Asymptotic Results for Conditional Measures of Association in the Classical Risk Model

Alexandru V. Asimit

Cass Business School, City University

Yiqing Chen

University of Liverpool

8^{th} Conference in Actuarial Science & Finance, Samos

29 May - 01 June, 2014

Outline

- Initial Motivation
- Background
- Main Results
- Future work
- Conclusions

Initial Motivation

Individual claim amount X with df F and number of claims N

Let S_T and $X_T^{(1)} \geq X_T^{(2)} \geq \ldots$

Aim: Understand the strength of dependence of concomitant extreme events for

$$\left(S_T, X_T^{(1)} | X_T^{(1)} > t\right) \quad \text{and} \quad \left(X_T^{(1)}, X_T^{(2)} | X_T^{(2)} > t\right)$$

How: Copula approach or finding some measures of association.

Background: Tail behaviour of X

Assume $x_F = \infty$. Denote $\bar{F} = 1 - F$.

- $F \in \mathcal{L}$ if $\lim_{t \to \infty} \dfrac{\bar{F}(t+x)}{\bar{F}(t)} = 1, \ x \in \Re$;

- $F \in \mathcal{S}$ if $\lim_{t \to \infty} \dfrac{\Pr(X_1 + X_2 > t)}{\Pr(X_1 > t)} = 2$;

- $F \in \mathcal{R}(\alpha)$ with $\alpha > 0$ if $\lim_{t \to \infty} \dfrac{\bar{F}(tx)}{\bar{F}(t)} = x^{-\alpha}, \ x > 0$;

- Note that $\bigcup_{\alpha > 0} \mathcal{R}(\alpha) \subset \mathcal{S} \subset \mathcal{L}$

- Pareto, InverseGamma, LogNormal, some Weibull etc

- $X_1, X_2, \ldots$ *iid* rvs with common *df* F
- $M_n = \max(X_i, i = 1, \ldots, n)$

If there exist a rv Y with nondegenerate df G and *normalizing constants* $a_n > 0$, b_n such that $a_n M_n + b_n \stackrel{w}{\to} Y$, then G belongs to the type of the distribution

$$H_\xi(x) = \begin{cases} \exp\left\{-(1+\xi x)^{-1/\xi}\right\}, & 1+\xi\ x > 0, \quad \xi \neq 0 \\ \exp\{-e^{-x}\}, & x \in \Re, \quad \xi = 0 \end{cases},$$

and we write $F \in MDA(G)$. H_ξ is known as the *generalized extreme value distribution*.

- $\Phi_\alpha(x) := H_{1/\alpha}(\alpha(x-1))$ is the standard Fréchet distribution
- $\Psi_\alpha(x) := H_{-1/\alpha}(\alpha(x+1))$ is the standard Weibull distribution
- $\Lambda(x) := H_0(x)$ is the standard Gumbel distribution.

Background: Generalized Pareto Distribution

- $F \in MDA(H_\xi)$ if and only if there exists a positive, measurable function $a(\cdot)$ such that

$$\lim_{t \uparrow x_F} \frac{\bar{F}(t + xa(t))}{\bar{F}(t)} = \begin{cases} (1+\xi x)^{-1/\xi}, \ 1+\xi & x > 0, \quad \xi \neq 0 \\ e^{-x}, & x \in \Re, \quad \xi = 0 \end{cases},$$

where x_F is the right endpoint of F

- *Generalized Pareto distribution* (GPD)

- The scaled excesses over high thresholds $\sim$ GPD

Background: A comment on tail behaviour

- If $X, Y \in \mathcal{S}$ and/or $\mathcal{L}$ are "independent" in the tail then

$$\lim_{t\to\infty} \frac{\Pr(X + Y > t)}{\Pr(X > t) + \Pr(Y > t)} = 1;$$

- If $X, Y \in MDA(\Lambda)$ are "positively dependent" in the tail then

$$\lim_{t\to\infty} \frac{\Pr(X + Y > 2t)}{\Pr(X > t) + \Pr(Y > t)} = \text{non-trivial constant.}$$

Background: Measures of association

Let $(Y_i, Z_i), i = 1, 2, 3$ be 3 iid copies of (Y, Z) with continuous marginals.

