Extreme versions of Wang risk measures and their estimation

Gilles STUPFLER (Aix Marseille Université)
Joint work with Jonathan EL METHNI (University Paris Descartes)

CIRM Workshop "Extremes – Copulas – Actuarial science"
Luminy, France, 23rd February 2016

Outline

- Wang risk measures and their extreme analogues
- Estimation
  - Intermediate case
  - Extreme case
- Finite-sample study
- Discussion

Setup

Let X be a positive rv having cdf F. The most famous risk measure related to X is arguably the Value-at-Risk (or quantile):

$$\forall \beta \in (0,1),\ \mathrm{VaR}(\beta) = q(\beta) = \inf\{t \in \mathbb{R} \,|\, F(t) \geq \beta\}.$$

An extreme VaR is obtained by letting $\beta \uparrow 1$.

- The estimation of a single extreme quantile of X only gives incomplete information on the extremes of X.
- The VaR lacks the coherency property (translation invariance, positive homogeneity, monotonicity and subadditivity) which can be an issue from the financial point of view.

This is why a family of other quantities, possibly taking into account the whole right tail of X, was developed and studied.

Wang risk measures

Definition (Wang, 1996)

The *Wang distortion risk measure* (DRM) of X, with distortion function g, is defined as

$$R_g(X) := \int_0^\infty g(1 - F(x))dx$$

for any nondecreasing, right-continuous function $g : [0,1] \to [0,1]$ with $g(0) = 0$ and $g(1) = 1$.

Under mild regularity conditions on q (*e.g.* continuity) then

$$R_g(X) = \int_0^1 q(1-\alpha)dg(\alpha).$$

A Wang DRM is then a weighted version of the expectation of X.
We recover the expectation of X by taking $g(x) = x$.

Examples of Wang DRMs:

- For $g(x) = \mathbb{I}\{\min\{1, x/(1-\beta)\} = 1\}$: the $\mathrm{VaR}(\beta)$;
- For $g(x) = \min\{1, x/(1-\beta)\}$: the $\mathrm{TVaR}(\beta)$, *i.e.* the average of all quantiles exceeding $\mathrm{VaR}(\beta)$, is a coherent DRM.

If moreover F is continuous:

- For $g(x) = \min\{1, x/(1-\beta)\}$ applied to X^a: the ath Conditional Tail Moment (CTM) of X (El Methni *et al.*, 2014), *i.e.* the average value of X^a in the $100(1-\beta)\%$ highest cases.

Any combination of CTMs and VaR, such as the conditional variance or the stop-loss premium above level $q(\beta)$ may also be computed by combining Wang DRMs.

The Wang DRM R_g is coherent iff g is a concave function (Wirch and Hardy, 2002).

Extreme versions of Wang DRMs

How can Wang DRMs be used to understand the extremes of X?

First idea (Vandewalle and Beirlant, 2004): consider the Wang DRMs of $\max(X - R, 0)$ for $R \uparrow \infty$:

$$\int_0^{\infty} g(1 - F(x + R))dx = \int_R^{\infty} g(1 - F(y))dy.$$

- This is adapted to the examination of excess-of-loss reinsurance policies for extreme losses, *e.g.* the stop-loss premium above level R is obtained for $g(x) = x$;
- Their work is restricted to concave functions g satisfying a regular variation condition in a neighborhood of 0: it excludes the simple VaR risk measure, for instance.

Our idea is rather to consider a:

- conditional construction, *i.e.* which looks at the extremes of X given that X lies above a high quantile;
- unifying framework to recover as many risk measures as possible.

Let g be a distortion function. Pick $\beta \in [0,1)$ and let g_β be the distortion function

$$g_\beta(x) := g\left(\min\left[1, \frac{x}{1-\beta}\right]\right).$$

The related Wang DRM is denoted by $R_{g,\beta}(X)$:

$$R_{g,\beta}(X) := \int_0^\infty g_\beta(1-F(x))dx = \int_0^\infty g(1-F_\beta(x))dx$$

$$\text{with } F_\beta(x) := \max\left[0, \frac{F(x)-\beta}{1-\beta}\right].$$

Because F_β is (usually) the cdf of X given $X > q(\beta)$, it follows that

$R_{g,\beta}(X)$ is the Wang DRM of X given $X > q(\beta)$.

For $\beta \uparrow 1$, it will be thought of as an extreme Wang DRM.

