ESTIMATION IN THE PARETO DISTRIBUTION

By Mette Rytgaard

Nordisk Reinsurance Company A/S, Copenhagen, Denmark

Abstract

In the present paper, different estimators of the Pareto parameter α will be proposed and compared to each others.

First traditional estimators of α as the maximum likelihood estimator and the moment estimator will be deduced and their statistical properties will be analyzed. It is shown that the maximum likelihood estimator is biased but it can easily be modified to an minimum-variance unbiased estimator of α. But still the coefficient of variance of this estimator is very large.

For similar portfolios containing same types of risks we will expect the estimated α-values to be at the same level. Therefore, credibility theory is used to obtain an alternative estimator of α which will be more stable and less sensitive to random fluctuations in the observed losses.

Finally, an estimator of the risk premium for an unlimited excess of loss cover will be proposed. It is shown that this estimator is a minimum-variance unbiased estimator of the risk premium. This estimator of the risk premium will be compared to the more traditional methods of calculating the risk premium.

1. Introduction

The Pareto model is very often used as a basis for Excess of Loss quotations as it gives a pretty good description of the random behaviour of large losses — see for example Benktander (1970).

The distribution function can be written as

$$F(x) = 1 - \left(\frac{c}{x}\right)^{\alpha}, \qquad x > c \tag{1.1}$$

with $\alpha > 0$ and $c > 0$. The mean value $E(X)$ exists if $\alpha > 1$ and

$$E(X) = \frac{\alpha}{\alpha - 1}\, c. \tag{1.2}$$

ASTIN BULLETIN, Vol. 20, No. 2

The variance Var (X) exists if $\alpha > 2$ and

$$\text{(1.3)} \qquad \operatorname{Var}(X) = \frac{\alpha}{(\alpha-1)^2(\alpha-2)}\, c^2 .$$

The density function can be written

$$f(x) = \alpha c^{\alpha} x^{-\alpha-1} .$$

The Pareto distribution belongs to the exponential family of distributions as the density function can be written

$$p_\theta(x) = C(\theta) \exp\left(\sum_{i=1}^{k} Q_i(\theta)\, t_i(x) \right) h(x),$$

with

$$\theta = \alpha,\ C(\theta) = \alpha c^{\alpha},\ Q_i(\theta) = -(\alpha+1),\ t_i(x) = \ln x,\ h(x) = 1 .$$

See f. ex. SILVEY (1970).

In the expression of the distribution function two parameters appear. Through the whole paper, we will assume that the lower limit c is known as very often will be the case in practice when the reinsurer receives information about all losses exceding a certain limit which could for instance be the priority of the excess of loss treaty.

If on the other hand c is unknown which is the case if the reinsurer receives only a list of the largest losses and does not know if the list contains all losses exceding 100,000, 120,000 or 150,000. In this case we have to estimate the c.

The maximum likelihood estimator of c is very simple:

$$\hat{c} = \min_i X_i .$$

In other words, we choose the parameter c to be equal to the smallest loss (see f. ex. MUKHOPADHYAY and EKWO (1987) about estimation problems for c).

All results in the following hold only true if c is known. If it is unknown and we have to estimate it, for instance all statements about unbiased estimators of α will not be true.

The Pareto distribution with the distribution funtion at the form (1.1) is the common used definition of the Pareto distribution in Europe. In HOGG and KLUGMANN (1984) we find a different definition of the Pareto distribution function

$$F(x) = 1 - \left(\frac{b}{b+x} \right)^{\alpha}, \qquad x > 0 .$$

This definition of the Pareto distribution is the common used in America. If X is "European" Pareto distributed with parameters (c, α), then $X-c$ is "Amer-

https://doi.org/10.2143/AST.20.2.2005443 Published online by Cambridge University Press

ican" Pareto distributed with parameters (b, α) where $b = c$. The results in the following are applicable to the American Pareto as well by replacing X_i by $X_i + c$ and putting $b = c$ in the formulas.

A third version of the Pareto distribution is known as the "shifted" Pareto (see HOGG and KLUGMAN (1984)). The distribution function is

$$F(x) = 1 - \left(\frac{l+d}{l+x}\right)^{\alpha}, \qquad x > d > 0$$

with unknown parameters α and l. If $X-d$ is American Pareto with parameters (b, α) then X is shifted Pareto with parameters (d, α, l) where $l = b-d$. Of course the shifted Pareto with $l = 0$ is equal to the European Pareto. If $l > 0$, two parameters have to be estimated. This estimation will require numerical techniques and the results in the following can no longer be applied.

2. ESTIMATION OF THE α-PARAMETER

Let $X_1, \ldots, X_n$ be independent identically Pareto distributed random variables. Hence, the maximum likelihood estimator of α is

$$\hat{\alpha} = n \Big/ \sum_{i=1}^{n} \ln \frac{X_i}{c} \tag{2.1}$$

where ln denotes the natural logarithm.