- Kendall's tau

  $$\tau = \Pr((Y_1 - Y_2)(Z_1 - Z_2) > 0) - \Pr((Y_1 - Y_2)(Z_1 - Z_2) < 0);$$

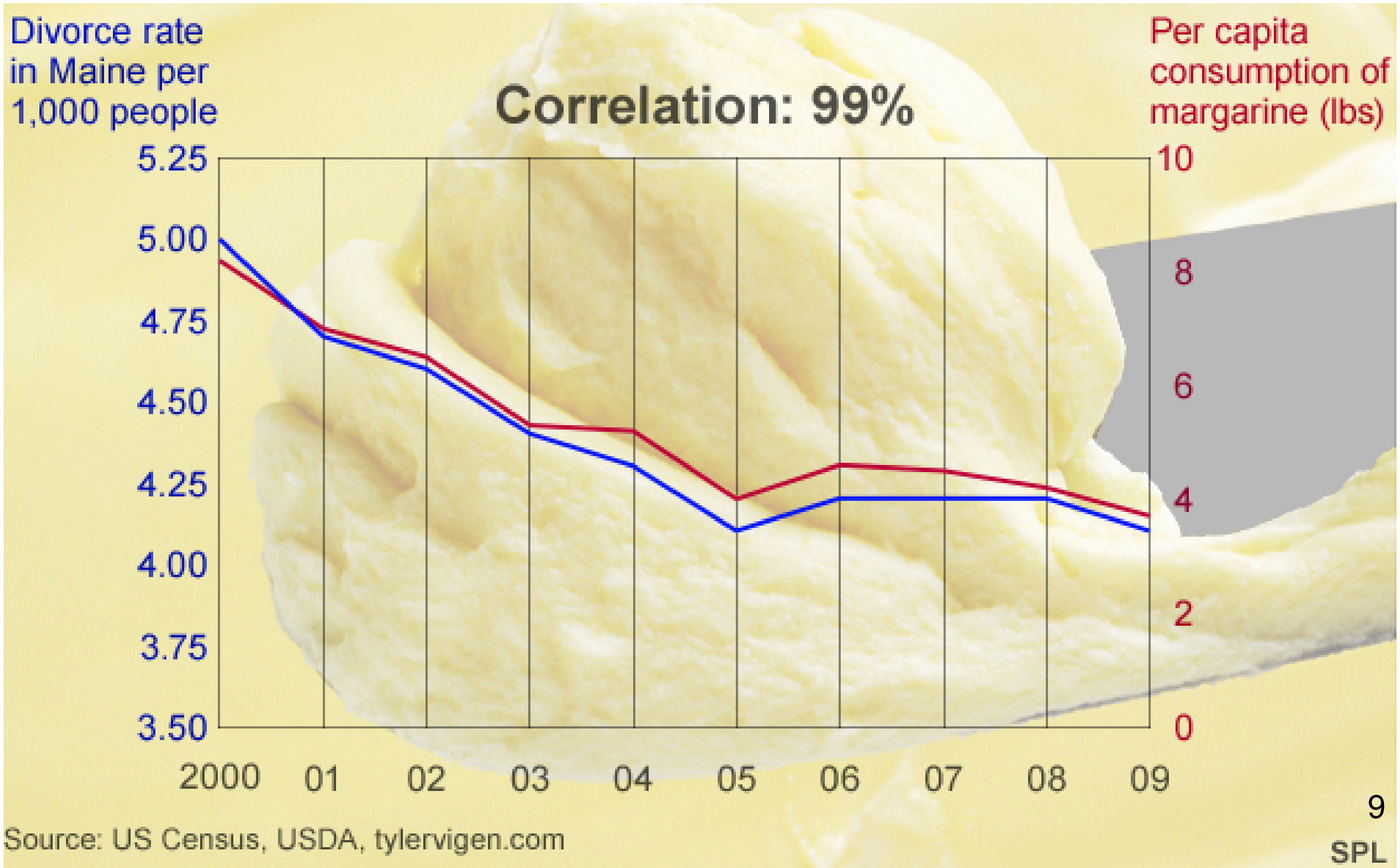
- Spearman's rho

  $$\rho_R = 3[\Pr((Y_1 - Y_2)(Z_1 - Z_3) > 0) - \Pr((Y_1 - Y_2)(Z_1 - Z_3) < 0)];$$

