Estimating extremal dependence using B-splines

Christian Genest & Johanna G. Nešlehová
McGill University, Montréal, Canada

Joint work with A. Bücher and D. Sznajder

Outline

1. Extreme-value copulas
2. A new diagnostic tool: The A-plot
3. An intrinsic estimator based on B-splines
4. Simulation results and data illustration

1. Extreme-value copulas

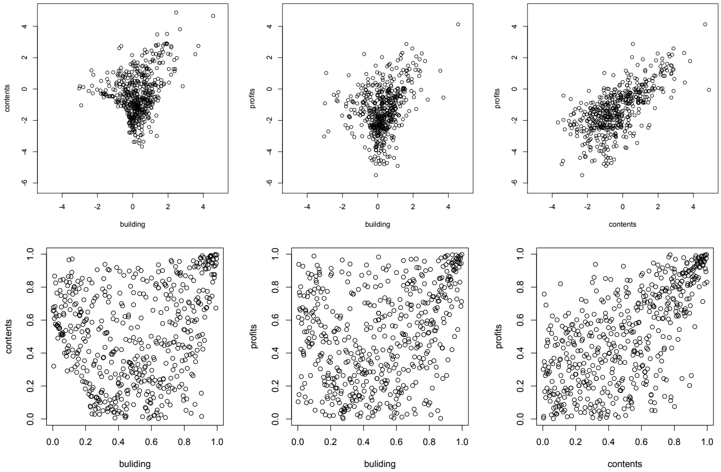
Extreme-value copulas are the asymptotic dependence structures of component-wise maxima.

Modeling joint extremes is a key issue in risk management. A classical example (McNeil 1997) is

- X: damage to buildings
- Y: loss to contents
- Z: loss of profits

from losses of 1M DKK to the Copenhagen Reinsurance company arising from fire claims between 1980 and 1990.

Danish Fire Insurance Data



Original data on the log-scale (top) and
pairs of normalized ranks (bottom)

Analytic form (Pickands 1981)

All extreme-value copulas are of the form

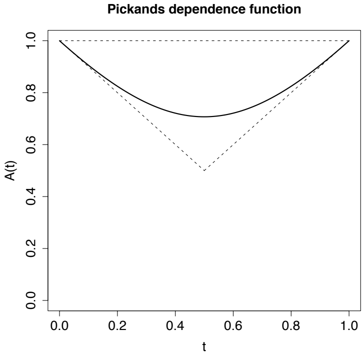
$$C(u,v) = \exp\left[\ln(uv) A\left\{\frac{\ln(v)}{\ln(uv)}\right\}\right],$$

where $A : [0,1] \to [0,1]$ is convex and

$$\forall_{t \in [0,1]} \quad \max(t, 1-t) \leq A(t) \leq 1.$$

The function A is called the Pickands dependence function.

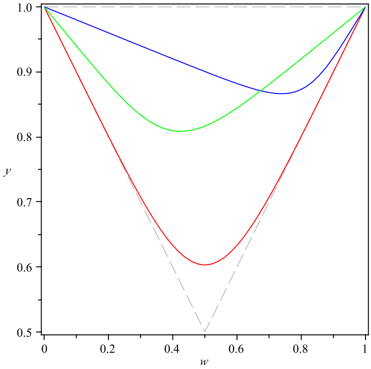
A generic Pickands dependence function



Its tail dependence coefficient:

$$\lambda = \lim_{u \uparrow 1} \Pr\{X > F^{-1}(u) | Y > G^{-1}(u)\} = 2\{1 - A(1/2)\}.$$

Parametric examples



Symmetric and asymmetric Galambos extreme-value copulas

Issues of (our) current interest

Suppose $(X_1, Y_1), \ldots, (X_n, Y_n)$ is a random sample from

$$H(x, y) = C\{F(x), G(y)\},$$

where F, G are continuous and C is a copula.

- How can one decide whether C is extreme-value?
- If an extreme-value copula model is appropriate, how can A be estimated intrinsically?

