat{\beta}_j$ are close to 1.

Long-range forecasts unreliable.

Possible alternative: jump processes (e.g. Fan and Wang, 2007)

Further reading: December 2014 special issue of *Extremes*: Extremes in Finance.

Thanks!

Tang, M. and Weller, G.B. Bivariate tail risk analysis for high-frequency returns via extreme value theory. submitted to *Quantitative Finance*.

Contact: gweller57@gmail.com

Website: grantbweller.com

References:

Fan, J. and Wang, Y. (2007). Multi-scale jump and volatility analysis for high-frequency financial data. *Journal of the American Statistical Association*, 102(480):1349–1362.

Huser, R., Davison, A. C., and Genton, M. G. (2015). A comparative study of likelihood estimators for multivariate extremes. *arXiv preprint arXiv:1411.3448*.

Müller, U. A., Dacorogna, M. M., and Pictet, O. V. (1998). Heavy tails in high-frequency financial data. In Adler, R. J., Feldman, R. E., and Taqqu, M. S., editors, *A Practical Guide to Heavy Tails: Statistical Techniques and Applications.*, pages 55–77. Birkhaüser, Boston.

Resnick, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer, New York.