It follows easily that $\ln (X/c)$ will be exponentially distributed with mean value $1/\alpha$ when X is Pareto distributed (c, α). Then

$$T = \sum_{i=1}^{n} \ln \frac{X_i}{c} \tag{2.2}$$

will be Γ-distributed with density function

$$f(t) = \frac{\alpha^n}{(n-1)!} t^{n-1} e^{-\alpha t}, \qquad t > 0.$$

As $\hat{\alpha} = n/T$ we get the following

$$E(\hat{\alpha}) = \frac{n\alpha}{(n-1)} \int_0^{\infty} \frac{\alpha^{n-1}}{(n-2)!} t^{n-2} e^{-\alpha t} \, dt = \frac{n}{n-1} \alpha \tag{2.3}$$

and

$$E(\hat{\alpha}^2) = \frac{n^2 \alpha^2}{(n-1)(n-2)} \int_0^{\infty} \frac{\alpha^{n-2}}{(n-3)!} t^{n-3} e^{-\alpha t} \, dt = \frac{n^2}{(n-1)(n-2)} \alpha^2.$$

Hence, the variance of $\hat{\alpha}$ is

$$\text{(2.4)} \qquad \operatorname{Var}(\hat{\alpha}) = \frac{n^2\alpha^2}{n-1}\left(\frac{1}{n-2} - \frac{1}{n-1}\right) = \frac{n^2}{(n-1)^2(n-2)}\alpha^2 .$$

The maximum likelihood estimator $\hat{\alpha}$ of α is not unbiased — but so is the estimator

$$\text{(2.5)} \qquad \alpha^* = \frac{n-1}{T}$$

with T given by (2.2).

Furthermore,

$$\text{(2.6)} \qquad \operatorname{Var}(\alpha^*) = \frac{1}{n-2}\alpha^2 < \operatorname{Var}(\hat{\alpha}) .$$

Thus, α^* is a better estimator of α than $\hat{\alpha}$ is — and in the following we shall concentrate on this estimator α^*.

As the joint density function of $X_1, \dots, X_n$ can be written

$$\text{(2.7)} \qquad p(x_1, \dots, x_n) = \alpha^n c^{n\alpha}\left(\prod_{i=1}^{n} X_i\right)^{-\alpha-1} = \alpha^n c^{-n} e^{-(\alpha+1)T}$$

with T as in (2.2). Therefore, T will be sufficient for α as the Pareto distribution belongs to the exponential family of distributions which are complete. Then every function $g(T)$ of T is a minimum-variance unbiased estimator of its mean value $Eg(T)$ (see SILVEY (1970), p. 33, or RAO (1973), p. 321). Thus, α^* is a minimum-variance unbiased estimator of $E(\alpha^*) = \alpha$.

As $X_1, \dots, X_n$ are independent, identically Pareto distributed random variables, then $Y_1, \dots, Y_n$ with $Y_i = \ln(X_i/c)$ are independent, identically exponentially distributed random variables with mean value $1/\alpha$ and variance $1/\alpha^2$. It follows from the Central Limit Theorem (RAO (1973), p. 127) that

$$Z_n = \frac{1}{n-1}\sum_{i=1}^{n} Y_i, \qquad n \to \infty ,$$

is asymptotically normally distributed $(1/\alpha,\ 1/(n\alpha^2))$. As $1/(n\alpha^2) \to 0$ for $n \to \infty$ and the function $f(y) = 1/y$ is differentiable with

$$f'(1/\alpha) = -\alpha^2 \neq 0 ,$$

it follows that $\alpha_n^* = f(z_n)$ is asymptotically normally distributed $\left(\alpha, \dfrac{\alpha^2}{n}\right)$

(See RAO (1973), p. 122-124).

Thus, α^* is consistent. That is, α^* converges in probability to its estimated value as n converges to infinity (see RAO (1973), p. 344).

3. THE MOMENT-ESTIMATOR OF α

When X is Pareto distributed (c, α), the mean value $E(X)$ is given by (1.2).

If we solve the equation

$$\bar{X} = \frac{\alpha^0}{\alpha^0 - 1} c \qquad \text{with} \qquad \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i,$$

we get the following estimator of α:

$$\alpha^0 = \frac{\bar{X}}{\bar{X} - c}. \tag{3.1}$$

We can only determine the asymptotical distribution of α^0 when $\alpha > 2$. In this case we get: Let $X_1, \ldots, X_n$ be independent identically Pareto distributed random variables with mean value and variance given by (1.2) and (1.3). Then according to the Central Limit Theorem (RAO (1973), p. 127)

$$Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

will be asymtotically normally distributed with parameters

$$\left(\frac{\alpha}{\alpha - 1} c, \frac{\alpha}{(\alpha - 1)^2 (\alpha - 2)} \frac{c^2}{n} \right), \quad \text{when} \quad \alpha > 2.$$

Consider the function $f(y) = y/(y - c)$. As $f'(y) = -c/(y - c)^2$, we get

$$f'\left(\frac{\alpha}{\alpha - 1} c \right) = -\frac{(\alpha - 1)^2}{c} \neq 0, \quad \text{when} \quad \alpha > 2.$$

The estimator $\alpha_n^0 = f(Y_n)$ is then asymptotically normallly distributed with parameters (RAO (1973), p. 122-124)

$$\left(\alpha, \frac{\alpha(\alpha - 1)^2}{n(\alpha - 2)} \right) = \left(\alpha, \frac{\alpha^2}{n} + \frac{\alpha}{n(\alpha - 2)} \right).$$

α^0 is then asymptotically unbiased for α, and usually, the asymptotical variance is

$$\text{As} \cdot \text{var}(\alpha^0) > \text{Var}(\alpha^*).$$

In some cases we do not know each and every single loss amount but only the total amount of losses and the number of losses exceding a certain lower limit c. Then it is only possible to calculate the α^0-estimator of α.

