A measure of dependence for stable distributions

Tuncay Alparslan & John Nolan

American University
Washington, DC, USA

Workshop on Extremes, Copulas and Actuarial Science
Luminy, France
23 Feb 2016

Outline

Spectral measure characterization

We will say $\mathbf{X} \sim \mathbf{S}(\alpha, \Lambda, \delta; j)$, $j = 0, 1$ if its joint characteristic function is given by

$$\phi(\mathbf{u}) = E \exp(i\langle \mathbf{u}, \mathbf{X}\rangle) = \exp\left(- \int_{\mathbb{S}} \omega\left(\langle \mathbf{u}, \mathbf{s}\rangle | \alpha; j\right) \Lambda(d\mathbf{s}) + i\langle \mathbf{u}, \delta\rangle \right),$$

where

$$\omega(t|\alpha; j) = \begin{cases} |t|^{\alpha}[1 + i \,\mathrm{sign}\,(t) \tan \frac{\pi\alpha}{2}(|t|^{1-\alpha} - 1)] & \alpha \neq 1, j = 0 \\ |t|^{\alpha}[1 - i \,\mathrm{sign}\,(t) \tan \frac{\pi\alpha}{2}] & \alpha \neq 1, j = 1 \\ |t|[1 + i \,\mathrm{sign}\,(t) \frac{2}{\pi} \log |t|] & \alpha = 1, j = 0, 1. \end{cases}$$

The 1-parameterization is more commonly used, but discontinuous in α. 0-parameterization is a continuous parameterization.

Projection parameterization

Every one dimensional projection $\langle \mathbf{u}, \mathbf{X} \rangle = u_1 X_1 + u_2 X_2 + \cdots + u_d X_d$ has a univariate stable distribution, with a constant index of stability α and skewness $\beta(\mathbf{u})$, scale $\gamma(\mathbf{u})$ and shift $\delta(\mathbf{u})$ that depend on the direction $\mathbf{u}$.

Projection parameterization

Every one dimensional projection $\langle \mathbf{u}, \mathbf{X} \rangle = u_1 X_1 + u_2 X_2 + \cdots + u_d X_d$ has a univariate stable distribution, with a constant index of stability α and skewness $\beta(\mathbf{u})$, scale $\gamma(\mathbf{u})$ and shift $\delta(\mathbf{u})$ that depend on the direction $\mathbf{u}$.

We will call the functions $\beta(\cdot)$, $\gamma(\cdot)$ and $\delta(\cdot)$ the projection parameter functions. They determine the joint distribution via the Cramér-Wold device, so we can parameterize $\mathbf{X}$ by these projection parameter functions: $\mathbf{X} \sim \mathbf{S}\left(\alpha, \beta(\cdot), \gamma(\cdot), \delta(\cdot); j\right)$, $j = 0$ or $j = 1$.

Projection parameterization

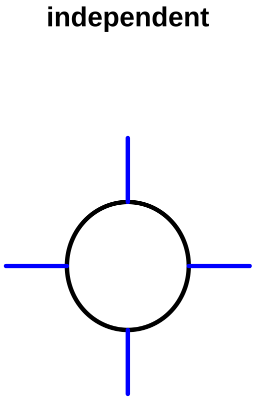
Every one dimensional projection $\langle \mathbf{u}, \mathbf{X} \rangle = u_1 X_1 + u_2 X_2 + \cdots + u_d X_d$ has a univariate stable distribution, with a constant index of stability α and skewness $\beta(\mathbf{u})$, scale $\gamma(\mathbf{u})$ and shift $\delta(\mathbf{u})$ that depend on the direction $\mathbf{u}$.