  That is, we want $\hat{A}_n$ to be convex and such that

$$\forall_{t \in [0,1]} \quad \max(t, 1-t) \leq \hat{A}_n(t) \leq 1.$$

2. A new diagnostic tool: The A-plot

Consider the transformation $T : (0,1)^2 \to (0,1)$ defined by

$$T(u,v) = \frac{\ln(v)}{\ln(uv)}\,.$$

If C is an extreme-value copula, then

$$\frac{\ln(v)}{\ln(uv)} = t \quad \Rightarrow \quad A(t) = \frac{\ln\{C(u,v)\}}{\ln(uv)}\,.$$

Transformation

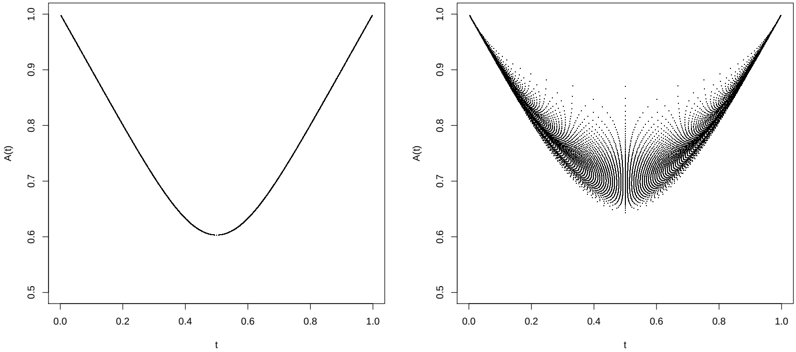
Define the set

$$\mathcal{S} = \left\{ \left(t = \frac{\ln(v)}{\ln(uv)}, A(t) = \frac{\ln\{C(u,v)\}}{\ln(uv)} \right) : u, v \in (0,1) \right\}.$$

When C is an extreme-value copula, the graph of $\mathcal{S}$ coincides with the Pickands dependence function.

When C is not extreme, this relationship breaks down!

Plots of the graph of S



Galambos(3) vs Gaussian(0.7)

The A-plot: A diagnostic tool

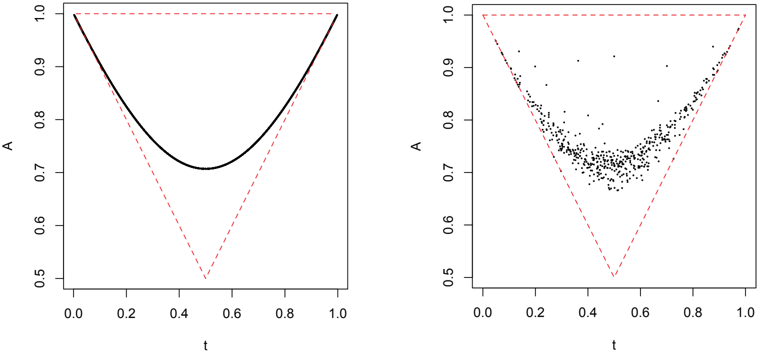
Plot the pairs $(T_1, Z_1), \ldots, (T_n, Z_n)$, where for each $i \in \{1, \ldots, n\}$,

$$T_i = \frac{\ln(\hat{V}_i)}{\ln(\hat{U}_i \hat{V}_i)}, \quad Z_i = \frac{\ln\{\hat{C}_n(\hat{U}_i, \hat{V}_i)\}}{\ln(\hat{U}_i \hat{V}_i)}.$$

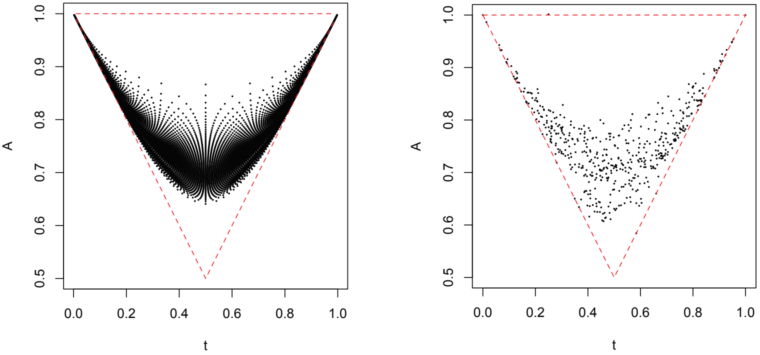
Extreme dependence appears reasonable if the points fall close to a convex curve.