Our previous examples may be recovered in this framework:

- For $g(x) = \mathbb{I}\{x = 1\}$: then $R_{g,\beta}(X)$ is $\mathrm{VaR}(\beta)$;
- For $g(x) = x$: then $R_{g,\beta}(X)$ is $\mathrm{TVaR}(\beta)$;
- When F is continuous, for $g(x) = x$ applied to X^a: then $R_{g,\beta}(X^a)$ is the ath CTM of X above level $q(\beta)$.

How can extreme Wang DRMs be estimated?

Intermediate case

In all what follows, we work in the case of a heavy-tailed X:

Second order condition $\mathcal{C}_2(\gamma, \rho, A)$

There are $\gamma > 0$ and $\rho \leq 0$ such that

$$\forall x > 0, \ \lim_{t\to\infty} \frac{1}{A(t)} \left(\frac{U(tx)}{U(t)} - x^\gamma \right) = x^\gamma \frac{x^\rho - 1}{\rho}$$

where $U(t) = q(1 - t^{-1})$ denotes the tail quantile function and A is a Borel measurable function having constant sign and converging to 0. If $\rho = 0$, the rhs is $x^\gamma \log x$.

If actually $U(t) = t^\gamma$, namely X is Pareto distributed with tail index γ, it can be shown that

$$R_{g,\beta}(X^a) = (1-\beta)^{-a\gamma} \int_0^1 s^{-a\gamma} dg(s) = [q(\beta)]^a \int_0^1 s^{-a\gamma} dg(s).$$

In our framework, this conclusion holds asymptotically for $\beta \uparrow 1$:

$$R_{g,\beta}(X^a) = [q(\beta)]^a \int_0^1 s^{-a\gamma} dg(s)(1 + o(1)).$$

We now:

- let $\beta = \beta_n \uparrow 1$ so that $R_{g,\beta_n}(X^a)$ is an extreme Wang DRM;
- replace $q(\beta_n)$ by its empirical counterpart $X_{n-\lfloor n(1-\beta_n)\rfloor,n}$, where $X_{1,n} \leq \cdots \leq X_{n,n}$ are the order statistics of an $n-$sample of iid copies of X and $\lfloor \cdot \rfloor$ is the floor function;
- plug in a consistent estimator $\widehat{\gamma}_n$ of γ.

Our first estimator is then

$$\widehat{R}^{AE}_{g,\beta_n}(X^a) = X^a_{n-\lfloor n(1-\beta_n)\rfloor,n} \int_0^1 s^{-a\widehat{\gamma}_n} dg(s).$$

Alternatively, one can show that for any $a > 0$:

$$R_{g,\beta_n}(X^a) = \int_0^1 [q(1-(1-\beta_n)s)]^a dg(s).$$

We may directly replace the function q by its empirical counterpart:

$$\forall \alpha \in (0,1], \ \widehat{q}_n(\alpha) = X_{n-\lfloor n(1-\alpha) \rfloor, n}$$

which yields our second, functional plug-in estimator:

$$\widehat{R}^{PL}_{g,\beta_n}(X^a) = \int_0^1 X^a_{n-\lfloor n(1-\beta_n)s \rfloor, n} dg(s).$$

Here, the empirical quantiles $X_{n-\lfloor n(1-\beta_n)s \rfloor, n}$ are tail order statistics for the control of which the second order framework is needed.

Theorem 1 (Intermediate case, AE estimator)

Assume that U satisfies $\mathcal{C}_2(\gamma, \rho, A)$. Assume further that

$$\beta_n \to 1, \; n(1-\beta_n) \to \infty \;\; \textit{and} \;\; \sqrt{n(1-\beta_n)}A((1-\beta_n)^{-1}) \to \lambda \in \mathbb{R}$$

and the following joint convergence holds:

$$\sqrt{n(1-\beta_n)}\left(\widehat{\gamma}_n - \gamma, \frac{X_{n-\lfloor n(1-\beta_n)\rfloor,n}}{q(\beta_n)} - 1\right) \xrightarrow{d} (\Gamma, \Theta).$$

Finally, let $g_1, \ldots, g_d$ be distortion functions and $a_1, \ldots, a_d > 0$ and suppose that

$$\exists \eta > 0, \; \forall j \in \{1, \ldots, d\}, \; \int_0^1 s^{-a_j\gamma - 1/2 - \eta} dg_j(s) < \infty.$$

