

# PREMIUM CALCULATION FOR FAT-TAILED RISK

BY

ROGER GAY

## ABSTRACT

When insurance claims are governed by fat-tailed distributions considerable uncertainty about the value of the tail-index is often inescapable. In this paper, using the theory of risk aversion, a new premium principle (the power principle – analogous to the exponential principle for thin-tailed claims) is established and its properties investigated. Applied to claims arising from generalized Pareto distributions, the resultant premium is shown to be the ratio of the two largest expected claims, for which the ratio of the actual claims is an unbiased as well as a consistent estimator. Whereas thin-tailed claim premiums are determined largely by the first two moments of the claims distribution, fat-tailed claim premiums are determined by the first two extremes. The context of risk-aversion leads to a natural model for incorporating tail-index uncertainty into premiums, which nevertheless leaves the basic ratio structure unaltered. To illustrate the theory, possible ‘premiums’ for US hurricane data are examined, which utilize the consistent pattern of observed extremes.

## KEYWORDS

Exponential principle, power principle, constant risk aversion, tail-index uncertainty, expected value of extremes.

## 1. INTRODUCTION

### 1.1. Ambit and structure of the paper

A new premium principle (the power principle) for fat-tailed risk is described and its properties discussed. While for this type of risk the precise value of the tail index is of first importance in determination of claim severity, its value is very difficult to estimate precisely, particularly with the sorts of sample sizes likely to be available to insurers in practice, (see e.g. De Haan (1994), Beirlant *et al.* (1994), Coles (2001), Hsieh (2004), Huisman *et al.* (2001)).

The power principle can be used to provide a coherent framework in which to determine premiums when uncertainty in the tail-index is suitably modelled. The ‘uncertainty premiums’ – risk-averse premiums which take account of tail-index imprecision – have a transparent and suggestive structure which may in

itself assist in the tail-index estimation problem. Uncertainty premiums can be regarded either as weighted mean values of risk-averse premiums derived from the power principle, or as Bayesian posterior means.

In their review of Bayesian methods in extreme value modelling, Coles and Powell (1996) write “extreme value problems are characterized by scarcity of data and the requirement of modelling where the data are most sparse. This presents a dilemma when considering a Bayesian approach to inference: the value of additional prior information is likely to be substantial, but the plausibility of formulating such prior knowledge is questionable”.

In consideration of a problem (not unrelated to that of this paper) – forecasting the next record catastrophic loss – Hsieh (2004) incorporates prior knowledge by employing the input of an expert, using prior gamma densities Hsieh (2001a, 2001b) to estimate key predictive parameters. Hill (1994) first proposed a forecasting model linking extreme value and Bayesian methods.

Coles and Tawn (1996) have demonstrated how, within a Bayesian framework, estimates supplemented by expert information lead to an improved analysis of extreme rainfall data. For other applications of Bayesian methods integrated with extreme value theory see Pickands (1994), Smith (2000) and Smith and Goodman (2000).

In this paper, transformed beta densities are used as priors for the tail-index. They are shown to arise naturally in the context of constant relative risk-aversion, under which fat-tailed premiums are determined. Plausibility issues *can* arise if priors are made ‘too informative’ (Appendix, Note 3, and for example issues raised in Pickands (1994), Yuan and Clarke (1999)).

The structure of the remainder of the paper is as follows. The exponential principle for thin-tailed risk is described; the power principle is introduced as the analogous principle for fat-tailed risk, and its properties discussed.

When applied to the generalised Pareto distribution (inclusive of a broad range of fat-tailed laws for ‘exceedances’ over a sufficiently high threshold) risk-neutral premiums for claims are shown to be ratios of expected values of the two largest order statistics or first two extremes. The premiums are thus linked to a basic idea in extreme value statistics that the ‘distance’ between two subsequent extreme order statistics can be used to characterise tail heaviness (an idea which can be traced back at least to Hill, (1975)).

This ratio structure of premiums remains when uncertainty in the tail-index is appropriately modelled.

An example employing U.S. hurricane data with a characteristic pattern of extremes is used to illustrate ideas, with a tail-index estimate derived from relations between extremes being preferred to the maximum likelihood estimate.

## 1.2. Thin-tailed risk and the exponential principle

When general insurance claims are *thin-tailed*, i.e. possess moments of all orders, the exponential principle (see Rolski *et al.* (1999), p. 80) with foundations in absolute risk aversion (Pratt (1964), and Arrow (1971), Appendix, Note 1) applies.

Under the exponential principle, premium  $P$  for claim  $X$  is set as

$$P = \ln\{M_X(s)\}/s, \quad (s > 0) \quad (1.1)$$

and

$$P \approx \mu + \frac{1}{2}s\sigma^2 + o(s) \quad (1.2)$$

where  $M_X(\cdot)$ ,  $\mu$ ,  $\sigma^2$  are respectively the moment generating function, the mean and variance of  $X$ . Parameter  $s$  measures the *constant absolute risk aversion* of the insurer. Bowers *et al.* (1986, p. 7) infer that insurers are *risk-neutral or only mildly risk-averse*. Thus  $s$  is small and approximation (1.2), in terms of the first two moments of the distribution follows; the notation “ $o(s)$ ” of (1.2) means that  $\{P - (\mu + \frac{1}{2}s\sigma^2)\}/s \rightarrow 0$  as  $s \rightarrow 0$ .

Rolski *et al.* (1999) p. 80 refer to (1.2) as the *variance principle* and catalogue some standard variants.

Indeed, the exponential principle (1.1) applies to *annual aggregate claims*  $S_N$  where

$$S_N = X_1 + X_2 + \dots + X_N \quad (1.3)$$

the  $\{X_i\}$  are identically and independently distributed (i.i.d), and  $N$  is the claim number independent of  $\{X_i\}$ , (frequently assumed to be a Poisson or negative binomial random variable), since

$$(M_{S_N}(s) = M_N[\ln\{M_X(s)\}]) \quad (1.4)$$

(dropping the subscript on the i.i.d.  $\{X_i\}$ ,  $i \geq 1$ ;  $X$  is any representative claim) so that

$$P = E[X] \times E[N] + \frac{1}{2}s\{\text{Var}(X)E[N] + E[X]^2 \text{Var}(N)\} + o(s) \quad (1.5)$$

In this sense a fairly complete theory exists for thin-tailed risk, simple and elegant results for the aggregate premium deriving from particular distributional choices of  $X$  and  $N$ , including results for stop-loss insurance.

The exponential principle is conveniently established using the *pricing function*

$$m_\theta(x) = \exp(\theta x) \quad (1.6)$$

with ‘risk parameter’  $\theta = s$ , and the *pricing rule*

$$m_\theta(P) = E[m_\theta(X)] \quad (1.7)$$

The pricing mechanism can be related to, but is more general than, the expected utility principle, Appendix; Note 1, and Gay (2004a).

Use of the pricing rule (1.7) in conjunction with a pricing function  $m_\theta(x)$  allows for some unification of premium principles:

- (i)  $m(x) = x$  leads to the risk-neutral premium  $P = E[X]$
- (ii)  $m_s(x) = \exp(sx)$  leads to the exponential principle and the variants deriving from it.
- (iii) For any claim  $X > 0$  with absolutely continuous distribution function  $F(x)$ , choosing

$$m_\theta(x) = \{F(x)\}^\theta, \quad (\theta > 0) \quad (1.8)$$

in conjunction with (1.7) leads to the *quantile principle* (Rolski *et al.*, (1999), p. 82). Under the quantile principle  $P$  is chosen as a suitable quantile of the claim distribution by the rule

$$F(P) = q(\theta) = (\theta + 1)^{-1/\theta} \quad (1.9)$$

Note that  $q(\theta)$  is strictly increasing with  $\theta$ , and  $e^{-1} < q(\theta) < 1$ . See Gay (2004a)

- (iv) The ‘power principle’ for fat-tailed claims derived below has  $m_\alpha(x) = x^{\alpha+1}$  ( $\alpha > 0$ ).

### 1.3. Fat-tailed risk and the power principle

Individual claims  $\{X_i\}$  of (1.3) are governed by fat-tailed distribution  $F(x)$  when

$$1 - F(x) = x^{-\delta} L(x), \quad (\text{for large } x > 0, \delta > 0) \quad (1.10)$$

$L(x)$  is ‘slowly varying at  $\infty$ ’; i.e. for any  $a > 0$  and  $x$  sufficiently large, i.e.

$$\lim_{x \rightarrow \infty} L(ax) / L(x) = 1$$

See for instance Feller, (1971, p. 278), Embrechts *et al.* (1997, p. 131).

Thus the claim random variable  $X$  has moments  $E[X^k]$  only of order  $k < \delta$ .