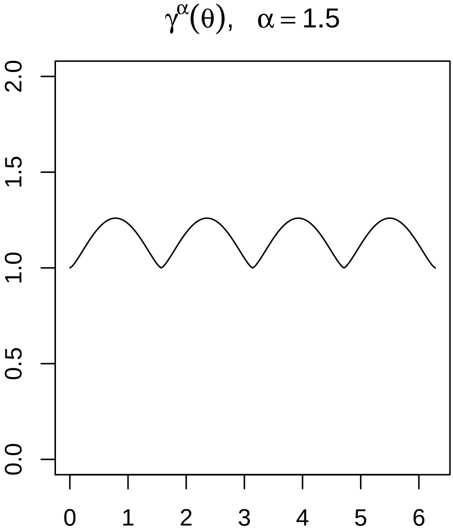
We will call the functions $\beta(\cdot)$, $\gamma(\cdot)$ and $\delta(\cdot)$ the projection parameter functions. They determine the joint distribution via the Cramér-Wold device, so we can parameterize $\mathbf{X}$ by these projection parameter functions: $\mathbf{X} \sim \mathbf{S}\left(\alpha, \beta(\cdot), \gamma(\cdot), \delta(\cdot); j\right)$, $j = 0$ or $j = 1$.

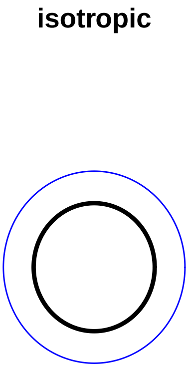
In this section, we will always assume that $d = 2$ and $\mathbf{X}$ has normalized components: $\gamma(1,0) = \gamma(0,1) = 1$.

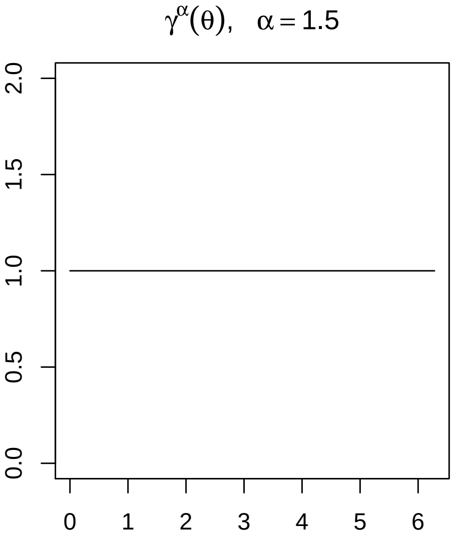
Will sometimes use polar notation: $\gamma(\theta) := \gamma(\cos\theta, \sin\theta)$ to specify a scale funtion on the unit circle.

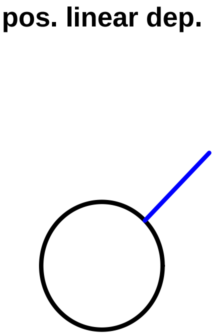
Spectral measure $\Lambda(\cdot)$ and scale function $\gamma(\cdot)$