It is a helpful complement to formal tests of extremeness (some of which are inconsistent, e.g., Ghoudi et al. 1998).

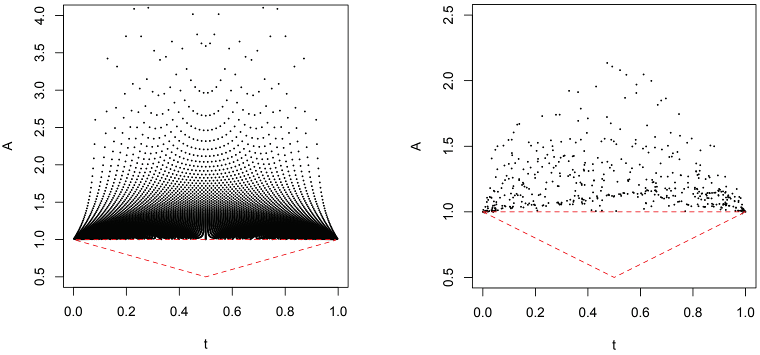
Example 1: Gumbel copula with $\tau = .5$



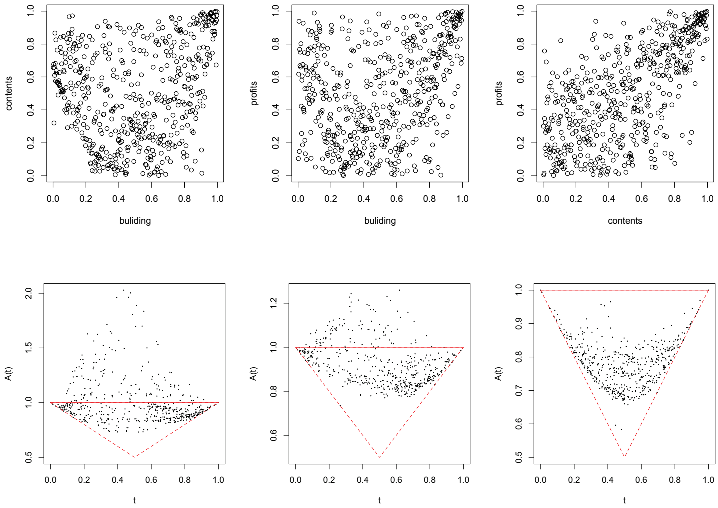
Example 2: Gaussian copula with $\tau = .5$



Example 3: Clayton copula with $\tau = -.25$



Danish Fire Insurance Data



$$(X, Y), \quad (X, Z), \quad (Y, Z)$$

Thresholding

The A-plot can be adapted to help see whether C is in the max-domain of attraction of an extreme-value copula, i.e.,

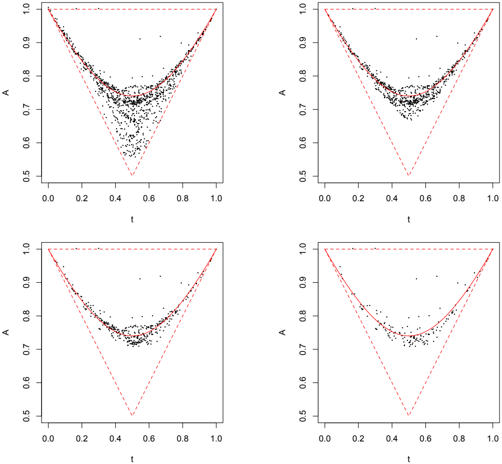
$$\lim_{\ell \to \infty} C^{\ell}(u^{1/\ell}, v^{1/\ell}) = C_0(u, v).$$

This condition implies that for sufficiently large $w \in (0, 1)$,

$$C(u, v) \approx C_0^{1/\ell}(u^{\ell}, v^{\ell}) = C_0(u, v)$$

for all $u, v > w$; see, e.g., Ledford & Tawn (1996).