- Pearson product-moment correlation coefficient

  $$\rho_L = \frac{cov(Y, Z)}{sd(Y)sd(Z)}.$$

Responsible use of Pearson correlation: Spurious Correlation between Margarine consumption and Divorce Rate in Main (US)



Main results

If $E(N) < \infty$ and $F \in \mathcal{L}$ is continuous, then

$$\tau^{+1}(t) \sim \rho_R^{+1}(t) \sim 1, \quad \text{where}$$

$$\begin{aligned}\tau^{+1}(t) = \Pr&\left(\left(S_{T,1} - S_{T,2}\right)\left(X_{T,1}^{(1)} - X_{T,2}^{(1)}\right) > 0 | X_{T,1}^{(1)}, X_{T,2}^{(1)} > t\right) \\ &- \Pr\left(\left(S_{T,1} - S_{T,2}\right)\left(X_{T,1}^{(1)} - X_{T,2}^{(1)}\right) < 0 | X_{T,1}^{(1)}, X_{T,2}^{(1)} > t\right)\end{aligned}$$

$$\begin{aligned}\rho_R^{+1}(t) = 3\Pr&\left(\left(S_{T,1} - S_{T,2}\right)\left(X_{T,1}^{(1)} - X_{T,3}^{(1)}\right) > 0 | X_{T,1}^{(1)}, X_{T,2}^{(1)}, X_{T,3}^{(1)} > t\right) \\ &- 3\Pr\left(\left(S_{T,1} - S_{T,2}\right)\left(X_{T,1}^{(1)} - X_{T,3}^{(1)}\right) < 0 | X_{T,1}^{(1)}, X_{T,2}^{(1)}, X_{T,3}^{(1)} > t\right)\end{aligned}$$

Main results (cont'd)

If $E(N) < \infty$, $p_N(2) > 0$ and F is continuous, then

$$\tau^{12}(t) \sim 1/3 \text{ and } \rho_R^{12}(t) \sim 7/15, \text{ where}$$

$$\tau^{12}(t) = 3\Pr\left(\left(X_{T,1}^{(1)} - X_{T,2}^{(1)}\right)\left(X_{T,1}^{(2)} - X_{T,2}^{(2)}\right) > 0 | X_{T,1}^{(2)}, X_{T,2}^{(2)} > t\right)$$
$$- 3\Pr\left(\left(X_{T,1}^{(1)} - X_{T,2}^{(1)}\right)\left(X_{T,1}^{(2)} - X_{T,2}^{(2)}\right) < 0 | X_{T,1}^{(2)}, X_{T,2}^{(2)} > t\right)$$

$$\rho_R^{12}(t) = 3\Pr\left(\left(X_{T,1}^{(1)} - X_{T,2}^{(1)}\right)\left(X_{T,1}^{(2)} - X_{T,3}^{(2)}\right) > 0 | X_{T,1}^{(2)}, X_{T,2}^{(2)}, X_{T,3}^{(2)} > t\right)$$
$$- 3\Pr\left(\left(X_{T,1}^{(1)} - X_{T,2}^{(1)}\right)\left(X_{T,1}^{(2)} - X_{T,3}^{(2)}\right) < 0 | X_{T,1}^{(2)}, X_{T,2}^{(2)}, X_{T,3}^{(2)} > t\right).$$

Main results (cont'd)

