On Univariate Extreme Value Statistics and the Estimation of Reinsurance Premiums

B. Vandewalle
J. Beirlant

Katholieke Universiteit Leuven
University Center of Statistics
Actuarial Science Team

Göteborg 2005
... on univariate extreme value statistics and the estimation of insurance premiums for excess-of-loss reinsurance policies in excess of a high retention level ... with special attention to heavy-tailed distributions and Wang’s premium principle as a generalization to the net premium principle.
Overview

- (Re)insurance premium calculation
  - Net premium principle
  - Wang’s premium principle
    - applied to excess-of-loss reinsurance setting
- Extreme value statistics
  - Motivation
  - Extreme value theory (first order framework)
  - Estimating reinsurance premiums
- Finite sample behavior
  - Simulated data (Fréchet, Burr)
  - Reinsurance premiums (net premium, dual-power transform)
Overview

- Asymptotic results
  - Motivation
  - Extreme value theory (second order framework)
  - Premium approximation error
  - Asymptotic normality and bias

- Finite sample behavior
  - Secura Belgian Re data
  - Simulated data (Fréchet, Burr)
  - Reinsurance premiums (net premium, dual-power transform)
Insurance premium calculation

Net premium principle

Insurance premium calculation ... at the heart of actuarial science ...

\( X \sim \) non-negative random variable denoting the total claim amount resulting from a single insurance policy

(Decumulative distribution function \( \bar{F}(x) = P(X > x) \))

Net premium principle

Under the assumption that risk is essentially non-existing if the insurer sells enough identically distributed and independent policies, on average the insurer will not lose any money when using premium

\[
\Pi = E(X) = \int_0^\infty \bar{F}(x) dx
\]
Insurance premium calculation

Wang’s premium principle

However ... experienced losses hardly ever equal expected losses ...

\[ \sim \] extra loading for risk is desirable if the insurer on average does not want to lose any money \((\Pi \geq E(X))\)

The search for sound premium calculation principles has been the subject of numerous actuarial papers and remains debatable with respect to choice ...

Wang’s premium principle (Wang, 1996)

With \(g\) an increasing, concave function, called the distortion, that maps \([0, 1]\) onto \([0, 1]\)

\[
\Pi = \int_{0}^{\infty} g\left(\bar{F}(x)\right) \, dx
\]
Insurance premium calculation

Wang’s premium principle

Examples:

- \( g(x) = x \) \( \rightsquigarrow \) Net premium principle (\( II = EX \))
- \( g(x) = x^{1/\alpha} \) (\( \alpha \geq 1 \)) \( \rightsquigarrow \) Proportional hazard transform principle
- \( g(x) = 1 - (1 - x)^\alpha \) (\( \alpha \geq 1 \)) \( \rightsquigarrow \) Dual-power transform principle
- \( g(x) = (1 + \alpha)x - \alpha x^2 \) (\( 0 \leq \alpha \leq 1 \)) \( \rightsquigarrow \) Gini principle
- \( g(x) = \frac{\sqrt{1+\alpha x} - 1}{\sqrt{1+\alpha} - 1} \) (\( \alpha > 0 \)) \( \rightsquigarrow \) Square root function principle
- \( g(x) = \frac{1-e^{-\alpha x}}{1-e^{-\alpha}} \) (\( \alpha > 0 \)) \( \rightsquigarrow \) Exponential function principle
- \( g(x) = \frac{\log(1+\alpha x)}{\log(1+\alpha)} \) (\( \alpha > 0 \)) \( \rightsquigarrow \) Logarithmic function principle
Insurance premium calculation

Wang’s premium principle
Reinsurance premium calculation

Excess-of-loss reinsurance setting

\[ X_R = (X - R)_+ = \max (0, X - R) \]

\[ \sim \] total claim amount resulting from a single excess-of-loss reinsurance policy in excess of a high retention level \( R \)

\( \left( \text{decumulative distribution function } F_R(x) = F(x + R) \right) \)

\[ \downarrow \]

\[ \Pi(R) = \int_R^\infty g \left( F(x) \right) \, dx \]

\[ \sim \] premium given as a function of the decumulative distribution function \( F \) of the original claim amount for the layer from \( R \) to infinity
Extreme value statistics

Motivation

In reinsurance applications, emphasis often lies on the modelling of extreme events, i.e. mostly events with

- low frequency
- high and often disastrous impact

A very practical tool in the analysis of such extreme events can be found in extreme value statistics, where the tail behavior of a distribution is characterized mainly by the extreme value index $\gamma$

- theoretical framework
- some estimators (extreme value index, small exceedance probabilities ... reinsurance premiums)
Extreme value statistics

Extreme value theory (first order framework)

Consider $X_{1,n} \leq \ldots \leq X_{n,n}$ independent and identically distributed random variables with common distribution function $F$.