#### *Pareto and Generalized Pareto*

Our principal concern is with Pareto claims from distribution

$$F(v) = 1 - (1 + v/\lambda)^{-\delta}, \quad (v > 0, \delta > 0) \quad (1.11)$$

Formula (1.11) describes virtually the simplest form of pure tail distribution. But its role in extreme value theory is much more important. For a large enough threshold  $u$ , function (1.11) approximates for a broad class of random variables, *to the distribution of exceedances above  $u$* , i.e. is the generalized Pareto distribution for fat tails (see for instance, Balkema and de Haan (1974), Pickands (1975), Embrechts *et al.* (1997, Chapter 3), Drees *et al.* (2004) for precise statement of convergence results). In these circumstances,  $\delta$  is independent of  $u$ , while  $\lambda$  is not (although Leadbetter (1991) provides conditions under which  $\lambda$  also is independent of  $u$ ).

If the claim  $V$  in (1.11) is measured in terms of scale parameter  $\lambda$ ; i.e.  $V = X\lambda$  the simplified distribution for  $X$

$$F(x) = 1 - (1 + x)^{-\delta}, \quad (x > 0, \delta > 0) \quad (1.12)$$

ensues. Fat-tailed claims with limited moments have no moment generating functions and hence no constant absolute risk-averse premiums under (1.1).

#### *Type I extreme value (Frechet) distribution*

Another distribution used to model fat-tailed claims in its own right, but more importantly, is inclusive of a broad class of fat-tailed distributions  $F(\cdot)$  of type (1.10) is the Frechet distribution.

The density of the  $k$ th extreme  $X_{(k)}^*$ , the limiting value of the normalized order statistic  $X_{(n-k+1)}/v_n$  from distribution  $F(\cdot)$  for large  $n$ , denoted by  $f_k(x)$  is

$$f_k(x) = \delta x^{-k\delta-1} \exp\{-x^{-\delta}\} / \Gamma(k) \quad (k \geq 1, x > 0, \delta > 0) \quad (1.13)$$

The sequence of normalizing constants  $\{v_n\}$  – ‘the tail-quantile function – derives from  $n\{1 - F(v_n)\} = 1$  for all  $k$  (Kendall and Stewart, (1969), p. 331, David and Nagaraja, (2003), p. 306).

The tail-quantile function plays a prominent role in extreme value theory, determining among other things membership of  $F(\cdot)$  within the domain of attraction of (1.13). Thus  $v_n = v_n(\rho)$ ,  $\rho = 1/\delta$  is distribution-specific. For example,  $v_n = \lambda n^\rho$  for Pareto (1.11). Further examples of  $v_n$  are to be found in Embrechts *et al.* (1997), Section 3.4, Teugels and Vanroelen (2004), Beirlant *et al.* (1996)

#### *Relation between Generalized Pareto and Frechet Premiums*

Our focus in this paper is on large claims from fat-tailed distributions. Use of Pareto distribution (1.11) will enable premium determination provided the claims exceed a sufficiently large threshold. The Frechet densities describe behaviour of the limiting values of ranked claims.

*The class of fat-tailed distributions from (1.10) which have generalized Pareto distributions when exceedances are above a sufficiently high level, is precisely the class within the domain of attraction of the Frechet distribution Pickands (1975), Embrechts et al. (1997, p. 131), (Drees, Ferreira and De Haan (2004), Nagaraja, (2004))*

#### *The Power Principle*

For fat-tailed risk, premium  $P$  can be set using the power principle. Deriving from the theory of risk-aversion it is the insurance premium implied by *constant relative risk-aversion* measured by parameter  $\alpha$  via the pricing function

$$m(x) = x^{\alpha+1} \quad (x > 0, \alpha \geq 0) \quad (1.14)$$

and pricing rule (1.7); (see Appendix: Note 1).

For any claims random variable  $X$  the power principle leads to a premium  $P$  determined by the risk-aversion coefficient  $\alpha$  as

$$P(\alpha) = \left\{ \mu'_{\alpha+1} \right\}^{\frac{1}{\alpha+1}} \tag{1.15}$$

where  $\mu'_k$  is the  $k$ th raw moment of  $X$ ; ( $\mu'_1 = \mu = E[X]$ )

*Properties of Power Principle premiums*

Premiums determined by (1.15) have certain desirable properties of a ‘good’ premium (Rolski, *et al.* 1999, p. 79). In obvious notation:

- (i)  $P(0) = \mu$ ;  $P(\alpha)$  is increasing in  $\alpha + 1 \geq 0$  (see for instance Puri and Sen (1971, p. 12), reflecting risk-aversion increasing with  $\alpha$ .)
- (ii) For any  $\alpha \geq 0$ ,  $P$  is subadditive:  $P_{X+Y}(\alpha) \leq P_X(\alpha) + P_Y(\alpha)$  (from Minkowski’s inequality) ‘splitting risks won’t reduce premiums’
- (iii) Stochastic dominance:  
 For any  $\alpha \geq 0$ ,  $X \succ Y \Rightarrow P_X(\alpha) > P_Y(\alpha)$   
 (follows from the fact that for  $X > 0$ ,  $E[X^{\alpha+1}] = \int_0^\infty (1 - F_X(x^{\frac{1}{\alpha+1}})) dx$ )
- (iv) For any  $\alpha \geq 0$ ,  $P_a(\alpha) = a$  for all constants  $a > 0$   
 ‘no unjustified premium loading’
- (v) Proportionality: For any  $\alpha \geq 0$ ,  $P_{aX}(\alpha) = aP_X(\alpha)$

(The integral  $\int_0^\infty x^{\alpha+1} f(x) dx$  implicit in (1.15) is *the Mellin transform* of  $f(x)$  for which extensive tables are available, e.g. Oberhettinger (1974)).

**1.4. Risk-averse premiums**

*Premiums for large fat-tailed claims*

When claim  $X$  has distribution  $F(\cdot)$  from (1.10) its conditional distribution above sufficient level  $u$  is Pareto (1.11) with the parameter  $\lambda$  reflective of the level.

To suit our further objectives, it will be helpful to incorporate  $\lambda$  in the claim. Thus applying (1.15) to Pareto (1.11) we have for claim  $(1 + X/\lambda)$  premium  $P$  is given by

$$\begin{aligned} P^{\alpha+1} &= \lambda^{-1} \int_0^\infty (1 + x/\lambda)^{\alpha+1} \delta (1 + x/\lambda)^{-\delta-1} dx \\ &= \delta \{ \delta - (\alpha + 1) \}^{-1} \end{aligned}$$

whence

$$P(\alpha) = \left\{ 1 - \frac{\alpha + 1}{\delta} \right\}^{-1/(\alpha+1)} \tag{1.16}$$

We re-parameterise (1.16) as

$$P_\beta = (1 - \rho/\beta)^{-\beta} \tag{1.17}$$

where  $\beta = (\alpha + 1)^{-1}$  determines the fee the insurer charges for bearing risk of claims generated from a distribution with tail index  $\rho = 1/\delta$ .

**Example:** If  $\rho = 0.75$ ,  $\beta = 0.95$ ,  $P_{0.95} = 4.39$ , compared with  $P_{1.0} = \mu = 4.00$ . Thus an insurer's  $\beta$  of 0.95 implies a charge of about 10% above the risk-neutral premium for bearing risk when  $\rho = 0.75$ .

**Remarks:**

(i) *Maximum tail-fatness*

Formula (1.17) shows that the insurer's level of constant relative risk-aversion  $\alpha$  determines the *maximum tail-index*  $\beta = (\alpha + 1)^{-1} = \rho_{\max} = 1/\delta_{\min}$  with which the insurer will deal for a particular class of fat-tailed insurance claims.

(ii) *When there is no mean claim  $E[X]$*

Formula (1.17) assumes  $\delta > 1$  ( $\rho < 1$ ) and existence of  $E[X]$ . If  $\delta \leq 1$  ( $\rho \geq 1$ ) then cover can be offered on part of Pareto claim  $(1 + X/\lambda)^\rho$ , ( $\rho < 1$ ).

Application of (1.15) results in premiums

$$P_\beta = (1 - \rho\phi/\beta)^{-\beta} \tag{1.18}$$

for the Pareto partial claims.

(iii) *Risk-neutral premiums*

(The insurer is risk-neutral or only mildly risk-averse). For Pareto (1.11) direct integration shows that  $E[1 + X/\lambda] = (1 - \rho)^{-1}$  confirming premium formula (1.17) with  $\beta = 1$ .

However, in Appendix: Note 2 it is shown that

$$(1 - \rho)^{-1} = \frac{E[1 + X_{(n)}/\lambda]}{E[1 + X_{(n-1)}/\lambda]} \tag{1.19}$$

where  $X_{(k)}$  is the  $k$ th Pareto order statistic, *the ratio being independent of sample size  $n$* .