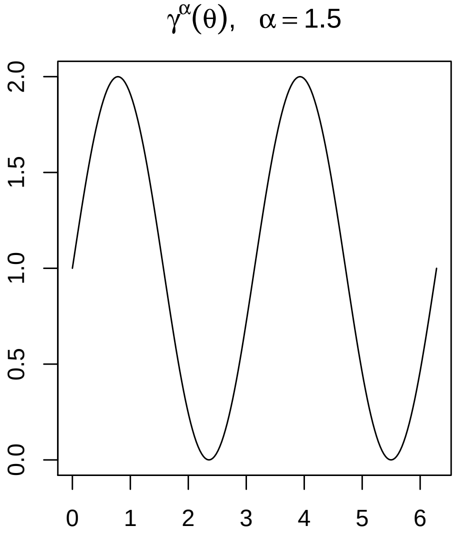




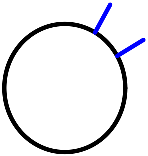

isotropic


$\gamma^{\alpha}(\theta), \quad \alpha = 1.5$
2.0
1.5
1.0
0.5
0.0
0
1
2
3
4
5
6

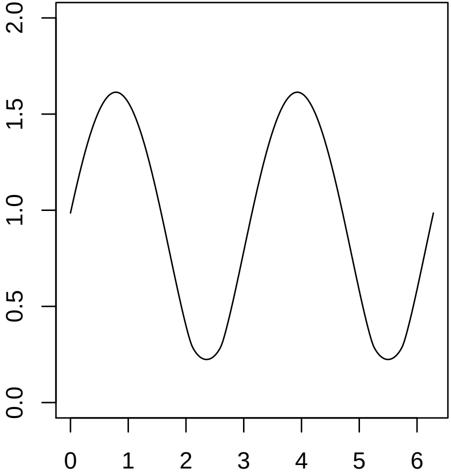

pos. linear dep.


$\gamma^{\alpha}(\theta), \quad \alpha = 1.5$
0.0
0.5
1.0
1.5
2.0
0
1
2
3
4
5
6

pos. associated



$\gamma^{\alpha}(\theta), \quad \alpha = 1.5$



Outline

Definition

Set $\gamma_{\perp}(\mathbf{u}) = (|u_1|^{\alpha} + |u_2|^{\alpha})^{1/\alpha}$ (independence), $p \in [1, \infty]$

$$\eta_p = \eta_p(X_1, X_2) = \|\gamma^{\alpha}(u_1, u_2) - \gamma_{\perp}^{\alpha}(u_1, u_2)\|_{L^p(\mathbb{S}, d\mathbf{u})}. \quad (1)$$

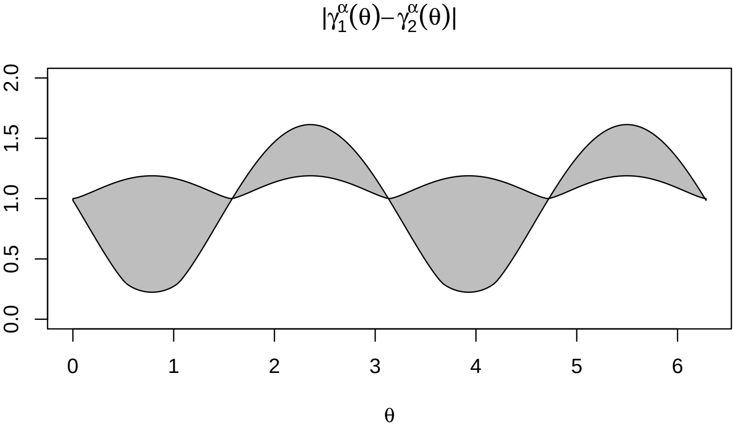
Here $d\mathbf{u}$ is (unnormalized) surface area on $\mathbb{S}$.

Definition

Set $\gamma_\perp(\mathbf{u}) = (|u_1|^\alpha + |u_2|^\alpha)^{1/\alpha}$ (independence), $p \in [1, \infty]$

$$\eta_p = \eta_p(X_1, X_2) = \|\gamma^\alpha(u_1, u_2) - \gamma_\perp^\alpha(u_1, u_2)\|_{L^p(\mathbb{S}, d\mathbf{u})}. \quad (1)$$

Here $d\mathbf{u}$ is (unnormalized) surface area on $\mathbb{S}$.