Theorem 1 (Intermediate case, AE estimator, cont'd)

Then the random vector $\sqrt{n(1-\beta_n)}\left(\dfrac{\widehat{R}^{AE}_{g_j,\beta_n}(X^{a_j})}{R_{g_j,\beta_n}(X^{a_j})}-1\right)_{1\leq j\leq d}$

asymptotically has the joint distribution of

$$\left(a_j\left[\Theta+\frac{\displaystyle\int_0^1 s^{-a_j\gamma}\left(\log(1/s)\Gamma-\lambda\frac{s^{-\rho}-1}{\rho}\right)dg_j(s)}{\displaystyle\int_0^1 s^{-a_j\gamma}dg_j(s)}\right]\right)_{1\leq j\leq d}.$$

Theorem 2 (Intermediate case, PL estimator)

Assume that U *satisfies* $\mathcal{C}_2(\gamma, \rho, A)$. *Assume further that*

$$\beta_n \to 1,\ n(1-\beta_n) \to \infty \ \text{ and } \ \sqrt{n(1-\beta_n)}A((1-\beta_n)^{-1}) \to \lambda \in \mathbb{R}.$$

Let $g_1, \ldots, g_d$ *be distortion functions and* $a_1, \ldots, a_d > 0$. *If*

$$\exists \eta > 0,\ \forall j \in \{1, \ldots, d\},\ \int_0^1 s^{-a_j\gamma - 1/2 - \eta} dg_j(s) < \infty,$$

then the random vector $\sqrt{n(1-\beta_n)} \left(\dfrac{\widehat{R}^{PL}_{g_j,\beta_n}(X^{a_j})}{R_{g_j,\beta_n}(X^{a_j})} - 1 \right)_{1 \le j \le d}$ *is asymptotically multivariate Gaussian centered with covariance matrix*

$$V_{i,j} = a_i a_j \gamma^2 \frac{\int_{[0,1]^2} \min(s,t) s^{-a_i\gamma - 1} t^{-a_j\gamma - 1} dg_i(s) dg_j(t)}{\int_0^1 s^{-a_i\gamma} dg_i(s) \int_0^1 t^{-a_j\gamma} dg_j(t)}.$$

Comments: intermediate case

The conditions on (β_n) say that the quantiles used for the computation are those whose order is intermediate *i.e.* "large but not extreme" and the asymptotic bias of the estimator is controlled.

About the integrability condition $\int_0^1 s^{-a\gamma-1/2-\eta} dg(s) < \infty$:

- it entails $\int_0^1 s^{-a\gamma-\eta} dg(s) < \infty$ so that the Wang DRM exists and is finite;
- the $s^{-1/2}$ term makes sure one can use an approximation of the tail quantile process by a standard Brownian motion W:

$$\frac{X_{n-\lfloor n(1-\beta_n)s\rfloor,n}}{U((1-\beta_n)^{-1})} - s^{-\gamma} \approx \frac{s^{-\gamma}}{\sqrt{n(1-\beta_n)}} \left(\gamma \frac{W(s)}{s} - \lambda \frac{s^{-\rho}-1}{\rho} \right).$$

We now focus on how to eliminate condition $n(1-\beta_n) \to \infty$.

Extreme case

By condition $\mathcal{C}_2(\gamma, \rho, A)$:

$$\frac{q(1-1/tx)}{q(1-1/t)} = \frac{U(tx)}{U(t)} \approx x^\gamma \text{ as } t \to \infty.$$

Pick then β_n, $\delta_n \uparrow 1$ and for n large enough, rewrite this as

$$\forall s \in (0,1),\ q(1-(1-\delta_n)s) \approx \left(\frac{1-\beta_n}{1-\delta_n}\right)^\gamma q(1-(1-\beta_n)s).$$

- This links extreme quantiles to intermediate quantiles by the means of γ.
- Plugging in the rhs a consistent estimator of γ and the empirical estimator of q results in the extrapolated Weissman estimator.
- This provides a method to estimate extreme Wang DRMs since it is actually enough to estimate extreme quantiles.