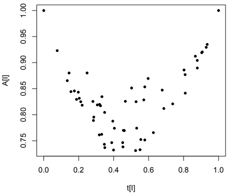
Illustration (Student t_2 with $\rho = 0.7$)



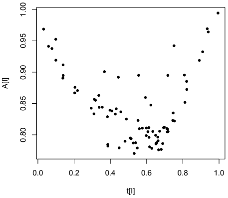
Threshold $w \in \{0, .25, .5, .75\}$, $n = 1000$

Illustration for the Danish data

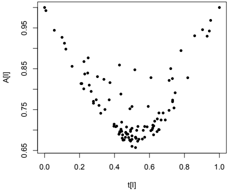
Buildings-Contents



Buildings-Profits



Profits-Contents



For all three pairs of risks, the probability that one loss exceeds a high threshold, given that the other loss has exceeded it, is about $\lambda_u \approx 1/2$!

3. An intrinsic estimator based on B-splines

Many estimators of A have been proposed so far; see, e.g.,

- Pickands (1981), Capéraà & Fougères & Genest (1997), Genest & Segers (2009)
- Deheuvels (1991), Hall & Tajvidi (2000), Jiménez, Villa-Diharce & Flores (2001), Segers (2007)
- Zhang, Wells & Peng (2007), Gudendorf & Segers (2012)
- Bücher, Dette & Volgushev (2011), Berghaus, Bücher & Dette (2012)
- Guillotte & Perron (2008), Guillotte, Perron & Segers (2011), Guillotte & Perron (2012)
- Ucer & Ahmadabadi (Bernstein polynomials, in progress)

A common limitation

Most of these estimators are not intrinsic "off the bat", i.e., one of these conditions is violated:

- $\hat{A}_n(0) = \hat{A}_n(1) = 1$;
- $\forall_{t \in [0,1]}\ \max(t, 1-t) \leq \hat{A}_n(t) \leq 1$;
- $\hat{A}_n$ is convex.

One can resort, e.g., to projections (Fils-Villetard, Guillou & Segers 2008), but this adds complexity.

Intrinsic estimators are not needed for diagnostics but essential to simulate from the corresponding extreme-value copula.

The new procedure

Cormier et al. (2014) propose to estimate A by fitting a B-spline of order $m = 3$ through the A-plot, viz.

$$\hat{A}_n = \sum_{j=1}^{m+k} \hat{\beta}_j \phi_{j,m},$$

where $\hat{\beta}_1, \ldots, \hat{\beta}_{m+k}$ are suitably selected scalars and

$$\phi_{1,m}, \cdots, \phi_{m+k,m}$$

denote the B-spline basis of order $m \geq 3$ with k interior knots.

Cox–de Boor recursion formula

To construct the basis $\phi_{1,m}, \ldots, \phi_{m+k,m}$ of order m with interior knots

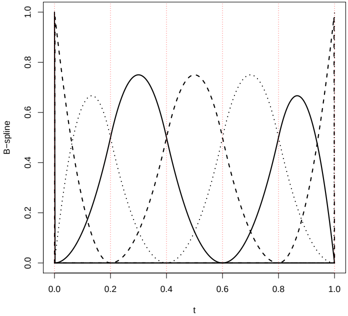
$$0 < \tau_{m+1} < \cdots < \tau_{m+k} < 1,$$

set $\tau_1 = \cdots = \tau_m = 0$, $\tau_{m+k+1} = \cdots = \tau_{2m+k} = 1$.