Properties of η_p

- $\mathbf{X}$ has independent components if and only if $\eta_p = 0$ for some (every) $p \in [1, \infty]$.
- α can be any value in $(0, 2)$ and $\mathbf{X}$ can have symmetric or non-symmetric components, and it can be centered or shifted.
- η_p is symmetric: $\eta_p(X_1, X_2) = \eta_p(X_2, X_1)$.
- η_p measures how far the scale function of $\mathbf{X}$ is from the scale function of a stable r. vector with independent components: when $\mathbf{X}$ is symmetric, earlier work shows

$$\sup_{\mathbf{x} \in \mathbb{R}^2} |f(\mathbf{x}) - f_\perp(\mathbf{x})| \le k_\alpha \|\gamma(\cdot) - \gamma_\perp(\cdot)\|.$$

Properties of η_p (continued)

- The p-norm in (1) is evaluated as an integral over the unit circle $\mathbb{S}$, not all of $\mathbb{R}^2$. In polar coordinates,

$$\eta_p = \left(2 \int_0^{\pi} |\gamma^{\alpha}(\cos\theta, \sin\theta) - \gamma^{\alpha}_{\perp}(\cos\theta, \sin\theta)|^p \, d\theta \right)^{1/p}, \quad (2)$$

  where the interval of integration has been reduced by using the fact that $\gamma(\cdot)$ is π-periodic
- $\eta_p \geq 0$ by definition, not measuring positive/negative dependence, just distance from independence. Don't think there is a general way of assigning a sign, e.g. rotate the indep. components case by $\pi/4$ and the resulting distribution bunches around both the lines $y = x$ and $y = -x$ for large values of $|\mathbf{X}|$.

Properties of η_p (continued)

- The definition makes sense in the Gaussian case: when $\alpha = 2$, the scale function for a bivariate Gaussian distribution with correlation ρ is $\gamma(\mathbf{u})^2 = 1 + 2\rho u_1 u_2$ and $\gamma_\perp = 1$. Then $\eta_p^p = |2\rho|^p \int_{\mathbb{S}} |u_1 u_2|^p d\mathbf{u}$, so $\eta_p = k_p |\rho|$.
- In elliptically contoured/sub-Gaussian case, can get an integral expression that can be evaluated numerically.
- Multivariate stable $\mathbf{X} = (X_1, \ldots, X_d)$ has mutually independent components if and only if all pairs are independent, so the components of $\mathbf{X}$ are mutually independent if and only if $\eta_p(X_i, X_j) = 0$ for all $i > j$.

Covariation in terms of $\gamma(\cdot)$

For $\alpha > 1$, the covariation is

$$[X_1, X_2]_\alpha = \int_{\mathbb{S}} s_1 s_2^{<\alpha-1>} \Lambda(d\mathbf{s})$$

Covariation in terms of $\gamma(\cdot)$

For $\alpha > 1$, the covariation is

$$[X_1, X_2]_\alpha = \int_{\mathbb{S}} s_1 s_2^{<\alpha-1>} \Lambda(d\mathbf{s}) = \frac{1}{\alpha} \left. \frac{\partial \gamma^\alpha(u_1, u_2)}{\partial u_1} \right|_{(u_1=0, u_2=1)}.$$

Thus the covariation depends only on the behavior of $\gamma(\cdot, \cdot)$ near the point $(1, 0)$. If X_1 and X_2 are independent, then $[X_1, X_2]_\alpha = 0$; but the converse is false.

Covariation in terms of $\gamma(\cdot)$

For $\alpha > 1$, the covariation is

$$[X_1, X_2]_\alpha = \int_{\mathbb{S}} s_1 s_2^{<\alpha-1>} \Lambda(d\mathbf{s}) = \frac{1}{\alpha} \left. \frac{\partial \gamma^\alpha(u_1, u_2)}{\partial u_1} \right|_{(u_1=0, u_2=1)} .$$

Thus the covariation depends only on the behavior of $\gamma(\cdot, \cdot)$ near the point $(1, 0)$. If X_1 and X_2 are independent, then $[X_1, X_2]_\alpha = 0$; but the converse is false.

Short diversion: when can covariation be 0?
If Λ_1 is any measure supported on $Q_1 \cup Q_3$, then covariation ≥ 0.
If Λ_2 is any measure supported on $Q_2 \cup Q_4$, then covariation ≤ 0.
Covariation of $c_1\Lambda_1 + c_2\Lambda_2 = c_1$ covariation $\Lambda_1 + c_2$ covariation Λ_2.
So by choosing $c_1, c_2 > 0$ appropriately we can get 0 covariation with many, many different measures.