Finally, two conditional versions of Pearson product-moment correlation coefficient are

$$\rho_L^{+1}(t) = \frac{cov\left(S_T, X_T^{(1)} | X_T^{(1)} > t\right)}{\sqrt{Var\left(S_T | X_T^{(1)} > t\right) Var\left(X_T^{(1)} | X_T^{(1)} > t\right)}}$$

and

$$\rho_L^{12}(t) = \frac{cov\left(X_T^{(1)}, X_T^{(2)} | X_T^{(2)} > t\right)}{\sqrt{Var\left(X_T^{(1)} | X_T^{(2)} > t\right) Var\left(X_T^{(2)} | X_T^{(2)} > t\right)}}.$$

- If $E(1+\epsilon)^N < \infty$ with $\epsilon > 0$ and $F \in MDA(\Phi_\alpha)$ or $\mathcal{R}(\alpha)$, then

$$\rho_L^{+1}(t) \sim 1, \quad \text{provided that } \alpha > 2;$$

- If $EN^2 < \infty$, $p_N(2) > 0$ and $F \in MDA(\Phi_\alpha)$ or $\mathcal{R}(\alpha)$, then

$$\lim_{t\to\infty} \rho_L^{12}(t) = \sqrt{\frac{\alpha(\alpha-2)}{(5\alpha-1)(\alpha-1)}}, \quad \text{provided that } \alpha > 2;$$

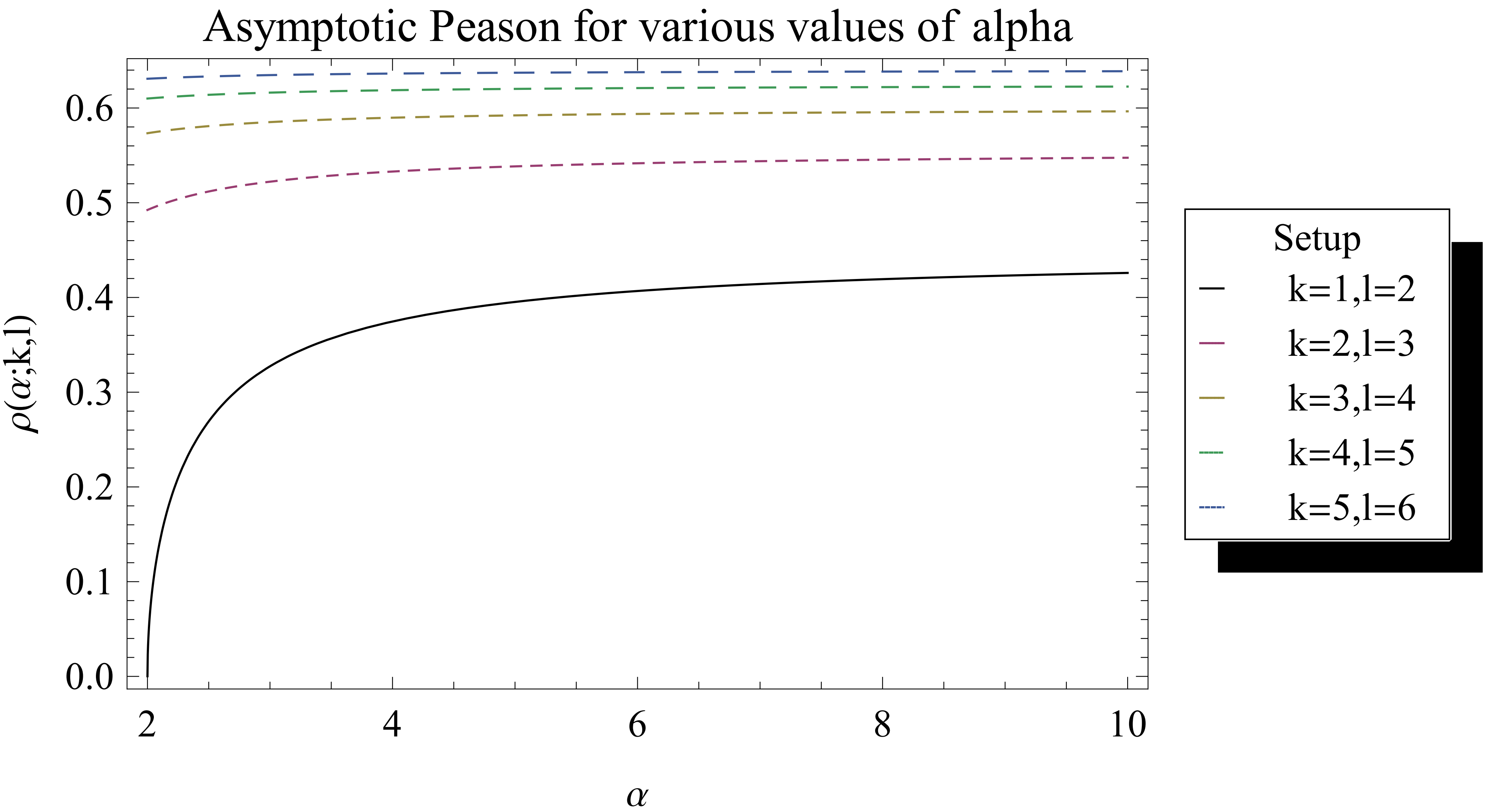
- If $E(1+\epsilon)^N < \infty$ with $\epsilon > 0$ and $F \in MDA(\Lambda) \bigcap \mathcal{S}$, then