- Maximum domain of attraction condition

$$
\lim_{n \to \infty} P \left( \frac{X_{n,n} - b_n}{a_n} \leq x \right) = H(x)
$$

- Limit necessarily of generalized extreme value type

$$
H_\gamma(x) = \begin{cases} 
\exp \left( - (1 + \gamma x)^{-\frac{1}{\gamma}} \right), & 1 + \gamma x > 0, \gamma \neq 0 \\
\exp \left( - \exp (-x) \right), & x \in \mathbb{R}, \gamma = 0
\end{cases}
$$

Then $F$ is said to belong to the maximum domain of attraction of the extreme value distribution $H_\gamma$, denoted as $F \in \mathcal{D}(H_\gamma)$.
Extreme value statistics

**Extreme value theory** (first order framework)

- $\gamma < 0 \implies F$ belongs to the Weibull class
  (e.g. uniform, beta and reversed Burr distribution)

- $\gamma = 0 \implies F$ belongs to the Gumbel class
  (e.g. exponential, normal and gamma distribution)

- $\gamma > 0 \implies F$ belongs to the Fréchet class
  (e.g. Fréchet, Pareto and Burr distribution)
Extreme value statistics

Extreme value theory (first order framework)

- $\gamma < 0 \implies F$ belongs to the Weibull class
  (e.g. uniform, beta and reversed Burr distribution)

- $\gamma = 0 \implies F$ belongs to the Gumbel class
  (e.g. exponential, normal and gamma distribution)

- $\gamma > 0 \implies F$ belongs to the Fréchet class
  (e.g. Fréchet, Pareto and Burr distribution)

\[
\bar{F}(x) = x^{-1/\gamma} l_F(x) \implies \text{Pareto-type tail}
\]
\[
(l_F \text{ slowly varying at infinity, i.e. } \frac{l_F(tx)}{l_F(t)} \xrightarrow{t \to \infty} 1 \text{ for all } x > 0)
\]
\[
U(x) = Q(1 - 1/x) = x^{\gamma} l_U(x)
\]
\[
(l_U \text{ slowly varying at infinity})
\]
Extreme value statistics

Estimating reinsurance premiums (Karamata theorem)

Assume \( G(t) = g(1/t) \) to be regularly varying at infinity with index of regular variation \( \beta \), i.e. \( G(t) = t^\beta l_G(t) \)

(\( l_G \) slowly varying at infinity)

\[
\downarrow \bar{F}(x) = x^{-1/\gamma} l_F(x)
\]

\[
\Pi(R) = \int_R^\infty x^{\beta/\gamma} l_{GF}(x) dx
\]

(\( l_{GF} \) slowly varying at infinity)

Karamata \( \downarrow \gamma < -\beta \)

\[
\Pi(R) \sim \dot{\Pi}(R) = \frac{1}{-\beta/\gamma - 1} R g\left(\bar{F}(R)\right)
\]

when the retention level \( R \) tends to infinity
\[ \Pi(R) \sim \hat{\Pi}(R) = \frac{1}{-\beta/\gamma - 1} Rg(\bar{F}(R)) \]

\[ \Downarrow \]

Premium estimator:  
\[ \hat{\Pi}(R) = \frac{1}{-\beta/\gamma - 1} Rg(\hat{p}_R) \]

\[ \sim \text{ large retention level} \]
\[ \sim \text{ regularly varying distortion} \]
Extreme value statistics

Estimating reinsurance premiums

\[ \Pi(R) \sim \hat{\Pi}(R) = \frac{1}{-\beta/\gamma - 1} Rg(\bar{F}(R)) \]

↓

Premium estimator:

\[ \hat{\Pi}(R) = \frac{1}{-\beta/\gamma - 1} Rg(\hat{p}_R) \]

\[ \sim \text{large retention level} \]

\[ \sim \text{regularly varying distortion} \]

**Net premium principle**

\[ g(x) = x \sim \left\{ \begin{array}{l}
G(x) = x^{-1}, \\
\beta = -1,
\end{array} \right. \]

\[ l_G(x) = 1. \]
Extreme value statistics

Estimating reinsurance premiums

\[ \Pi(R) \sim \hat{\Pi}(R) = \frac{1}{-\beta/\gamma - 1} Rg(\hat{F}(R)) \]

\[ \downarrow \]