1. For $j \in \{1, \ldots, k + 2m - 1\}$, let $\phi_{j,1} = \mathbf{1}_{[\tau_j, \tau_{j+1})}$.
2. For $\ell \in \{2, \ldots, m\}$, $j \in \{1, \ldots, k + 2m - \ell\}$, let

$$\phi_{j,\ell}(t) = \frac{t - \tau_j}{\tau_{j+\ell-1} - \tau_j}\, \phi_{j,\ell-1}(t) + \frac{\tau_{j+\ell} - t}{\tau_{j+\ell} - \tau_{j+1}}\, \phi_{j+1,\ell-1}(t).$$

Illustration: Third-order B-spline basis



This basis has $k = 4$ equally-spaced interior knots and consists of $m + k = 7$ B-spline polynomials of degree $m - 1 = 2$.

Fitting procedure

Assume that for unknown $\beta = (\beta_1, \ldots, \beta_{m+k})^\top$,

$$\forall_{t \in [0,1]} \quad A(t) = \sum_{j=1}^{m+k} \beta_j \phi_{j,m}(t) = \beta^\top \Phi(t),$$

where $\Phi(t) = (\phi_{1,m}(t), \ldots, \phi_{m+k,m}(t))^\top$.

View this as a regression $\mathrm{E}(Z) = \beta^\top X$ for which we have data

$$(X_1, Y_1) = (\Phi(T_1), Z_1), \ldots, (X_n, Y_n) = (\Phi(T_n), Z_n).$$

with $T_i = \ln(\hat{V}_i)/\ln(\hat{U}_i \hat{V}_i)$, $Z_i = \ln\{\hat{C}_n(\hat{U}_i, \hat{V}_i)\}/\ln(\hat{U}_i \hat{V}_i)$.

Penalized absolute-deviation (L_1) criterion

Given the pairs $(T_1, Z_1), \ldots, (T_n, Z_n)$, find

$$\hat{\beta}_n = \mathrm{argmin}_{B \in \mathcal{B}} \, ||Z - \beta^\top \Phi(T)||_1 + \lambda_n ||\beta^\top \Phi''||_\infty,$$

where $\mathcal{B}$ is the set of vectors $\beta \in \mathbb{R}^{m+k}$ such that

(A) $\beta^\top \Phi(0) = \beta^\top \Phi(1) = 1$;

(B) $\beta^\top \Phi''(\tau_j) \geq 0$ for every $j \in \{1, \ldots, k\}$;

(C) $|\hat{A}'_n(t)| \in [0, 1]$ at $t = 0$ and 1.

Technical details

(A)–(C) guarantee that $\hat{A}_n$ is intrinsic if $m \in \{3, 4\}$ because $\hat{A}_n''$ is then linear between the knots. Hence

$$\forall_{j \in \{1,\ldots,m\}} \; \beta^\top \Phi''(\tau_j) \geq 0 \quad \Rightarrow \quad \forall_{t \in (0,1)} \; \hat{A}_n''(t) \geq 0.$$

The penalization term $\lambda_n ||\beta^\top \Phi''||_\infty$ is needed to make the solution smooth when the knots are unknown (always!).

Minimization is performed over a large number of equally spaced empirical quantiles derived from $T_1, \ldots, T_n$.

Asymptotic result (Bücher et al.)

Suppose that $\lambda_n = o(\sqrt{n})$. Then, as $n \to \infty$,

$$\hat{B}_n = \sqrt{n}\,(\hat{\beta}_n - \beta) \rightsquigarrow \check{B} = \operatorname*{argmin}_{B \in \mathbb{R}^{m+k}} \left\{ \int_{(0,1)^2} J_B(u,v)\mathrm{d}C(u,v) \right\},$$

where

$$J_B(u,v) = \left| \frac{\mathbb{C}_C(u,v)}{\ln(uv)C(u,v)} - B^\top \Phi\left\{ \frac{\ln(v)}{\ln(uv)} \right\} \right| - \left| \frac{\mathbb{C}_C(u,v)}{\ln(uv)C(u,v)} \right|.$$

Asymptotic result (part 2)

Given that for all $t \in (0, 1)$,

$$\hat{A}_n(t) = \beta_n^\top \Phi(t),$$

one can conclude that if $\lambda_n = o(\sqrt{n})$, then as $n \to \infty$,

$$\mathbb{A}_n = \sqrt{n}\,(\hat{A}_n - A) \rightsquigarrow \check{B}^\top \Phi$$

in $\ell^\infty([0, 1])$ equipped with the sup-norm.