$$\rho_L^{+1}(t) \sim 1;$$

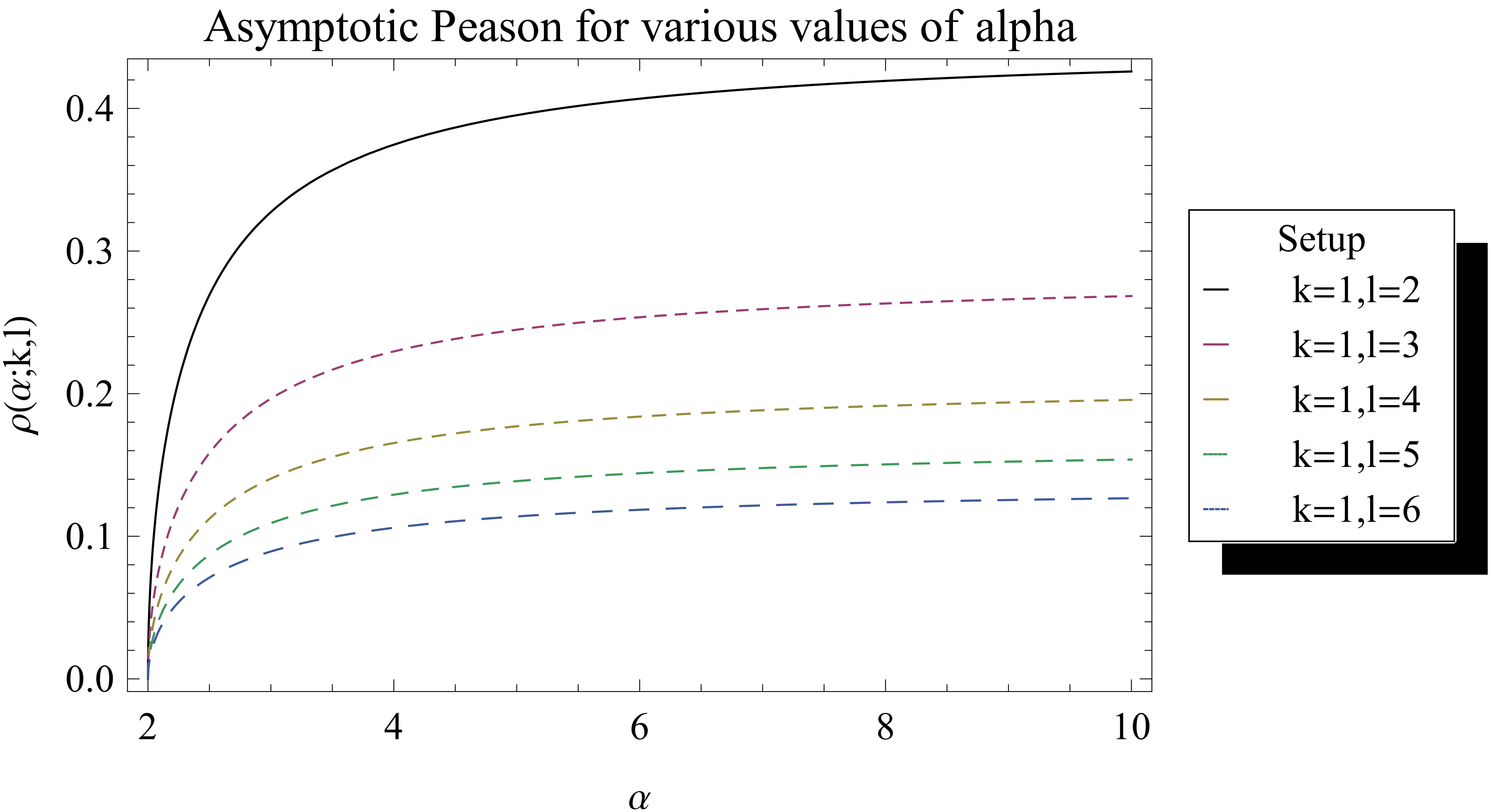
- If $EN^2 < \infty$, $p_N(2) > 0$ and $F \in MDA(\Lambda)$, then