Premium estimator: \[ \hat{\Pi}(R) = \frac{1}{-\beta/\gamma - 1} Rg(\hat{p}_R) \]

\[ \leadsto \text{large retention level} \]

\[ \leadsto \text{regularly varying distortion} \]

Dual-power transform principle (\( \alpha > 1 \))

\[ g(x) = 1 - (1 - x)^\alpha \leadsto \begin{cases} 
G(x) = x^{-1} \{ \alpha - \frac{\alpha(\alpha - 1)}{2} x^{-1} + o(x^{-1}) \} \\
\beta = -1 \\
l_G(x) = \alpha - \frac{\alpha(\alpha - 1)}{2} x^{-1} + o(x^{-1}) 
\end{cases} \]
Extreme value statistics

Estimating reinsurance premiums

\[ \Pi(R) \sim \hat{\Pi}(R) = \frac{1}{-\beta/\gamma - 1} \text{Rg}(\bar{F}(R)) \]

\[ \downarrow \]

Premium estimator: \[ \hat{\Pi}(R) = \frac{1}{-\beta/\gamma - 1} \text{Rg}(\hat{p}_R) \]

\[ \sim \] large retention level
\[ \sim \] regularly varying distortion

Needed:

- estimator \( \hat{p}_R \) of small exceedance probability \( p_R = \bar{F}(R) \)
- estimator \( \hat{\gamma} \) of the tail index \( \gamma \)
Extreme value statistics

Estimating reinsurance premiums \( (\rho_{R_n} \to 0, \text{ as } n \to \infty) \)

\[
\bar{F}(x) = x^{-1/\gamma} l_F(x) \text{ with } l_F \text{ slowly varying at infinity}
\]

\[
\lim_{t \to \infty} \frac{\bar{F}(\lambda t)}{\bar{F}(t)} = \lambda^{-1/\gamma} \lim_{t \to \infty} \frac{l_F(\lambda t)}{l_F(t)} = \lambda^{-1/\gamma}
\]

for all \( \lambda > 0 \)

- for large values \( y = \lambda t \) we can approximate \( \bar{F}(y) \) by

\[
\bar{F}(y) \approx \bar{F}(t) \left( \frac{y}{t} \right)^{-1/\gamma}
\]

- with a large threshold \( t = X_{n-k,n} \) and estimator \( \hat{\gamma}_k \) this leads to

\[
\hat{p}_{k,R_n} = \left( \frac{k + 1}{n + 1} \right) \left( \frac{R_n}{X_{n-k,n}} \right)^{-1/\hat{\gamma}_k}
\]
Consider scaled log-spacings \( Z_{jk} = j \left( \log X_{n-j+1,n} - \log X_{n-j,n} \right) \)
for \( j = 1, \ldots, k \) above a large threshold \( t = X_{n-k,n} \)

- \( \gamma > 0 \) : Feuerverger and Hall (1999), Beirlant et al. (1999)

\[
Z_{jk} \approx_d \gamma g_{jk}
\]

with \( g_{jk} \) i.i.d. standard exponential random variables

- Maximum likelihood estimator on the scaled log-spacings \( Z_{jk} \)
for large threshold \( X_{n-k,n} \) \((1 \leq j \leq k)\)

\[
\hat{\gamma}_{k,H} = \frac{1}{k} \sum_{j=1}^{k} Z_{jk}
\]

Hill (1975)
Finite sample behavior

Fréchet distribution

We consider 100 simulated samples of size 500 from a Fréchet(\(\eta\))
distribution with \(\eta = 4\) \((\gamma = 1/4)\)

\[
\bar{F}(x) = 1 - \exp(-x^{-\eta})
\]

and retention level \(R = U(x)\) with \(x = 1.025n\)
Finite sample behavior

Burr distribution

We consider 100 simulated samples of size 500 from a Burr($\beta$, $\tau$, $\lambda$) distribution with $\beta = 1$, $\tau = 4$ and $\lambda = 1$ ($\gamma = 1/4$)

$$
\bar{F}(x) = \beta^\lambda (\beta + x^\tau)^{-\lambda}
$$

and retention level $R = U(n)$

(Net)

(DP $\alpha = 1.366$)
Asymptotic results

Motivation

In order for these extreme value estimators of reinsurance premiums to become practically applicable, an optimal selection of the tail sample fraction $k$ seems to be desirable.

A bias-variance trade off through a minimizing asymptotic mean squared error criterion seems preferable in order to find an optimal tail sample fraction $k_o$.