Let then:

- $\beta_n \uparrow 1$ such that $n(1-\beta_n) \to \infty$, *i.e.* an intermediate level;
- $\delta_n \uparrow 1$ such that $(1-\delta_n)/(1-\beta_n) \to 0$, *i.e.* an extreme level;
- $\widehat{\gamma}_n$ be a consistent estimator of γ, *e.g.* the Hill (1975) estimator;
- $\widehat{R}_{g,\beta_n}(X^a)$ be any $\sqrt{n(1-\beta_n)}$−relatively consistent estimator of $R_{g,\beta_n}(X^a)$, such as the intermediate AE or PL estimator,

and estimate the extreme Wang DRM $R_{g,\delta_n}(X^a)$ by

$$\widehat{R}^W_{g,\delta_n}(X^a;\beta_n) = \left(\frac{1-\beta_n}{1-\delta_n}\right)^{a\widehat{\gamma}_n} \widehat{R}_{g,\beta_n}(X^a)$$

⇒ This is a Weissman-type estimator again!

Theorem 3 (Extreme case)

Work under the conditions of Theorem 2. Assume that $\rho < 0$ *and the following convergences hold:*

$$\sqrt{n(1-\beta_n)}\left(\frac{\widehat{R}_{g_j,\beta_n}(X^{a_j})}{R_{g_j,\beta_n}(X^{a_j})}-1\right)_{1\le j\le d} = O_{\mathbb{P}}(1)$$

$$\text{and} \quad \sqrt{n(1-\beta_n)}(\widehat{\gamma}_n-\gamma) \xrightarrow{d} \xi.$$

Then:

$$\frac{\sqrt{n(1-\beta_n)}}{\log([1-\beta_n]/[1-\delta_n])}\left(\frac{\widehat{R}^W_{g_j,\delta_n}(X^{a_j};\beta_n)}{R_{g_j,\delta_n}(X^{a_j})}-1\right)_{1\le j\le d} \xrightarrow{d} \begin{pmatrix} a_1\xi \\ \vdots \\ a_d\xi \end{pmatrix}.$$

We recover the behavior of the Weissman estimator for extreme quantiles by letting $d=1$ and $g(x)=\mathbb{1}\{x=1\}$.

Finite-sample study

Consider the two following distributions for X:

- Fréchet distribution: $F(x) = e^{-x^{-1/\gamma}}$, $x > 0$;
- Burr distribution: $F(x) = 1 - (1 + x^{-\rho/\gamma})^{1/\rho}$, $x > 0$.

These distributions are heavy-tailed and:

- both have extreme-value index γ;
- the Fréchet distribution has $\rho = -1$;
- the Burr distribution has second-order parameter ρ.

We take $\gamma \in \{1/6, 1/4\}$ and $\rho \in \{-2, -1\}$.

In each case we generate 5000 independent samples of X having size $n \in \{100, 300\}$.

We consider the following distortion functions:

- the expectation function $g(x) = x$ weighs all quantiles equally. It generates the extreme Conditional Tail Expectation (CTE);
- the Dual Power (DP) function $g(x) = 1 - (1 - x)^m$, with $m \in \mathbb{N} \setminus \{0\}$, gives higher but bounded weight to large quantiles. Here it yields the expectation of $\max(X_1, \ldots, X_m)$ when the iid copies $X_1, \ldots, X_m$ of X all lie above an extreme quantile;
- the Proportional Hazard (PH) transform function $g(x) = x^\alpha$, with $\alpha \in (0, 1)$, gives higher and unbounded weight to large quantiles.

Intuitively, the most difficult case is the PH case: the most extreme quantiles are those whose weight is the highest in the Wang DRM and whose estimation is the hardest.

Selecting β_n

Our estimators of extreme Wang DRMs are based upon the choice of the intermediate level β_n, used for:

- the estimation of γ;
- the preliminary estimation of the intermediate risk measure $R_{g,\beta_n}(X)$, which is itself used in the extrapolation procedure.

This parameter is chosen using the Hill estimator of γ,

$$\widehat{\gamma}_{\beta_n} = \frac{1}{\lceil n(1-\beta_n)\rceil} \sum_{i=1}^{\lceil n(1-\beta_n)\rceil} \log\left(X_{n-i+1,n}\right) - \log\left(X_{n-\lceil n(1-\beta_n)\rceil,n}\right)$$

which shall also be the estimator of γ in our procedure.