Co-difference in terms of $\gamma(\cdot)$

The co-difference is defined for symmetric α-stable vectors, and can be written as

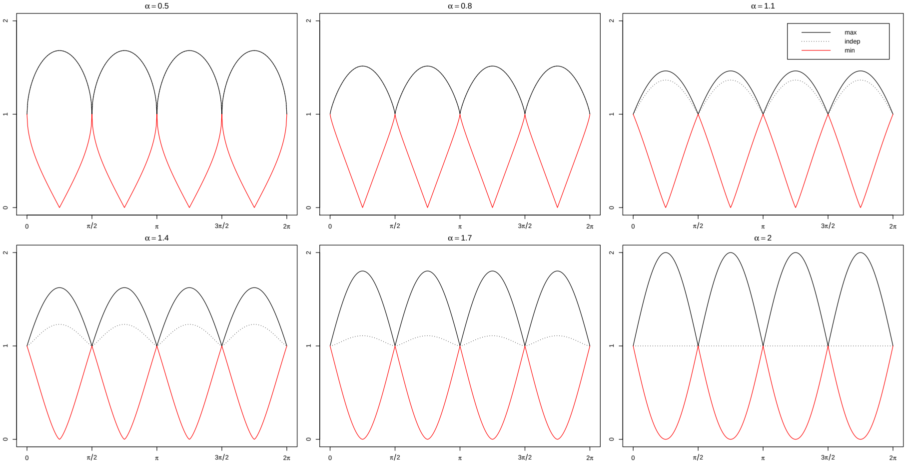
$$\tau = \gamma^{\alpha}(1,0) + \gamma^{\alpha}(0,1) - \gamma^{\alpha}(1,-1),$$

for any $\alpha \in (0,2)$. If X_1 and X_2 are independent, then $\tau = 0$, but again the converse is false.

Short diversion: when can co-difference be 0? Many ways, as on previous page when $\alpha > 1$; when $\alpha \leq 1$, can only have $\tau \geq 0$.

Envelope of scale function $\gamma^{\alpha}(\cdot)$

Find $\gamma^{\alpha}_{\min}(\cdot) = \min\ \gamma^{\alpha}(\cdot)$, $\gamma^{\alpha}_{\max}(\cdot) = \max\ \gamma^{\alpha}(\cdot)$ like Pickand's function



max is known and sharp; min is conjectured (and achieved)

Outline

Sample measure $\widehat{\eta}_2$

Use max. likelihood estimation of the marginals and set $\widehat{\alpha} = (\alpha_1 + \alpha_2)/2$, normalize each component.

For angles $0 \leq \theta_1 < \theta_2 < \cdots < \theta_m \leq \pi$, define $\widehat{\gamma}_j = \widehat{\gamma}(\cos\theta_j, \sin\theta_j) =$ ML estimate of the scale of the projected data set $\langle \mathbf{Y}_i, (\cos\theta_j, \sin\theta_j) \rangle$, $i = 1, \ldots, n$

Sample measure $\widehat{\eta}_2$

Use max. likelihood estimation of the marginals and set $\widehat{\alpha} = (\alpha_1 + \alpha_2)/2$, normalize each component.