- asymptotic bias term due to estimating $\hat{I}(R)$ by $\hat{\hat{I}}(R)$.
- asymptotic error term due to approximating $\hat{I}(R)$ by $\hat{\hat{I}}(R)$. 
Asymptotic results

Extreme value theory (second order framework)

Most extreme value estimators suffer from bias due to a slow convergence rate in

\[ \lim_{x \to \infty} \frac{l(\lambda x)}{l(x)} = 1 \]

for all \( \lambda > 0 \)
Asymptotic results

**Extreme value theory** (second order framework)

Most extreme value estimators suffer from bias due to a slow convergence rate in

\[
\lim_{x \to \infty} \frac{l(\lambda x)}{l(x)} = 1
\]

for all \( \lambda > 0 \)

\[ \rightsquigarrow \text{Parametrization of the rate of convergence:} \]

A slowly varying function \( l \) satisfies \((R_{\rho,b})\) for some constant \( \rho < 0 \) and rate function \( b \) satisfying \( b(x) \to 0 \), as \( x \to \infty \), if for all \( \lambda > 1 \)

\[
\frac{\left( \frac{l(\lambda x)}{l(x)} - 1 \right)}{b(x)} \to \frac{\lambda^\rho - 1}{\rho}, \quad \text{as} \quad x \to \infty
\]

Bingham et al. (1987)
**Lemma 1.** Suppose that $l_U \in \mathcal{R}_{\tilde{\rho}, \tilde{b}}$ and $l_G \in \mathcal{R}_{\rho^*, b^*}$, then as $R \to \infty$, the rate of convergence of the premium approximation $\dot{\Pi}(R)$ can be characterized as

\[
\frac{\Pi(R)}{\dot{\Pi}(R)} = 1 - \frac{1}{\beta + \rho^* + \gamma} b^* \left(1/\bar{F}(R)\right) \\
+ \frac{\beta}{\gamma(\beta + \tilde{\rho} + \gamma)} \tilde{b} \left(1/\bar{F}(R)\right) + o\left(r \left(1/\bar{F}(R)\right)\right)
\]

where the remainder function $r$ is defined as

\[
\sim \quad r = \tilde{b} \text{ in case } b^*(x) = o\left(\tilde{b}(x)\right) \\
\sim \quad r = b^* \text{ in case } \tilde{b}(x) = o\left(b^*(x)\right) \\
\sim \quad r = \tilde{b} \text{ or } r = b^* \text{ in case } \tilde{b}(x) \sim cb^*(x), \text{ with } c \neq 0
\]
Asymptotic results

Asymptotic normality and bias

**Theorem 1.** Suppose that \( l_U \in \mathcal{R}_{\tilde{\rho}, \tilde{b}} \). Let \( ABias(\hat{\gamma}_k) \sim I_{\gamma, \tilde{\rho}} \tilde{b}(n/k) \) and assume that as \( k, n \to \infty \) such that \( k/n \to 0 \) and \( \sqrt{k} \tilde{b}(n/k) \to 0 \),

\[
\sqrt{k} (\hat{\gamma}_k - \gamma) \to_d N(0, \sigma_\gamma^2)
\]

for some asymptotic variance \( \sigma_\gamma^2 \). Then the tail estimator \( \hat{p}_{k,R_n} \) satisfies

- as \( R_n \to \infty \) such that \( \tilde{a}_n = \frac{k+1}{(n+1)p_{R_n}} \to \infty \) and \( \frac{\log \tilde{a}_n}{\sqrt{k}} \to 0 \), that

\[
ABias(\hat{p}_{k,R_n}) \sim \frac{p_{R_n} \log \tilde{a}_n}{\gamma} I_{\gamma, \tilde{\rho}} \tilde{b}(n/k)
\]

- and furthermore, as \( k, n \to \infty \) such that \( k/n \to 0 \) and \( \sqrt{k} \tilde{b}(n/k) \to 0 \), that

\[
\frac{\gamma}{\log \tilde{a}_n} \sqrt{k} \left( \frac{\hat{p}_{k,R_n}}{p_{R_n}} - 1 \right) \to_d N(0, \sigma_\gamma^2)
\]
Theorem 2. Suppose that \( l_U \in \mathcal{R}_{\hat{\rho}, \tilde{b}} \) and \( l_G \in \mathcal{R}_{\rho^*, b^*} \). Moreover let \( ABias(\hat{\gamma}_k) \sim I_{\gamma, \hat{\rho}}\tilde{b}(n/k) \), and assume that as \( k, n \to \infty \) such that \( k/n \to 0 \) and \( \sqrt{k} \tilde{b}(n/k) \to 0 \), \( \sqrt{k} (\hat{\gamma}_k - \gamma) \to_d N(0, \sigma^2_{\gamma}) \) for some asymptotic variance \( \sigma^2_{\gamma} \). Then, when \( \gamma < -\beta \), the premium estimator \( \hat{\Pi}(R) \) based on tail estimator \( \hat{p}_{k,R_n} \) satisfies