This limit for the ratio of the two largest extremes is true for all fat-tailed distributions in the Frechet domain of attraction, i.e. as  $n \rightarrow \infty$

$$\frac{E[1 + X_{(n)}/\lambda]}{E[1 + X_{(n-1)}/\lambda]} \rightarrow \frac{E[v_n X_{(1)}^*]}{E[v_n X_{(2)}^*]} = \frac{E[X_{(1)}^*]}{E[X_{(2)}^*]} = (1 - \rho)^{-1}$$

Similar results hold for ratios of other expected order statistics and extremes. More generally as  $n \rightarrow \infty$

$$\frac{E[1 + X_{(n)} / \lambda]}{E[1 + X_{(n-k)} / \lambda]} \rightarrow \frac{E[X_{(1)}^*]}{E[X_{(1+k)}^*]} = kB(k, 1 - \rho) \quad (1.20)$$

where  $B(a, b) = \Gamma(a) \Gamma(b) / \Gamma(a + b)$ ,  $a > 0$ ,  $b > 0$ , see Appendix, Note 2).

As  $\rho \uparrow 1$  the ratio (1.19) becomes large. ‘The expected value of the second largest claim is small compared with the expected value of the largest’.

*Total expected claims and largest expected claims*

From (1.20) we obtain relationship

$$\begin{aligned} \frac{E[1 + X_{(n-1)} / \lambda]}{E[1 + X_{(n)} / \lambda]} &= (1 - \rho) \\ \frac{E[1 + X_{(n-2)} / \lambda]}{E[1 + X_{(n)} / \lambda]} &= (1 - \rho)(1 - \rho/2) \\ \frac{E[1 + X_{(n-3)} / \lambda]}{E[1 + X_{(n)} / \lambda]} &= (1 - \rho)(1 - \rho/2)(1 - \rho/3) \end{aligned}$$

etc.

In this sense annual aggregate claims  $S_N$  of (1.3) for classes of insurance with  $\rho \in [\frac{1}{2}, 1)$  can be expected to be mainly attributable to the largest or largest few claims. From these relationships, numerical values of the series:

- (i) for  $\rho = \frac{1}{2}$  gives ratios  $\frac{1}{2}, \frac{3}{8}, \frac{5}{16}, \dots$  relating ‘descending order’ expected large claims to the largest expected claim,
- (ii) for  $\rho = \frac{3}{4}$  the series is  $\frac{1}{4}, \frac{5}{32}, \frac{15}{128}, \dots$ .

(iv) *Existence of an unbiased estimator*

More remarkably, for Pareto claims, the risk-neutral premium for claim  $(1 + X/\lambda)$  admits an unbiased estimator. We set  $\lambda = 1$  for convenience. Then

$$\begin{aligned} E\left[\frac{1 + X_{(n)}}{1 + X_{(n-k)}} \middle| X_{(n-k)}, X_{(n-k-1)} \dots\right] &= E\left[\frac{1 + X_{(n)}}{1 + X_{(n-k)}} \middle| X_{(n-k)}\right] \\ &= E\left[\frac{1 + X_{(n)}}{1 + X_{(n-k)}}\right] \\ &= \frac{E[1 + X_{(n)}]}{E[1 + X_{(n-k)}]} \\ &= kB(k, 1 - \rho) \end{aligned}$$



and

$$E \left[ \frac{1 + X_{(n)}}{1 + X_{(n-k)}} \right] = E \left[ \frac{X_{(1)}^*}{X_{(1+k)}^*} \right] = \frac{E[X_{(1)}^*]}{E[X_{(1+k)}^*]} = kB(k, 1 - \rho) \quad (1.21)$$

(The proof appears in Appendix Note 2, B6).

A *consistent* and *unbiased* estimator  $\hat{P}$  for Pareto (1.11) claims in the form  $(1 + X/\lambda)$  is

$$\hat{P} = \frac{X_{(1)}^*}{X_{(2)}^*}$$

Hence the following general proposition:

*'Premiums for thin-tailed claims are determined largely by the first two moments of the distribution; for fat-tailed claims they are determined largely by the first two extremes'*

(v) Representation of premiums applicable under different conditions involves a number of notational changes. A glossary of premium notation appears at the end of the Appendix.

## 2. RISK-AVERSE PREMIUMS WHEN THE TAIL-FATNESS INDEX IS UNCERTAIN

### 2.1. Fat-tail claims models: large claims

A principal application of fat-tailed distributions is to model classes of insurance with potentially a few very large individual claims. We have in mind catastrophe, natural disaster, public liability, professional indemnity, industrial fire and the like. For (1.3) to provide a useful model for very large claims, the variance of individual claims must be infinite; otherwise aggregate claims will be governed by central limit theorems. This means that tail-fatness index  $\delta$  of (1.10), (1.11) and (1.12) cannot exceed 2 (or  $\rho \geq \frac{1}{2}$ ). The largest sample order statistic (extreme) has no second moment; all other order statistics (extremes) do have.

Various authors have remarked (e.g. Mikosch, (1997), Embrechts *et al.*, (1997) that the principal cost of annual aggregate claims in the classes of insurance involving fat-tailed claims above a high level (i.e. reinsurance claims) is usually attributable to a few of the largest claims. Embrechts *et al.* (1997), Section 8.2 mention the '20-80' rule-of-thumb used by actuaries, i.e. that in many large claim portfolios, 20% of claims account for 80% of total claim amount. Feller (1971), p. 288 and Mikosch (1997) provide probabilistic justification for this. Remark (iii) of Section 1.4 above provides more precise evidence from a different perspective.

Tail-fatness index estimation has been and is extensively studied, usually assuming large sample sizes are available. Some benchmark papers are Hill

(1975), Smith, (1987), Grimshaw (1990), Dekkers and De Haan (1993), Pickands (1994), Drees and Kaufmann (1998), Dekkers *et al.* (1999), Feuerverger and Hall (1999), Danielsson *et al.* (2001) and references therein). The index  $\rho$  is difficult to estimate, particularly in this range  $[\frac{1}{2}, 1)$  in which claim  $X$  is without a second moment (Smith (1985, 1987)), and when sample sizes available to an insurer, even with pooled data across an entire national industry are likely to be small. The difficulty remains when parametric or semi-parametric assumptions are made.

Many familiar principles used to set premiums for thin-tailed claims (the exponential principle, the variance principle and its variants) cannot be applied to claims without a second moment.

Furthermore, for claims  $X$  from Pareto (1.11) the mean  $\mu = E[1 + X/\lambda] = (1 - \rho)^{-1}$  is 'a rare event' as  $\rho \uparrow 1$  in the sense that  $\Pr(X > \mu) \rightarrow 0$ . The mean is larger than any quantile of the distribution as  $\rho$  approaches 1. Since  $\rho$  (and so  $\mu$ ) is never known precisely, it can be hazardous to use the quantile principle (1.9) to set premiums employing an estimate of  $\rho$  when  $\rho$  is supposed to be in the range  $[\frac{1}{2}, 1)$ . Premium  $P$  must be chosen as a quantile  $q(\alpha)$  with  $\alpha > \alpha_\mu$  where  $F(\mu) = q(\alpha_\mu)$ , but  $\mu$  is unknown, see Gay (2004a). In practical terms this difficulty can be managed by modeling tail-index uncertainty. In the sequel it is assumed that the claims mean  $E[X]$  exists so that  $\rho < 1$ .

## 2.2. A model for tail-fatness uncertainty

*“Statistical analyses of large claim data are based on extreme value theory and related methods. These are known to be very sensitive with respect to the tails of the distributions, and therefore the existence of one very large claim may justify the fit of a Pareto instead of a lognormal distribution, say.”* (Mikosch, 1997).

The functional form (1.17) of the premium  $P_\beta = (1 - \rho/\beta)^{-\beta}$  with  $\rho$  bounded by  $\beta = \rho_{\max}$  suggests the family of transformed beta variates

$$f_\rho(x) = vx^{v-1}/\beta^v, \quad (0 \leq x < \beta, v \geq 1) \quad (2.1)$$

as a natural model for uncertainty in  $\rho$ .

Allowance for uncertainty in  $\rho$  can be made by calculation of 'uncertainty premiums' as mean values of risk-averse premiums  $P_\beta$  weighted by (2.1)

Alternatively, distribution (2.1) can be regarded as a Bayesian prior for  $\rho$  with premium  $P(v, \beta) = E[P_\beta]$  the resultant posterior mean for specified  $v$ .

## 2.3. The nature of the uncertainty premiums

*Fat-tailed premiums.* The premium  $P(v, \beta)$  for claim  $[1 + X/\lambda]$  is obtained as the mean value or posterior mean of a Pareto premium, i.e. the expected value of the risk-averse premium  $P_\beta$  using density (2.1).

Thus

$$P(v, \beta) = E[P_\beta]$$

i.e.

$$P(v, \beta) = \int_0^\beta v\beta^{-v} x^{v-1} (1 - x/\beta)^{-\beta} dx \tag{2.2}$$

Substitution  $u = x/\beta$  in the integral leads immediately to

$$P(v, \beta) = vB(\theta, 1 - \beta) \tag{2.3}$$

Also by inspection of the integral (2.2), (replacing exponent  $\beta$  by  $2\beta$ );

$$E[P_\beta^2] = vB(v, 1 - 2\beta); \text{ for } x \in [0, \beta)$$

There is no posterior second moment (and no posterior variance) of the premium  $P_\beta = (1 - \rho/\beta)^{-\beta}$  for any  $\beta > \frac{1}{2}$ , the claims themselves having no variance for these values. We call  $P(v, \beta)$  *the uncertainty premium*.