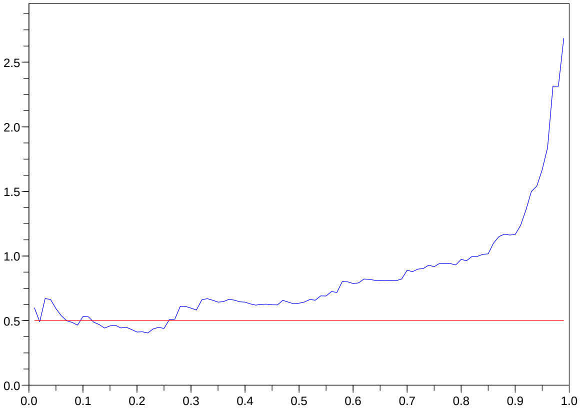

Figure 1: Sample path of the Hill estimator, computed with $n = 100$ Burr rvs for which $\gamma = 1/2$ and $\rho = -1$. x−axis: $1 - \beta$.

Our procedure is the following:

- pick $\beta_0 > 0$ and a window-width $h > 0$ (we take $h = 0.1$ here);
- for $\beta_0 < \beta < 1 - h$, define $I(\beta, h) = [\beta, \beta + h]$ and compute the standard deviation $\sigma(\beta, h)$ of the block $\{\widehat{\gamma}_\beta, \beta \in I\}$;
- find the last β such that $\sigma(\beta, h)$ realizes a local minimum less than the mean of the $\{\sigma(\beta, h), \beta \in (\beta_0, 1 - h)\}$;
- choose a value β^* in the interval $I(\beta, h) = [\beta, \beta + h]$, *e.g.*
  - ◊ its center,
  - ◊ the one giving the median estimate in the block.

Think of this as selecting β in the middle of the first interval in the extremes of the sample where the estimation is stable.

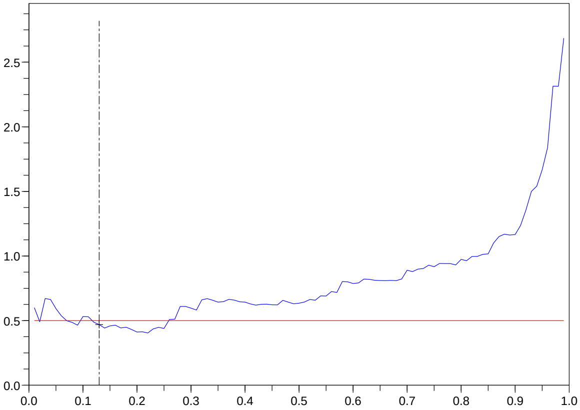

Figure 2: Sample path of the Hill estimator; taking $h = 0.1$ and recording the median estimate in the stability region yields $\beta^* = 0.86$.

Results: Conditional Tail Expectation

γ	δ	Est.	Fréchet $n=100$	Fréchet $n=300$	Burr $\rho=-1$ $n=100$	Burr $\rho=-1$ $n=300$	Burr $\rho=-2$ $n=100$	Burr $\rho=-2$ $n=300$
$\frac{1}{6}$	0.99	AE	0.0325	0.0098	0.0374	0.0133	0.0291	0.0095
		PL	0.0317	0.0097	0.0357	0.0127	0.0286	0.0094
	0.995	AE	0.0457	0.0137	0.0540	0.0191	0.0401	0.0130
		PL	0.0446	0.0135	0.0518	0.0184	0.0395	0.0129
	0.999	AE	0.0891	0.0258	0.1115	0.0386	0.0752	0.0236
		PL	0.0871	0.0255	0.1073	0.0375	0.0741	0.0235
$\frac{1}{4}$	0.99	AE	0.0973	0.0285	0.1028	0.0349	0.0834	0.0248
		PL	0.0900	0.0278	0.0944	0.0332	0.0835	0.0246
	0.995	AE	0.1411	0.0402	0.1515	0.0509	0.1190	0.0341
		PL	0.1305	0.0392	0.1395	0.0484	0.1202	0.0337
	0.999	AE	0.3039	0.0787	0.3350	0.1063	0.2492	0.0631
		PL	0.2807	0.0768	0.3102	0.1017	0.2604	0.0622

Table 1: Case of the CTE: relative MSEs.