- as \( R_n \to \infty \) such that \( \tilde{a}_n = \frac{k+1}{(n+1)p_{R_n}} \to \infty \) and \( \frac{\log \tilde{a}_n}{\sqrt{k}} \to 0 \), that

\[
ABias(\hat{\Pi}(R)) \sim -\frac{\beta \log \tilde{a}_n \Pi(R)}{\gamma} I_{\gamma, \hat{\rho}}\tilde{b}(n/k) + \frac{\Pi(R)}{\beta + \rho^* + \gamma} b_{\rho^*} (1/F(R))
\]

and furthermore, as \( k, n \to \infty \) such that \( k/n \to 0 \), \( \sqrt{k} \tilde{b}(n/k) \to 0 \) and \( \frac{\sqrt{k}}{\log \tilde{a}_n} b^* (1/p_R) \to 0 \), that

\[
-\frac{\gamma}{\beta \log \tilde{a}_n} \sqrt{k} \left( \frac{\hat{\Pi}(R)}{\Pi(R)} - 1 \right) \to_d N(0, \sigma^2_{\gamma})
\]
Asymptotic results

Asymptotic normality and bias (AMSE)

In order to find an optimal threshold for $\hat{\Pi}(R)$, the asymptotic mean squared error that has to be minimized with respect to $k$ then is proportional to

$$\frac{\beta^2 (\log \tilde{a}_n)^2 \sigma^2_{\gamma}}{\gamma^2} + \left\{ \frac{1}{\beta + \rho^* + \gamma} b_{\rho^*} \left( \frac{1}{p_R} \right) - \frac{\beta \log \tilde{a}_n}{\gamma} I_{\gamma, \tilde{\rho}} \tilde{b}(n/k) \right\}^2$$

Example: Hill based estimator (Haeusler and Teugels, 1985)

- $\sigma^2_{\gamma} = \gamma^2$
- $I_{\gamma, \tilde{\rho}} = \frac{1}{1-\tilde{\rho}}$

Need are second order estimates of $\gamma$, $\tilde{b}(n/k)$, $\tilde{\rho}$ and $p_R$.
Asymptotic results

Asymptotic normality and bias (second order estimates)

- $\gamma > 0$ : Feuerverger and Hall (1999), Beirlant et al. (1999)

  \[ Z_{jk} \approx d \left( \gamma + \tilde{b}_{n,k} \left( \frac{j}{k+1} \right)^{-\tilde{\rho}} \right) g_{jk}, \]

  Maximum likelihood estimator on the scaled log-spacings $Z_{jk}$ for large threshold $X_{n-k,n}$ ($1 \leq j \leq k$)

- Matthys et al. (2004) numerical inversion of approximation for quantile function $x_p = Q(1 - p)$ given by

  \[ x_p \approx X_{n-k,n} \left( \frac{k+1}{(n+1)p} \right)^\gamma \exp \left( \tilde{b}_{n,k} \frac{\left( \frac{k+1}{(n+1)p} \right)^{\tilde{\rho}} - 1}{\tilde{\rho}} \right) \]
Secura Belgian Re data

We consider the Secura Belgian Re data set on automobile claims from 1998 until 2001, as can be found in Beirlant et al. (2004). The data set consists of \( n = 371 \) claims which are at least as large as 1,200,000 Euro.

\[ R = 5,000,000 \text{ Euro} \implies (\text{Net}) \quad k_o = 95 \quad \hat{H}_{k_o}(R) = 41,798.13 \]
Finite sample behavior

Fréchet $(4)$ distribution

$$\alpha = 1.366$$

Dual Power premium

Net premium
Finite sample behavior

Burr(1,4,1) distribution

(Net)

(DP $\alpha = 1.366$)
Conclusion

... on univariate extreme value statistics and the estimation of insurance premiums for excess-of-loss reinsurance policies in excess of a high retention level ... with special attention to heavy-tailed distributions and Wang’s premium principle as a generalization to the net premium principle ...