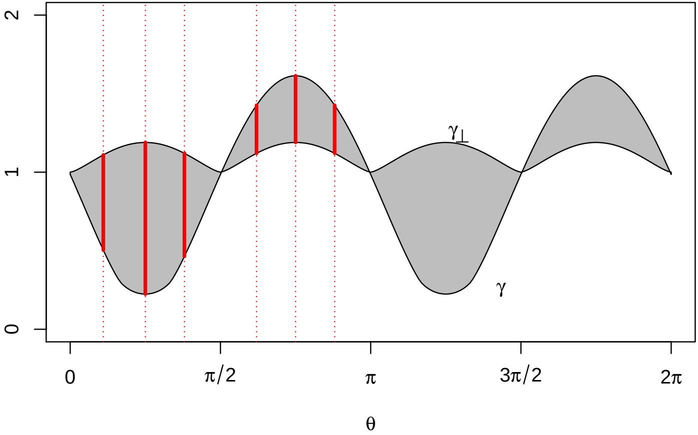
For angles $0 \leq \theta_1 < \theta_2 < \cdots < \theta_m \leq \pi$, define $\widehat{\gamma}_j = \widehat{\gamma}(\cos\theta_j, \sin\theta_j) =$ ML estimate of the scale of the projected data set $\langle \mathbf{Y}_i, (\cos\theta_j, \sin\theta_j)\rangle$, $i = 1, \ldots, n$

Define

$$\widehat{\eta}_2 = \left(\sum_{j=1}^{m} \left(\widehat{\gamma}_j^{\widehat{\alpha}} - \gamma_{\perp,j}^{\widehat{\alpha}} \right)^2 \right)^{1/2},$$

where $\gamma_{\perp,j}^{\widehat{\alpha}}$ is the scale in direction θ_j when components are independent.

Uniform grid with $m = 6$ directions



Suggest uniform grid in first and second quadrant that avoid 0, $\pi/2$, π

Get critical values by simulation, depends on α, the skewness parameters β_1 and β_2 of the marginals, grid size and sample size n.

The power to detect dependence increases as the grid size increases, but only for a while. The power plateaus near 5 points in each quadrant.

Fast approximation to critical values based on $\chi^2(1)$ distribution.

Power calculation via simulation, $\alpha = 1.5$, 5 grid points per quadrant, 1000 simulations

n	isotropic	indep. $\circlearrowleft \pi/4$	indep. $\circlearrowleft \pi/8$	indep. $\circlearrowleft \pi/16$	exact linear dep.
25	0.191	0.322	0.243	0.213	1
50	0.223	0.624	0.381	0.183	1
100	0.344	0.918	0.644	0.214	1
200	0.636	0.998	0.937	0.440	1
300	0.874	1	0.997	0.627	1
400	0.960	1	1	0.791	1
500	0.989	1	1	0.893	1

Outline

1. Multivariate stable
2. Dependence measure η_p
3. Sample dependence measure $\widehat{\eta}_p$
4. Related topics

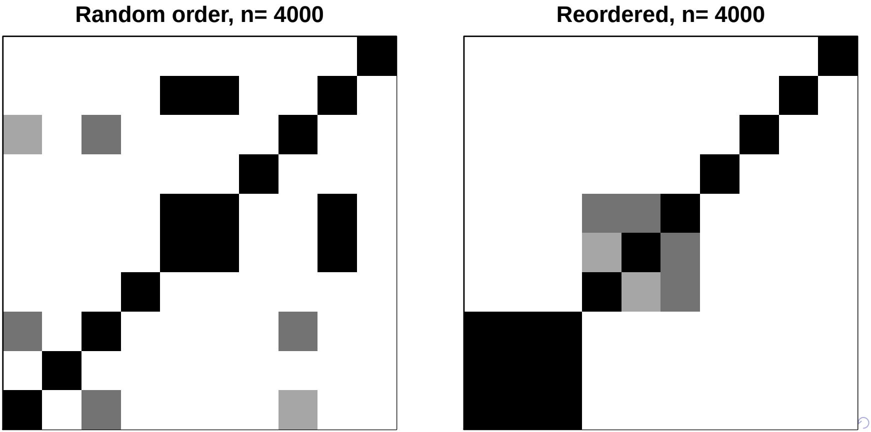
Dimension $d > 2$

Generated $d = 10$ dim data: 3 components were strongly dependent, 3 were somewhat dependent, other 4 were independent. Then permuted the components. Can we find the structure?