$$\lim_{t\to\infty} \rho_L^{12}(t) = \sqrt{\frac{1}{5}}.$$

Main results (cont'd)



Main results (cont'd)



Background: Hill estimator

Let $X_1, X_2, \ldots, X_n$ be iid rv's with $X_1 \in \mathcal{R}(\alpha)$. Then, the *Hill estimator* is

$$\frac{1}{k(n)} \sum_{i=1}^{k(n)} \left(\log X_n^{(i)} - \log X_n^{(k(n)+1)} \right) \to \alpha^{-1}, \text{ as } n \to \infty$$

where $X_n^{(1)} \geq X_n^{(2)} \geq \ldots \geq X_n^{(n)}$, provided that $k(n) \to \infty$ and $k(n)/n \to 0$ as $n \to \infty$.

Future work

Let $X_1, X_2, \ldots, X_{2n}$ be iid rv's with $X_1 \in \mathcal{R}(\alpha)$ with $\alpha > 2$. Denote

$$Y_i = \max\left(X_{2i-1}, X_{2i}\right), Z_i = \min\left(X_{2i-1}, X_{2i}\right) \text{ for all } i = 1, 2, \ldots, n.$$

Then, the most natural estimator is

$$\frac{A(k(n),n)}{\sqrt{B(k(n),n)C(k(n),n)}} \to \sqrt{\frac{\alpha(\alpha-2)}{(5\alpha-1)(\alpha-1)}}, \text{ as } n \to \infty,$$

provided that $k(n) \to \infty$ and $k(n)/n \to 0$ as $n \to \infty$. Note that

$$A(k(n),n) = \sum_{i=1}^{n} Y_i Z_i I\left(Z_i > Z_n^{(k(n)+1)}\right) - \frac{1}{k(n)} \sum_{i=1}^{n} Y_i I\left(Z_i > Z_n^{(k(n)+1)}\right) \sum_{i=1}^{n} Z_i I\left(Z_i > Z_n^{(k(n)+1)}\right),$$

Future work (cond'd)

$$B(k(n),n) = \sum_{i=1}^{n} Y_i^2 I\left(Z_i > Z_n^{(k(n)+1)}\right) - \frac{1}{k(n)} \left(\sum_{i=1}^{n} Y_i I\left(Z_i > Z_n^{(k(n)+1)}\right)\right)^2$$

and

$$C(k(n),n) = \sum_{i=1}^{n} Z_i^2 I\left(Z_i > Z_n^{(k(n))+1}\right) - \frac{1}{k(n)} \left(\sum_{i=1}^{n} Z_i I\left(Z_i > Z_n^{(k(n)+1)}\right)\right)^2,$$

where $Z_n^{(1)} \geq Z_n^{(2)} \geq \ldots \geq Z_n^{(n)}$.

Conclusions

1. Qualitative and Quantitative piece of information

2. Potential alternative estimator(s) for the RV index

3. More work on the estimation side