Results: Dual Power, $m = 3$

γ	δ	Est.	Fréchet $n = 100$	Fréchet $n = 300$	Burr $\rho = -1$ $n = 100$	Burr $\rho = -1$ $n = 300$	Burr $\rho = -2$ $n = 100$	Burr $\rho = -2$ $n = 300$
$\frac{1}{6}$	0.99	AE	0.0487	0.0169	0.0629	0.0215	0.0458	0.0140
		PL	0.0448	0.0160	0.0549	0.0194	0.0443	0.0142
	0.995	AE	0.0653	0.0225	0.0866	0.0295	0.0609	0.0182
		PL	0.0597	0.0212	0.0757	0.0267	0.0586	0.0184
	0.999	AE	0.1177	0.0394	0.1658	0.0549	0.1084	0.0307
		PL	0.1073	0.0371	0.1456	0.0499	0.1033	0.0306
$\frac{1}{4}$	0.99	AE	0.1558	0.0449	0.2175	0.0570	0.1327	0.0376
		PL	0.1397	0.0439	0.1707	0.0501	0.1252	0.0388
	0.995	AE	0.2182	0.0602	0.3161	0.0787	0.1818	0.0494
		PL	0.1932	0.0582	0.2471	0.0690	0.1698	0.0503
	0.999	AE	0.4485	0.1086	0.7089	0.1508	0.3561	0.0854
		PL	0.3899	0.1038	0.5482	0.1323	0.3279	0.0852

Table 2: Case of the DP(3): relative MSEs.

Results: PH transform, $\alpha = 2/3$

γ	δ	Est.	Fréchet $n=100$	Fréchet $n=300$	Burr $\rho=-1$ $n=100$	Burr $\rho=-1$ $n=300$	Burr $\rho=-2$ $n=100$	Burr $\rho=-2$ $n=300$
$\frac{1}{6}$	0.99	AE	0.0517	0.0162	0.0618	0.0207	0.0487	0.0141
		PL	0.0395	0.0145	0.0421	0.0157	0.0382	0.0133
	0.995	AE	0.0699	0.0216	0.0848	0.0282	0.0654	0.0184
		PL	0.0534	0.0191	0.0584	0.0215	0.0511	0.0172
	0.999	AE	0.1290	0.0383	0.1612	0.0523	0.1196	0.0311
		PL	0.0993	0.0334	0.1143	0.0406	0.0932	0.0286
$\frac{1}{4}$	0.99	AE	0.1920	0.0461	0.2432	0.0678	0.1516	0.0405
		PL	0.1008	0.0347	0.1122	0.0438	0.0927	0.0355
	0.995	AE	0.2669	0.0613	0.3421	0.0921	0.2055	0.0529
		PL	0.1384	0.0453	0.1595	0.0594	0.1242	0.0452
	0.999	AE	0.5454	0.1088	0.7137	0.1727	0.3928	0.0906
		PL	0.2760	0.0796	0.3409	0.1136	0.2330	0.0748

Table 3: Case of the PH(2/3): relative MSEs.

Conclusion

- We built extreme versions of Wang DRMs;
- They can be said to constitute a unifying framework for the study of risk above a high level;
- We built estimators of these risk measures:
  - ◊ in the intermediate case with the empirical counterpart of the quantile function;
  - ◊ in the extreme case by an extrapolation method;
- Our estimators have decent finite-sample performance for moderate γ;
- The PL estimator seems preferable overall especially when n is small and/or the integrability constraint on γ is strong.

Forthcoming studies

- Study analogue estimators for distributions being:
  - ◊ light-tailed: Gaussian distribution, Weibull distribution...
  - ◊ short-tailed: Uniform distribution, Beta distribution...
- Obtain the consistency of the estimators with weaker integrability conditions;
- Study a PL estimator which puts less weight on the most extreme quantiles, *e.g.* trimmed/Winsorized PL estimators;
- Study the behavior of the estimators with correlated data as input.

References

El Methni, J., Gardes, L., Girard, S. (2014). Nonparametric estimation of extreme risks from conditional heavy-tailed distributions, *Scandinavian Journal of Statistics* **41**(4): 988–1012.

El Methni, J., Stupfler, G. (2015). Extreme versions of Wang risk measures and their estimation for heavy-tailed distributions, preprint.

Vandewalle, B., Beirlant, J. (2006). On univariate extreme value statistics and the estimation of reinsurance premiums, *Insurance: Mathematics and Economics* **38**: 441–459.

Wang, S.S. (1996). Premium calculation by transforming the layer premium density, *ASTIN Bulletin* **26**: 71–92.

Thanks for listening!