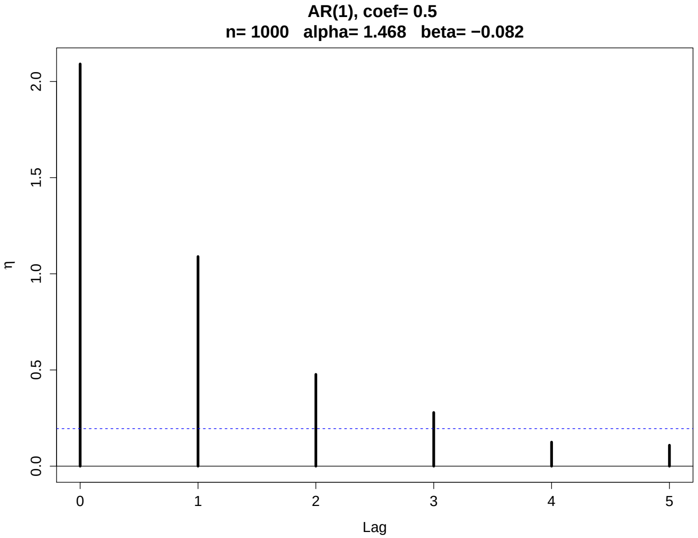
Dimension $d > 2$

Generated $d = 10$ dim data: 3 components were strongly dependent, 3 were somewhat dependent, other 4 were independent. Then permuted the components. Can we find the structure?

Compute pairwise $\widehat{\eta}_2$ and plot as a grayscale image (left), then cluster (right).



Time series - plot $\widehat{\eta}_2(X_i, X_{i+j})$ similar to ACF plot

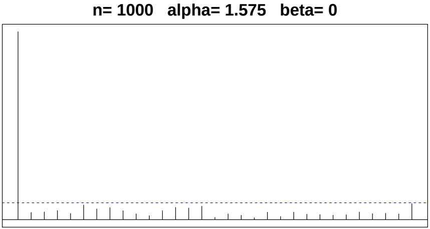


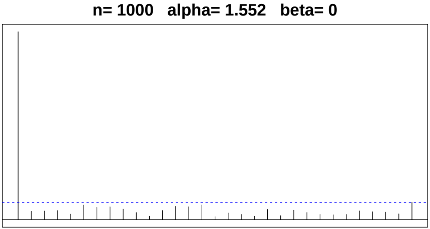
model selection - see dependence in an AR(1) simulated time series

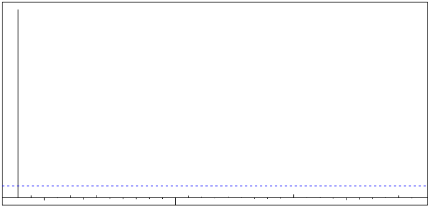
Time series - robustness

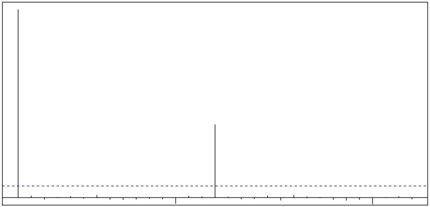
ACF is sensitive to extremes

$\widehat{\eta}^2$ ACF









$\widehat{\eta}^2$ with change at lag 15 ACF with one change at lag 15

Székely and Rizzo (2007, 2009) defined a distance covariance based on weighted L^2 difference of joint empirical characteristic function and product of marginal empirical distribution function. It is very general, characterizes independence. Simulations shows that it also works well with bivariate stable data. In fact, it is a bit more powerful than the above η_2. (We do not understand this.)

Domain of attraction modifications Have to use bootstrap samples to compute critical values, noticeably less power.

Similar measure of dependence for multivariate extreme value laws - difference between tail dependence function and the one corresponding to independence.