**2.4. On a pragmatic measure taken by insurers**

In particular, when  $v = 1$  ('law of equal ignorance' for  $\rho$  on  $[0, \beta)$ )

$$\begin{aligned} P(1, \beta) &= (1 - \beta)^{-1} \\ &= \frac{E[1 + X_{(n)} / \lambda]}{E[1 + X_{(n-1)} / \lambda]} \\ &= E\left[\frac{X_{(1)}^*}{X_{(2)}^*}\right] \end{aligned} \tag{2.4}$$

- the ratio of the expected value (or expected value of the ratio) of the two largest Pareto order statistics/extremes.

*However the extremes are now based on tail-index  $\beta = \rho_{max}$*

Thus (2.4) provides a theoretical endorsement of the pragmatic measure insurers will almost certainly take; i.e. choose a prudently large value  $\beta$  of the tail-index to allow for imprecise knowledge of  $\rho$ .

The implication of this measure is now made explicit. *It is equivalent to an assumption of a uniform (i.e. the least informative) prior distribution of  $\rho$  on  $[0, \beta)$ .*

Also implicit in this assumption is that *the insurer is not ruling out the possibility of claims arising partly from a mixture of thin-tailed risks ( $\rho = 0$ )*. Issues of plausibility of the prior distribution *can* arise if one tries for instance to restrict the range of  $\rho$  values to  $[\frac{1}{2}, 1)$ . See Appendix Note 3.

**2.5. Ratio structure of the uncertainty premium  $P(v, \beta)$**

For Pareto claims, if  $k$  is an integer  $P(k, \beta) = kB(k, 1 - \beta)$  where

$$kB(k, 1 - \beta) = \frac{E[1 + X_{(n)}]}{E[1 + X_{(n-k)}]} = E\left[\frac{X_{(1)}^*}{X_{(1+k)}^*}\right] \tag{2.5}$$

independently of  $n$  (see Appendix: Note 2), the order statistics being based on tail-index  $\beta = \rho_{\max}$ .

*The ratio structure of premiums is preserved in the uncertainty premiums.*

However premiums  $P(v, \beta)$  which utilise only integer values of  $v$  will have large gaps.

If  $v$  is not an integer, the notional structure of  $P(v, \beta)$  as a ratio of extremes can be preserved by defining fractional order statistics and fractional extremes via their densities.

For a sample from distribution  $F(\cdot)$  (density  $f(\cdot)$ ), define the  $v$ th fractional order statistic  $X_{(k,n)}^{(v)}$  interpolating between  $X_{(k)}$  and  $X_{(k+1)}$  where  $k = [v]$  ('the integer part of  $v$ ') for any value  $v$  such that  $1 \leq k \leq v < k + 1 \leq n$  via its density  $f_v(x)$  where

$$f_v(x) = \{F(x)\}^{v-1} \{1 - F(x)\}^{n-v} f(x) / B(v, n + 1 - v) \tag{2.6}$$

where  $B(v, n + 1 - v) = \Gamma(v)\Gamma(n + 1 - v) / \Gamma(n + 1)$ .

Fractional extremes can be similarly defined. See Appendix Note 2.

Then the ratio structure is preserved and provides continuously changing uncertainty premiums in the form

$$\begin{aligned} P(v, \beta) &= vB(v, 1 - \beta) \\ &= \frac{E[1 + X_{(n)} / \lambda]}{E[1 + X_{(n-k,n)}^{(v)} / \lambda]} \\ &= E\left[\frac{X_{(1)}^*}{X_{(k)}^{*(v)}}\right] \end{aligned} \tag{2.7}$$

Gay (2004b) provides a numerical illustration.

We now investigate how the relationships embodied in (1.21) and (2.7) can assist in tail-index estimation and premium determination *when a consistent pattern exists among the few largest extremes*. Conventional estimation methods generally require large sample sizes for consistency and error analysis (e.g. Embrechts *et al.* (1997), Section 6.4, Drees *et al.*, (2004)).

The usual assumptions require that:

- (i) the number of extremes  $k \rightarrow \infty$ ,
- (ii) compared with sample size  $n$ ,  $k/n \rightarrow 0$ .

Under these conditions  $\frac{X_{(k,n)}}{X_{(k+1,n)}} \xrightarrow{P} 1$  (Embrechts *et al.* (1997), Example 4.1.11) compared with our Equation (2.5).

The methodology is applicable for instance if the mechanism generating the largest extremes possibly differs from that producing ‘ordinary’ large claims.

Beirlant *et al.* (2004) report fitting a three-parameter extension of Generalized Pareto to a well-known large data set which provides a good fit *for all but the largest claims*.

### 3. APPLICATION: THE FEW LARGEST EXTREMES HAVE CLASSIC PATTERN

*‘Let the tail speak for itself’*

(Embrechts *et al.* (1997), Section 6.4.1)

As an example of the risk-averse method, we examine possible ‘premiums’ for the US hurricane data in Table 1.

The data exhibits a classic pattern of extremes. The data set is treated as “*this year’s 10 largest claims*”. Our objective is to set premiums *for next year’s 10 largest claims on the basis of this data alone* (in this year’s dollars).

A validation of the Pareto assumption for natural disaster data of this sort is provided in Hsieh (2004).

TABLE 1

US HURRICANE DATA; THE 10 MOST COSTLY LOSSES (1970-2002) ADJUSTED TO 2002 US DOLLARS.

*Source:* Insurance Information Institute, Feb. 2002

| Rank | Date           | Country                                    | Event            | Insured loss |
|------|----------------|--|------------------|--------------|
| 1    | Aug. 23, 1992  | FL, LA, MS                                 | Hurricane Andrew | 19,875       |
| 2    | Sept. 15, 1989 | Peurto Rico, U.S. et al.                   | Hugo             | 6,087        |
| 3    | Sept. 21, 1998 | PR, U.S. Virgin Is. FL, LA, MS             | George           | 3,201        |
| 4    | Oct. 4 1995    | FL, AL, GA, NC, SC, TN                     | Opal             | 2,479        |
| 5    | Sept. 14, 1999 | NC, NJ, VA, MD, WV, PA, OH 10 other states | Floyd            | 2,117        |
| 6    | Sept. 11, 1992 | Kauai, Oahu, HI.                           | Iniki            | 2,052        |
| 7    | Sept. 12, 1979 | MS, AL, FL, LA, TN, KY, WV, OH, PA, NY     | Frederic         | 1,865        |
| 8    | Sept. 5, 1996  | NC, SC, VA, MD, WV, PA, OH                 | Fran             | 1,835        |
| 9    | Aug. 17, 1983  | TX   | Alicia           | 1,220        |
| 10   | Sept. 15, 1995 | PR, U.S. Virgin Is.                        | Marilyn          | 1,033        |

Hurricanes

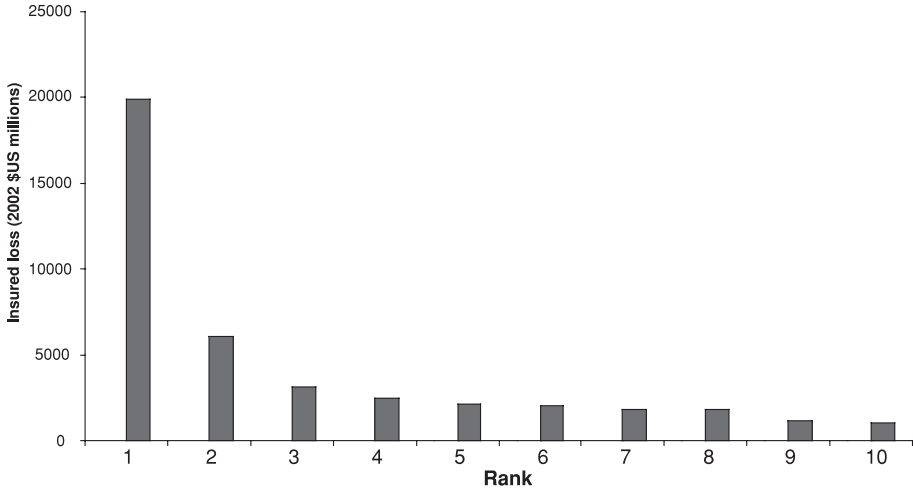


FIGURE 1: The 10 most costly US hurricanes in 2002 dollars.

The consistent pattern of the extremes (Figure 1) could be expected to assist the insurer in choosing a value of  $\rho_{\max} = \beta$  for premium determination.