Bonus: Spectral distribution estimation

For the spectral distribution L of an extreme-value copula,

$$A(t) = 1 - t + 2\int_0^t L(w)\mathrm{d}w \quad \Leftrightarrow \quad A'(t) = 2L(t) - 1$$

when A' exists; see, e.g., Einmahl & Segers (2009).

$$\hat{L}_n(t) = \{\hat{A}_n'(t) + 1\}/2, \quad \hat{L}_n'(t) = \hat{A}_n''(t)/2,$$

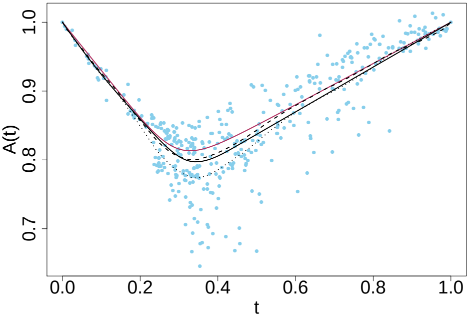
are easily computed and $\hat{A}_n'(0)$ and $\hat{A}_n'(1)$ estimate the spectral masses at the end-points.

Computer implementation ($m = 3$)

- ✓ The procedure is coded in R using the "COBS" package.
- ✓ It is fully automated and only requires the user to define the constraints and the number of knots.
- ✓ From experience, between 10 and 15 knots suffice to capture the complexity of the data.
- ✓ The derivatives are calculated using the "FDA" package after knot and coefficient abstractions from "COBS".

Contrasting $m = 3$ vs $m = 4$

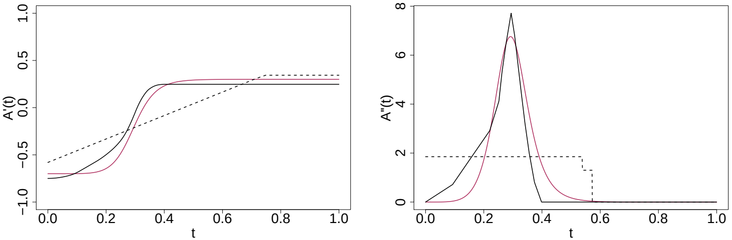
Little difference when estimating A only.



Asymmetric logistic model with $\alpha = .3$, $\beta = .7$, $\theta = 6$
B-spline (solid), Pickands (dotted), CFG (dashed), $n = 400$

Contrasting $m = 3$ vs $m = 4$ (cont'd)

Bigger difference when estimating L, and especially L'.



B-splines estimates of A' (left) and A'' (right)
$m = 3$ (dashed) and $m = 4$ (solid)
Same data, same set of knots

4. Simulation results and data illustration

The B-spline estimators of L with $m = 3$ and $m = 4$ were compared to the estimator of Einmahl & Segers (2009).

- ✓ 9 extreme-value and 5 other copulas;
- ✓ various degrees of asymmetry and dependence;
- ✓ various sample sizes and $N = 1000$ repetitions.

Performance measure used:

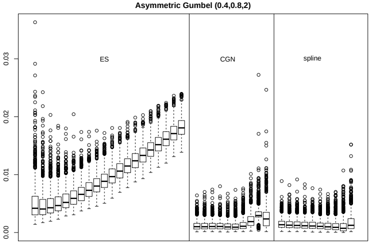
$$D_n = \frac{1}{n} \sum_{i=1}^{n} \{L(T_i) - \hat{L}_n(T_i)\}^2.$$