**3.1. Estimates of the tail index from extreme ratios**

The relations between expected values of ratios of the largest extremes, from equation (1.21) can assist in deciding on a maximum value  $\beta$  of tail-fatness  $\rho$  with which to set premiums. We have:

$$\begin{aligned}
 E[v_n X_{(1)}^* / v_n X_{(2)}^*] &= (1 - \rho)^{-1} \\
 E[v_n X_{(1)}^* / v_n X_{(3)}^*] &= (1 - \rho/2)^{-1} (1 - \rho)^{-1} \\
 E[v_n X_{(1)}^* / v_n X_{(4)}^*] &= (1 - \rho/3)^{-1} (1 - \rho/2)^{-1} (1 - \rho)^{-1} \\
 &\dots \\
 E[v_n X_{(1)}^* / v_n X_{(1+9)}^*] &= 9B(9, 1 - \rho)
 \end{aligned}$$

These nine relationships (all of which involve the important  $X_{(1)}^*$ ) provide estimates for  $\hat{\rho}$  for  $\rho$ ; i.e.

$$(0.6937, 0.7436, 0.7379, 0.7292, 0.7053, 0.6997, 0.6830, 0.7389, 0.7517)$$

(c.f. the L-moment methodology of Pandey *et al.* (2001)).

The mean of these estimates is  $\hat{\rho} = 0.7203$ , their mean absolute deviation (MAD) is 0.0221.

### 3.2. Maximum likelihood based on the $k$ largest extremes

Our assumptions are as follows:

- (i) the  $k$  largest claims arise from a Pareto distribution with fixed tail-index  $\delta$ , ( $\rho = 1/\delta$ ).
- (ii) The total number  $n$  of claims is large; (*we do not assume  $n$  is known*). So that the distribution of the  $k$ th largest claim is approximately Frechet (1.13).
- (iii) Other claims in the set may arise from different distributions.

Drees *et al.* (2004) outline use and discuss properties of conditional maximum likelihood estimation in the sense of Cox and Hinkley (1974, p. 17) predicated on exceedances above a sufficiently high level having a Generalized Pareto distribution.

The conditional likelihood function is of the form:

$$L_1[\mathbf{X} | X_{(n-k)} = x_{(n-k)}] = \prod_{j=0}^{k-1} f(y_{(n-j)}) \tag{3.1}$$

where  $\mathbf{X} = [X_{(n)}, X_{(n-1)} \cdots X_{(n-k+1)}]$ ,  $y_{(n-j)} = x_{(n-j)} - x_{(n-k)}$  and  $f(\cdot)$  is the Pareto density (1.11)

An alternative form of the conditional maximum likelihood which follows from the Markovian character of the order statistics is

$$L_2[X_{(n)}, X_{(n-1)} \cdots X_{(n-k+1)} | X_{(n-k)} = x_{(n-k)}] = \prod_{j=0}^{k-1} f(x_{(n-j)}) / [1 - F(x_{(n-k)})]^k \tag{3.2}$$

where  $F(\cdot)$  is Pareto (1.11) and  $f(\cdot)$  is its density.

For application to the data set in Table 1,  $L_1(\cdot)$  is general in that it is applicable to any distribution from (1.10). The second likelihood  $L_2(\cdot)$  presupposes Pareto for the loss distribution.

Notwithstanding, both assumptions lead to the same Frechet extremes.

### 3.3. Results from Conditional Maximum Likelihood Estimation (CMLE)

Application of CMLE to the data of Table 1 produces:

- (i) using  $L_1$ :  $\hat{\rho} = 0.5553$ ,  $\hat{\lambda}_1 = 3052.3$   
Exceedances above  $x_{(10)} = 1033$  are assumed to follow Pareto (1.11).
- (ii) using  $L_2$ :  $\hat{\rho} = 0.5553$ ,  $\hat{\lambda}_2 = 2019.3$   
(Raw losses are assumed to follow Pareto (1.11)).

The results are coherent in that  $\hat{\lambda}_1 = \hat{\lambda}_2 + x_{(10)}$  is to be anticipated under assumptions; i.e. when comparing Pareto  $f(x)$  and  $f(x - k | X > k)$ , the maximum of  $f(x)$  is at  $x = 0$  and is  $\delta/\lambda$ , whereas that of  $f(x - k | X > k) = (\delta/\lambda)(1 + x/\lambda)^{-\delta-1}/(1 + k/\lambda)^{-\delta}$  at  $x = k$  and is  $\delta/(\lambda + k)$ .

### 3.4. Comparing ratio and MLE results

Figure 2 below depicts graphs of three sets of ratios for  $k = 1$  to  $9$ ;

- (i) Actual ratios  $\frac{X_{(1)}^*}{X_{(1+k)}^*}$ ,
- (ii) Expected values of the ratios when  $\rho = 0.72$ , i.e.  $kB(k, 1 - 0.72)$ ,
- (iii) Expected values of the ratio when  $\rho = 0.5553$ , i.e.  $kB(k, 1 - 0.5553)$ .

The inadequate fit provided by the Conditional ML estimate ( $\rho = 0.5553$ ) is quite apparent.

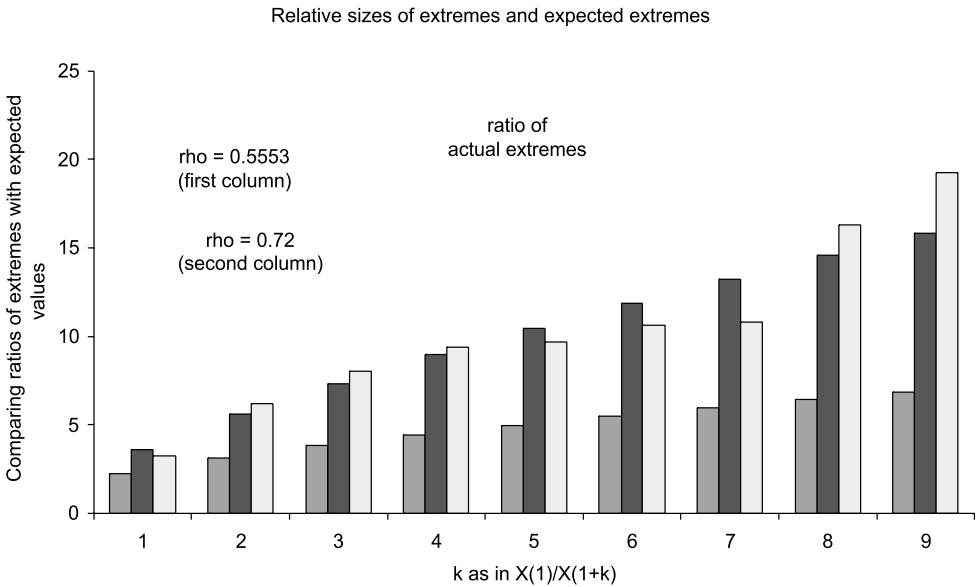


FIGURE 2: Ratios of expected extremes with  $\rho = 0.5553$  (CMLE) and moment estimators  $\rho = 0.72$ , together with ratios of actual extremes.

Smith, (1985), Smith and Naylor, (1987), Nagaraja, (2004) have cautioned against difficulties which may arise in applying MLE to extremes, particularly with  $\rho \in [\frac{1}{2}, 1)$



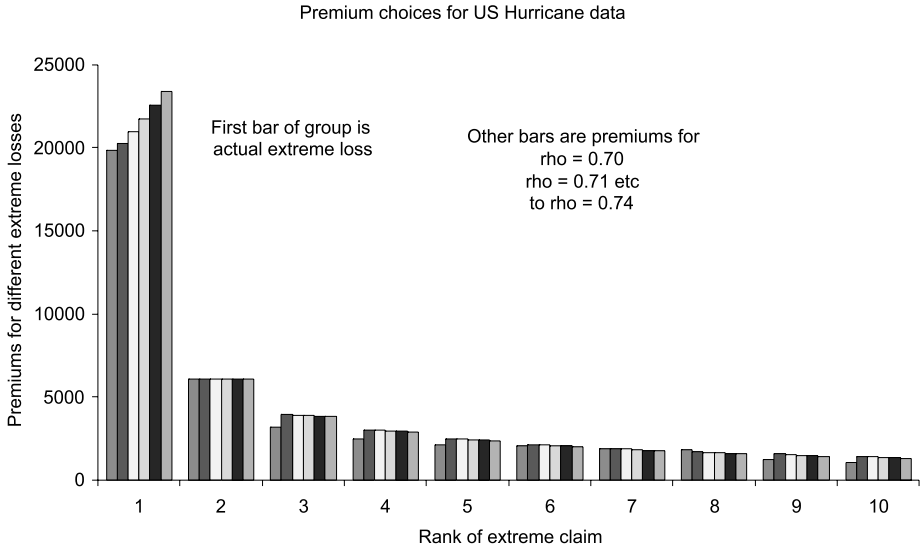


FIGURE 3: Premiums for US Hurricanes set using tail-index values.  $\beta = \rho_{\max} = 0.70(0.01)0.74$  and the actual value  $X_{(2)}^* = 6087$ .

### 3.5. Premiums for extremes

Bearing in mind that the largest claims are the most dangerous, prudence in choice of  $\beta$  should be particularly reflected in the premium for the largest claim (increasing  $\beta$  decreases premiums for lesser extremes in relative terms). The premium for the largest claim is set, the remainder of premiums following from relations (2.5).

The premiums depicted in Figure 3 are calculated using  $\beta = 0.70(0.01)0.74$  with the premium for the second extreme set at its actual value. Total premiums are given in Table 2.

### 3.6. Confidence interval for the premium

Given  $X_{(2)}^*$ , confidence intervals for the premium for  $X_{(1)}^*$  can be set once  $\beta$  is specified, since  $\{X_{(1)}^*/X_{(2)}^*\}^{-\delta}$  is  $B(1,1)$  i.e. uniformly distributed (Appendix, Note 2).

TABLE 2  
TOTAL PREMIUMS FOR US HURRICANES (2002 USD MILLIONS)

| Actual | $\beta = 0.70$ | $\beta = 0.71$ | $\beta = 0.72$ | $\beta = 0.73$ | $\beta = 0.74$ |
|--------|----------------|----------------|----------------|----------------|----------------|
| 41764  | 44635          | 45054          | 45526          | 46054          | 46652          |

Thus while the premium for the largest extreme is  $X_{(2)}^*/(1-\beta)$  its one-sided upper confidence limit is  $X_{(2)}^*/(0.05)^\beta$ . Values are given in Table 3.

The lower limit is the actual largest extreme  $X_{(1)}^*$  (= 19875).

TABLE 3  
PREMIUMS FOR LARGEST CLAIM GIVEN  $X_{(2)}^* = 6087$  WITH 95% CONFIDENCE INTERVAL  
BASED ON DIFFERENT VALUES OF  $\beta = \rho_{\max}$ .

| Tail-index b               | $\beta = 0.70$ | $\beta = 0.71$ | $\beta = 0.72$ | $\beta = 0.73$ | $\beta = 0.74$ |
|----------------------------|----------------|----------------|----------------|----------------|----------------|
| Premium                    | 20290          | 20990          | 21739          | 22544          | 23412          |
| 95% upper confidence limit | 49559          | 51066          | 52619          | 54219          | 55868          |

The confidence intervals provide additional information about the inherent menace of tails as heavy as these.

#### *Remark*

There is considerable discretion available to the insurer in setting premiums. The present methodology focuses insurer attention on the expected structure of potential claims (fully determined in terms of relative sizes, once  $\beta$  is chosen), challenging insurer judgement, experience and intuition.

Input of expert advice and ancillary information can be accommodated. In the hurricane context, the changing nature of land usage, population concentration, building codes and preferences are likely to influence the financial impact of hurricane damage.

Hsieh (2004) points to a 1994 report by Insurance Services Office, Inc. suggesting that population density on the storm-prone southwestern Atlantic coast of the US increased nearly 75% from 1970 to 1990, a much greater increase than the 20% countrywide figure. This trend could be to influence the structure of expected value of future claims over time. Hall and Tajvidi (2000) discuss how such trends in extreme values over time may be estimated.

### SUMMARY AND CONCLUSIONS

A new premium principle with foundations in the theory of risk-aversion applicable to fat-tailed claims is described and its properties examined. Premiums for generalized Pareto claims are expected values of ratios of the largest extremes, so that an unbiased estimator exists for risk-neutral premiums.

The ratio structure of premiums is preserved when imprecision in the tail index is modelled using transformed beta prior distributions.

An example illustrating how insurers can use the presence of a strong pattern in the structure of the largest claims to assist in choice of tail-index and to set premiums is provided.

#### ACKNOWLEDGEMENTS

The author would like to acknowledge helpful comments and recommendations from two unknown referees, as well as help of both Editors. In particular thanks are due to Paul Embrechts who suggested looking for estimators of the premium.

*The research was carried out in part while Roger Gay was on a sabbatical visit to Heriot-Watt School of Mathematical and Computer Sciences.*

#### APPENDIX

##### NOTE 1.

##### *Pricing functions and a pricing rule deriving from constant risk aversion*

The 'no-arbitrage' principle widely used in pricing risk is inappropriate for pricing general insurance premiums. In economic terms, insurers are *price-setters* (see Albrecht, 1992). Premiums can be set using the classic theory of risk-aversion.

Arrow (1971) defined two measures of risk-aversion described below:

*Absolute risk aversion*  $R_A$  where

$$R_A = - \frac{U''(x)}{U'(x)} \quad (\text{A.1})$$

and

*Relative risk aversion*  $R_R$  where

$$R_R = -x \frac{U''(x)}{U'(x)} \quad (\text{A.2})$$

If  $R_A = s$  ( $> 0$ , constant), the utility function

$$U(x) = -\exp(-sx) \quad (\text{A.3})$$

obtains. The *exponential principle* which determines the risk-averse premium  $P$  for claim  $X$  as

$$P = \ln\{M_X(s)\} / s \quad (\text{A.4})$$

(where  $M_X(s)$  is the moment generating function of  $X$ ) follows by using a *version* of the expected utility principle (see for instance Bowers *et. al.* 1986, p. 9); i.e. a version of

$$U(P) = E[U(X)] \quad (\text{A.5})$$

For instance, the minimum acceptable premium  $P$  to an insurer with utility function  $U$  and wealth  $W$  for insurance against random claim  $X$  is given as in (A.4) by using utility function (A.3) and

$$U(W) = E[U(W + P - X)] \quad (\text{A.6})$$

Or it derives from the ‘zero utility principle’,

$$E[U(P - X)] = U(0) \quad (\text{A.7})$$

(see for instance, Rolski *et al.* 1999, p. 91), in conjunction with utility function (A.3).

*Equivalently*, the exponential function  $m(x) = \exp(sx)$  could be regarded as the *pricing function* of an insurer with constant absolute risk-aversion (see Gay, 2004b) and the principle follows directly from a *pricing rule*.

$$m(P) = E[m(X)] \quad (\text{A.8})$$

analogous to (A.5).

The pricing rule and the pricing principle can be obtained by manipulation of the expected utility principle and utility functions *to apply directly to premium determination rather than asset pricing*. Think of the graph of  $U(x) = -\exp(-sx)$  in the third Cartesian quadrant ( $x < 0$ ,  $U < 0$ ). Now transfer the shape back to the first quadrant ( $x > 0$ ,  $U > 0$ ) i.e. measure loss positively and negative utility positively (where  $-U(x)$  is “degree of discomfort”  $m(x)$ , say). Then  $m(x) = \exp(sx)$  is used to price insurance premiums in the face of constant absolute risk aversion.

For *constant relative risk aversion*, the process requires amendment of the definition of  $R_R$  for pricing insurance premiums to

$$R_R = x \frac{U''(x)}{U'(x)} \quad (\text{A.9})$$

resulting in pricing function

$$m(x) = x^{\alpha+1} \quad (\text{A.10})$$

when  $R_R$  is set to constant  $\alpha$ .

Evidently  $\alpha = 0$  provides for risk-neutrality,  $\alpha > 0$  for positive risk-aversion.

However it is derived, *the exponential principle*:

- has its roots in constant absolute risk-aversion,
- is applicable to thin-tailed distributions of exponential order, and
- results in setting premiums set with particular regard to the first two moments of the distribution

The *power principle* derives from constant relative risk aversion, is applicable to fat-tailed (power law) distributions. Premiums are determined by expected values of the first two extreme claims.

NOTE 2.

*Pareto order statistics, Frechet extremes and conditional distributions*

For Pareto (1.12) and any value  $v$  such that  $1 \leq k \leq v < k + 1 \leq n$ , the  $v$ th fractional order statistic  $X_{(k,n)}^{(v)}$  interpolating from  $v = k$  up to  $v < k + 1$  is defined via its density on  $(0, \infty)$  given by

$$f_v(x) = \delta [1 - (x + 1)^{-\delta}]^{v-1} (x + 1)^{-(n-v)\delta} (x + 1)^{-\delta-1} / B(v, n + 1 - v) \tag{B.1}$$

by analogy with the  $k$ th order statistic, where  $B(v, n + 1 - v) = \Gamma(v)\Gamma(n + 1 - v) / \Gamma(n + 1)$ .

Thus the  $v$ th order statistic is the  $k$ th order statistic when  $v = k$  an integer, but interpolates from the  $k$ th up to but excluding the  $(k + 1)$ th order statistic ( $1 \leq k \leq n - 1$ ) as implied by its density (B1) otherwise.