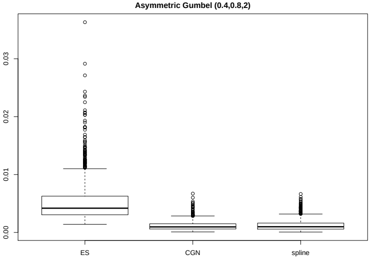
Clarifications and conclusions

- ✓ The ES estimator uses thresholding; 20 values were used: `w = seq(10,88,4)`.
- ✓ For fairness, the B-spline estimators were also applied to thresholded data; 10 levels used: `w = seq(0,.8,10)`.
- ✓ In total: $N = 1000$ values of D_n for 40 estimators:

$$20 \text{ ES}, \quad 10 \text{ CGN } (m = 3), \quad 10 \text{ BGNS } (m = 4).$$

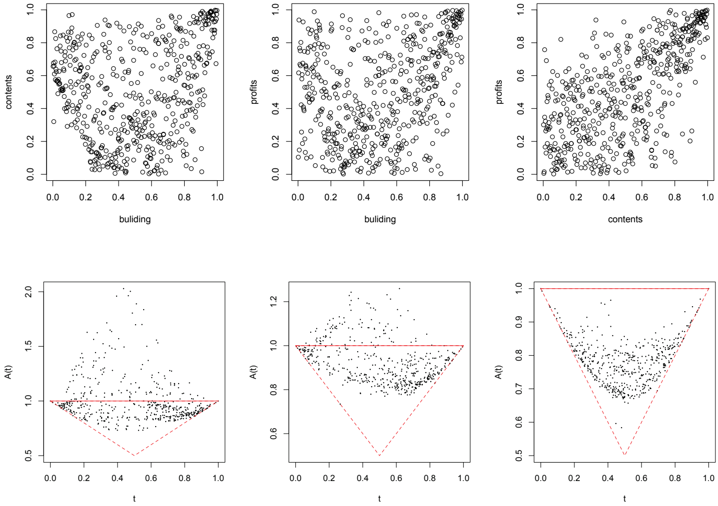
Typical outcome (Asymmetric Gumbel)





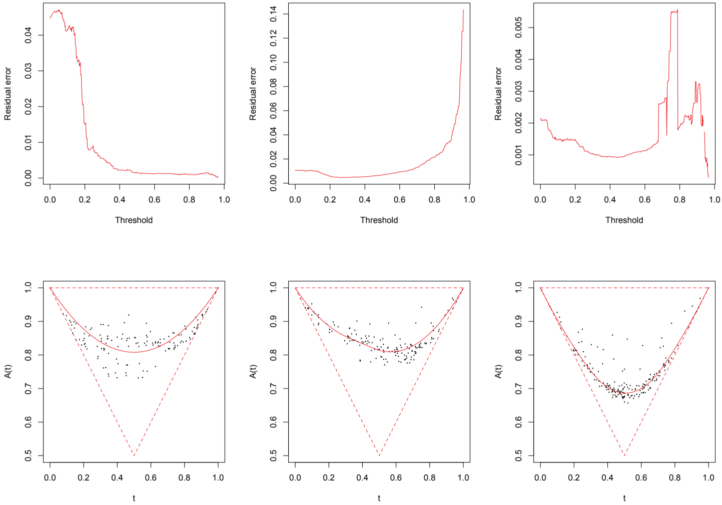
Conclusion: The CGN and BCJS estimators are typically superior to the ES estimator.

Danish Fire Insurance Data



$$(X, Y), \quad (X, Z), \quad (Y, Z)$$

Thresholding and final estimates



$$(X,Y), \quad (X,Z), \quad (Y,Z)$$

Take-home message

- ✓ The A-plot is useful for detecting extreme-value dependence.
- ✓ An intrinsic estimator of A can be based on B-splines.
- ✓ B-splines of order $m = 3$ are adequate for estimating A (off-the-shelf solution with COBS and FDA packages).
- ✓ B-splines of order $m = 4$ yield better estimates of L and L' than the approach of Einmahl & Segers (2009).
- ✓ Asymptotic theory is available and non-extreme data can be handled via thresholding (no asymptotics in support).