Using (B1) it is easy to show that

$$\begin{aligned} E[1 + X_{(k,n)}^{(v)}] &= B(v, n + 1 - v - \rho) / B(v, n + 1 - v) \\ &= \{\Gamma(v)\Gamma(n + 1 - v - \rho) / \Gamma(n + 1 - \rho)\} \times \Gamma(n + 1) / \Gamma(v) / \Gamma(n + 1 - v) \\ &= \Gamma(n + 1 - v - \rho)\Gamma(n + 1) / \{\Gamma(n + 1 - \rho)\Gamma(n + 1 - v)\} \end{aligned}$$

For instance, when  $v = n - k + 1$ ,

$$E[1 + X_{(n-k+1)}] = \Gamma(k - \rho)\Gamma(n + 1) / \{\Gamma(n + 1 - \rho)\Gamma(k)\} \tag{B.2}$$

$$\begin{aligned} &\sim n^\rho \Gamma(k - \rho) / \Gamma(k) \text{ (see for instance Feller (1968, p. 66))} \tag{B.3} \\ &= E[X_{(1)}^*] \end{aligned}$$

and  $E[1 + X_{(n-1)}] = (1 - \rho) E[1 + X_{(n)}]$

We also have for Pareto

$$\begin{aligned} \frac{1 + E[X_{(n)}]}{1 + E[X_{(n-k,n)}^{(v)}]} &= \{\Gamma(1 - \rho)\Gamma(n + 1) / \Gamma(n + 1 - \rho)\} / \\ &\quad [\Gamma(v + 1 - \rho)\Gamma(n + 1) / \{\Gamma(n + 1 - \rho)\Gamma(v + 1)\}] \\ &= \{\Gamma(1 - \rho)\Gamma(v + 1)\} / \Gamma(v + 1 - \rho) \\ &= vB(1 - \rho, v) \end{aligned}$$

which is independent of  $n$ . In particular, for  $v = 1, k = n - 1, \frac{E[1 + X_{(n)}]}{(1 - \rho)^{-1}} = \frac{E[1 + X_{(n-1)}]}{(1 - \rho)^{-1}}$

Similarly, the  $v$ th Frechet extreme  $X_{(k)}^{*(v)}$  interpolating between  $X_{(k)}^*$  and  $X_{(1+k)}^*$  is defined via its density

$$f_v(x) = \delta x^{-v\delta-1} \exp\{-x^{-\delta}\} / \Gamma(v) \quad (k \geq 1, v \geq 1, x > 0, \delta > 0)$$

from which it follows that

$$E[X_{(k)}^{*(v)}] = \Gamma(v - \rho) / \Gamma(v) \tag{B.4}$$

and that

$$\begin{aligned} \frac{E[v_n X_{(1)}^*]}{E[v_n X_{(k)}^{*(v)}]} &= \Gamma(1 - \rho) \Gamma(v + 1) / \Gamma(v + 1 - \rho) \\ &= vB(1 - \rho, v) \end{aligned} \tag{B.5}$$

The convergence of the expected values of normalised order statistics to the corresponding expected value of extremes is established in Resnick (1987), Chapter 2).

*Conditional expectation*

For Pareto (1.12) consider the conditional expectation:

$$E\left[\frac{1 + X_{(n)}}{1 + X_{(n-k)}} \mid X_{(n-k)}\right] = \int_{x_{(n-k)}}^{\infty} \frac{1 + x_{(n)}}{1 + x_{(n-k)}} f(x_{(n)} \mid x_{(n-k)}) dx_{(n)} \tag{B.6}$$

We have  $f(x_{(n)} \mid x_{(n-k)}) = B^{-1}(k, 1) \times \{F(x_{(n)}) - F(x_{(n-k)})\}^{k-1} f(x_{(n)}) / \{1 - F(x_{(n-k)})\}^k$  (see for instance, Arnold, Balakrishnan and Nagaraja, (1993), p. 24)

Now put  $u = \frac{1 - F(x_{(n)})}{1 - F(x_{(n-k)})} = \frac{(1 + x_{(n)})^{-\delta}}{(1 + x_{(n-k)})^{-\delta}}$  for Pareto (1.12).

The integral in (B6) simplifies to  $\int_0^1 u^{-\rho} (1 - u)^{k-1} du / B(k, 1)$   
 $= B(k, 1 - \rho) / B(k, 1)$   
 $= kB(k, 1 - \rho)$

Note that this is also the value of  $\frac{1 + E[X_{(n)}]}{1 + E[X_{(n-k)}]}$ .

Furthermore since  $E \left[ \frac{1 + X_{(n)}}{1 + X_{(n-k)}} \middle| X_{(n-k)} \right]$  is constant and independent of  $X_{(n-k)}$  it is also *the unconditional expectation*. (The result can also be obtained by direct integration using the joint distribution of  $X_{(n)}, X_{(n-k)}$ .) Thus we have:

$$E \left[ \frac{1 + X_{(n)}}{1 + X_{(n-k)}} \right] = kB(k, 1 - \rho) \tag{B.7}$$

Analogous results hold for fractional Pareto order statistics, and Pareto version (1.11). The result (B7) is independent of  $n$ . Heuristically, this means it remains unchanged if we put

(i)  $1 + X_{(n)} = n^\rho X_{(1)}^*$ , (ii)  $1 + X_{(n-k)} = n^\rho X_{(1+k)}^*$ , their limiting values in distribution. Then

$$E \left[ \frac{X_{(1)}^*}{X_{(1+k)}^*} \right] = kB(k, 1 - \rho) \tag{B.8}$$

Alternatively this follows from the existence of  $\sup E \left[ \frac{1 + X_{(n)}}{1 + X_{(n-k)}} \right]^{1+\varepsilon}$  ( $1 + \varepsilon < 1/\rho$ ), and hence the uniform integrability of  $\left[ \frac{1 + X_{(n)}}{1 + X_{(n-k)}} \right]$ , and convergence of the expectation in (B7) to the right side, using Billingsley (1995), Theorem 25.12 Corollary, p. 338.

*Risk-averse premium for the largest Frechet claim*

For  $X_{(1)}^*$ , application of (1.15) to Frechet density (1.13) leads to

$$P_\beta = \{\Gamma(1 - \rho/\beta)\}^{-\beta} = \{\Gamma(2 - \rho/\beta)\}^{-\beta} (1 - \rho/\beta)^{-\beta}$$

Then, weighting with (2.1) requires

$$P(v, \beta) = \int_0^\beta v\beta^{-v} x^{v-1} (1 - x/\beta)^{-\beta} (\Gamma(2 - x/\beta))^\beta dx$$

Now  $0.8856 < \Gamma(2 - y) \leq 1$  for  $0 \leq y \leq 1$  (the minimum at  $2 - y \approx 1.46163$ , see for instance Wrench, 1968); replacing the gamma function by its unit upper bound we obtain

$$P(v, \beta) \leq \int_0^\beta v\beta^{-v} x^{v-1} (1 - x/\beta)^{-\beta} dx$$

i.e.

$$P(v, \beta) = K \times vB(1 - \beta, v)$$

In the example of Section 3, this bound is applied with  $v = 1$ ,  $K = X_{(2)}^*$ .

## NOTE 3.

*Implausible priors for  $\rho$* 

No inherent technical difficulties arise if instead of family (2.1) tail-index uncertainty is modeled by

$$f_{\rho}(x) = v(x-a)^{v-1}/(\beta-a)^v \quad (a \leq x < \beta, v \geq 1) \quad (\text{C.1})$$

The premium  $P_{\beta}$  of (1.17) is then replaced by

$$P_{a,\beta} = (1-a/\beta)^{-\beta} P_{\beta} \quad (\text{C.2})$$

As  $a \uparrow \beta$ , (greater certainty about  $\rho$  is assumed)  $P_{a,\beta} \gg P_{\beta}$ . However, such a prior is *essentially implausible* in the light for instance of the empirical realities expressed in the Mikosch quote at the start of Section (2.1).

## GLOSSARY OF PREMIUMS

| Notation           | Application; (page of first occurrence)  |
|--------------------|--|
| $P$                | (i) premium in general; (p. 165)<br>(ii) premium applicable to thin-tailed claims distributions; (p. 165)  |
| $P(\alpha)$        | Power principle premium (PPP) determined by constant relative risk aversion $\alpha (\geq 0)$ ; (p. 168)   |
| $P_X(\alpha)$      | PPP for claim $X$ determined by constant relative risk aversion $\alpha$ ; (p. 168)  |
| $P_{\beta}$        | (i) PPP for Pareto claim $(1+X/\lambda)$ for $\rho < \beta = \rho_{\max} = (1+\alpha)^{-1}$ ; (p. 169)<br>(ii) PPP for Pareto claim $(1+X/\lambda)^{\phi}$ ( $\rho \geq 1, \rho\phi < \beta < 1$ ); (p. 169) |
| $P(v,\beta)$       | Uncertainty premium. Expected value of $P_{\beta}$ when tail-index $\rho$ is weighted by $f_{\rho}(x) = vx^{v-1}/\beta^v, (0 \leq x < \beta, v \geq 1)$ ; (p. 173)   |
| $P_{\alpha,\beta}$ | Uncertainty premium. Expected value of $P_{\beta}$ when tail-index $\rho$ is weighted by $f_{\rho}(x) = v(x-\alpha)^{v-1}/(\beta-\alpha)^v, (\alpha \leq x < \beta, v \geq 1)$ ; (p. 186)                    |

## REFERENCES

- ALBRECHT, P. (1992) Premium calculation without arbitrage, *ASTIN Bulletin*, **22**, 247-256.  
 ARNOLD, B., BALAKRISHNAN, N. and NAGARAJA H. (1993) *A first course in order statistics*, J. Wiley.  
 ARROW, K.J. (1971) Theory of risk aversion in *Essays on the theory of risk bearing*, 90-120, Markham, Chicago.  
 BALKEMA, A.A. and DE HAAN, L. (1974) Residual lifetime at great age, *Annals of Probability*, **2**, 792-804.  
 BEIRLANT, J., JOOSSENS, E. and SEGERS, J. (2004) A new model for large claims, *North American Actuarial Journal*, to appear.  
 BEIRLANT, J., TEUGELS, J.L. VYNCKIER, P. (1996) *Practical analysis of extreme values*, Leuven University Press.



- BEIRLANT, J., TEUGELS, J.L., and VYNCKIER, P. (1994) Extremes in non-life insurance, Galambos, J. *et al.* eds, *Extreme value theory and applications*, 489-510, Dordrecht: Kluwer Academic.
- BOWERS, N.L., GERBER, H.U., HICKMAN, J.C., JONES, D.A. and NESBITT, C.J. (1986) *Risk Theory* Society of Actuaries, USA.
- COLES, S. (2001) *An Introduction to Statistical Modelling of Extreme Values*, Springer-Verlag, London.
- COLES, S.G. and POWELL, E.A. (1996) Bayesian methods in extreme value modelling; a review and new developments, *International Statistical Review*, **64**, 119-136.
- COLES, S.G. and TAWN, J.A. (1996) A Bayesian analysis of extreme rainfall data, *Applied Statistics*, **45**, 463-478.
- COX, D.R. and HINKLEY, D.V. (1974) *Theoretical Statistics*, Chapman and Hall, London.
- DANIELSSON, J., DE HAAN, L., PENG, L. and DE VRIESS, C.G. (2001) Using the bootstrap method to choose the sample fraction in tail index estimation, *J. Multivariate Anal.*, **76**, 226-248.
- DAVID, H.A. and NAGARAJA, H.N. (2003) *Order Statistics*, 3<sup>rd</sup> Edition, Wiley-Interscience.
- DE HAAN, L. (1994) Extreme value statistics, Galambos, J. *et al.* eds, *Extreme value theory and applications*, 93-122, Dordrecht: Kluwer Academic.
- DEKKERS, A.L.M. and DE HAAN, L. (1993) Optimal choice of sample fraction for extreme value estimation. *J. Multivariate Anal.*, **47**, 173-195.
- DEKKERS, A.L.M., EINMAHL, J.H.J., and DE HAAN, L. (1989) A moment estimator for the index of an extreme value distribution, *Ann. Statist.*, **17**, 1833-1855.
- DREES, H., FERREIRA, A. and DE HAAN, L. (2004) On maximum likelihood estimation of the extreme value index, *Annals of Applied Probability*, **14**, 1179-1201.
- DREES, H. and KAUFMANN, E. (1998) Selecting the optimal sample fraction in univariate extreme value estimation, *Stochastic Proc. Appl.*, **75**, 149-172.
- EMBRECHTS, P., KLUPPELBERG, C. and MIKOSCH, T. (1997) *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag.
- FELLER, W. (1968) An introduction to probability theory and its applications, Vol. I, 3<sup>rd</sup> edn. J. Wiley.
- FELLER, W. (1971) An introduction to probability theory and its applications, Vol. II, J. Wiley.
- FEUERVERGER, A. and HALL, P. (1999) Estimating the tail exponent by moment departure from a Pareto distribution, *Ann. Statist.*, **27**, 760-781.
- GAY, R. (2004a) Pricing risk when the distributions are fat-tailed. *J. Appl. Prob.* **41A**, Special Volume: *Festschrift for Chris Heyde; Stochastic Methods and Their Applications*, 157-175.
- GAY, R. (2004b) The power principle and tail-fatness uncertainty, Monash University, Dept of Econometrics, Working Paper Series, (WP01/04). Available online at: <http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/>
- GRIMSHAW, D. (1990) A unification of tail-estimators, *Comm. Statist. Theory Methods*, **19**, 4841-4857.
- HALL, P. and TAJVIDI, N. (2000) Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data, *Statistical Science*, **15**, 153-167.
- HILL, B.M. (1975) A simple general approach to inference about the tail of a distribution, *Ann. Statist.*, **3**, 1163-1174.
- HILL, B.M. (1994) Bayesian forecasting of extreme values in an exchangeable sequence, *Journal of Research of the National Institute of Standards and Technology*, **99**, 521-538.
- HSIEH, P.-H. (2001a) Robustness of conditional moments: an application to premium calculation of reinsurance treaties, *Risk Analysis*, **21**, 225-234.
- HSIEH, P.-H. (2001b) On Bayesian predictive moments of the next record value using three-parameter gamma priors, *Communications in Statistics; Theory and Methods*, **30**, 729-728.
- HSIEH, P.-H. (2004) A data-analytic method for forecasting the next record catastrophic loss, *Journal of risk and insurance*, **71**, 309-322.
- HUISMAN, R., KOEDIJK, K., KOOL, C. and PALM, F. Tail index estimation in small samples, *Journal of business and economic statistics*, **19**, 208-16.
- KENDALL, M.G. and STUART, A. (1969) *The advanced theory of statistics*, Vol. 1. Griffin.
- LEADBETTER, M. (1991) On a basis for 'peaks over threshold' modeling, *Statistics and Probability Letters*, **12**, 357-362.
- MIKOSCH, T. (1997) Heavy-tailed modelling in insurance, *Commun. Statist. – Stochastic models*, **13**, 799-815.

- NAGARAJA, H.N. (2004) An introduction to extreme order statistics and actuarial applications, Paper presented at 2004 ERM Symposium April 26, 2004.
- OBERHETTINGER, F.L. (1974) Tables of Mellin Transforms, Springer-Verlag, N.Y.
- PANDEY, M.D., VAN GELDER P.H.A.J.M. and VRIJLING, J.K. (2001) The estimation of extreme quantiles of wind velocity using L-moments in the peaks-over-threshold approach, *Structural Safety*, **23**, 179-192.
- PICKANDS, J. (1975) Statistical inference using extreme order statistics, *Annals of Statistics*, **3**, 119-131.
- PICKANDS, J. III (1994) Bayes quantile estimation and threshold selection in tail index estimation; in Galambos, J. *et al.* eds, *Extreme value theory and applications* (Dordrecht: Kluwer Academic).
- PRATT, J.W. (1964) Risk aversion in the small and in the large, *Econometrica*, **32**, 122-136.
- PURI, M.L. and SEN, P.K. (1971) Non-parametric methods in multivariate analysis, J. Wiley.
- RESNICK, S.I. (1987) *Extreme Values, Regular Variation and Point Processes*, Springer, New York.
- ROLSKI, T., SCHMIDL, H., SCHMIDT, V. and TEUGELS, J. (1999) *Stochastic processes for insurance and finance*, Wiley.
- SMITH, R.L. (2000) Measuring risk with extreme value theory. In *Risk Management: Theory and Practice*, edited by M. Dempster, Cambridge University Press, also as Chapter Two of *Extremes and Integrated Risk Management*, edited by P. Embrechts, Risk Books London, 19-35.
- SMITH, R.L. (1987) Estimating the tails of probability distributions, *Annals of Statistics*, **15**, 1174-1207.
- SMITH, R.L. (1985) Maximum likelihood estimation in a class of non-regular cases, *Biometrika*, **72**, 67-90.
- SMITH, R.L. and GOODMAN, D. (2000) Bayesian Risk Analysis, Chapter 17 of *Extremes and Integrated Risk Management*, P. Embrechts, Ed. Risk Books London, 235-251.
- SMITH, R.L. and NAYLOR, J.C. (1987) A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution, *Applied Statistics*, **36**, 358-369.
- TEUGELS, J.L. and VANROELEN, G. (2003) Box-Cox transformations and heavy-tailed distributions, *J. App. Prob.* (Special Volume) **43a** 'Festschrift for C.C. Heyde, 2004', Applied Probability Trust.
- WRENCH, J.W. Jr. (1968) Concerning two series for the gamma function, *Math. Comput.*, **22**, 616-626.
- YUAN, A. and CLARKE, B.S. (1999) A minimally informative likelihood for decision analysis, *Canad. J. Statist.*, **27**, 649-665.

ROGER GAY

*Dept. of Accounting and Finance*

*Monash University*

*Clayton, Australia, 3168*

*E-mail: roger.gay@buseco.monash.edu.au*