In the tables below, we compare the mean value and variance of the two estimators α^* and α^0 for different values of $\alpha > 2$.

TABLE 3.1

COMPARISON OF STANDARD DEVIATIONS AND COEFFICIENTS OF VARIANCE FOR $\alpha = 2.1$

n	$\sqrt{\mathrm{Var}(\alpha^*)}$	$\frac{\sqrt{\mathrm{Var}(\alpha^*)}}{E(\alpha^*)}$	$\sqrt{\mathrm{Var}(\alpha^0)}$	$\frac{\sqrt{\mathrm{Var}(\alpha^0)}}{E(\alpha^0)}$	Efficiency of α^0
5	1.212	0.577	(2.254)	(1.073)	(1.859)
10	0.742	0.354	(1.594)	(0.759)	(2.147)
15	0.582	0.277	(1.302)	(0.620)	(2.234)
20	0.495	0.236	(1.127)	(0.537)	(2.277)
25	0.438	0.209	1.008	0.480	2.303
50	0.303	0.144	0.713	0.339	2.353
100	0.212	0.101	0.504	0.240	2.377

TABLE 3.2

COMPARISON OF STANDARD DEVIATIONS AND COEFFICIENTS OF VARIANCE FOR $\alpha = 2.5$

n	$\sqrt{\mathrm{Var}(\alpha^*)}$	$\frac{\sqrt{\mathrm{Var}(\alpha^*)}}{E(\alpha^*)}$	$\sqrt{\mathrm{Var}(\alpha^0)}$	$\frac{\sqrt{\mathrm{Var}(\alpha^0)}}{E(\alpha^0)}$	Efficiency of α^0
5	1.443	0.577	(1.500)	(0.600)	(1.039)
10	0.864	0.354	(1.061)	(0.424)	(1.200)
15	0.683	0.277	(0.866)	(0.346)	(1.249)
20	0.589	0.236	(0.750)	(0.300)	(1.273)
25	0.521	0.209	0.671	0.268	1.287
50	0.361	0.144	0.474	0.190	1.315
100	0.253	0.101	0.335	0.134	1.328

TABLE 3.3

COMPARISON OF STANDARD DEVIATIONS AND COEFFICIENTS OF VARIANCE FOR $\alpha = 3$

n	$\sqrt{\mathrm{Var}(\alpha^*)}$	$\frac{\sqrt{\mathrm{Var}(\alpha^*)}}{E(\alpha^*)}$	$\sqrt{\mathrm{Var}(\alpha^0)}$	$\frac{\sqrt{\mathrm{Var}(\alpha^0)}}{E(\alpha^0)}$	Efficiency of α^0
5	1.732	0.577	(1.549)	(0.516)	(0.894)
10	1.061	0.354	(1.095)	(0.365)	(1.033)
15	0.832	0.277	(0.895)	(0.298)	(1.075)
20	0.707	0.236	(0.775)	(0.258)	(1.095)
25	0.626	0.209	0.693	0.231	1.108
50	0.433	0.144	0.490	0.163	1.132
100	0.303	0.101	0.346	0.115	1.143

The (asymptotically) efficiency is defined as the squareroot of the asymptotic variance of α^0 divided with the variance of α^*, i.e.

$$\sqrt{\frac{\text{As} \cdot \text{var}\,(\alpha^0)}{\text{Var}\,(\alpha^*)}}\,.$$

Of course it makes no real sense to calculate the asymptotic variance and efficienty of α^0 for small values of n.

4. ESTIMATION OF α AS AN AVERAGE OF THE ACTUAL ESTIMATE AND THE "MARKET" VALUE

The minimum-variance unbiased estimator α^* is the best estimator, but even the variance of this estimator is large if only few losses are known or available, which is often the case for practical purposes. Very often the basis for estimation of α is only 5 or 10 losses. But on the other hand, we often have a certain expectation about the right level of the α-value, having experience from other cases in the market. For fire losses, we will usually expect an α near 1.5 — for motor liability an α perhaps near 2.5. Some use of credibility theory would therefore be natural.

Again we assume $X_1, \ldots, X_n$ being independent, identically Pareto distributed random variables. Let

$$S_i = \ln \frac{X_i}{c}\,.$$

Then

$$\bar{S} = \frac{1}{n} \sum_{i=1}^{n} S_i$$

is an estimator of $1/\alpha$. We now regard $\Theta = 1/\alpha$ as a random variable.

Given $\Theta = \theta = 1/\alpha$, S_i is exponentially distributed with a mean value θ. Furthermore,

$$E(S_i|\Theta = \theta) = \mu(\theta) = \theta, \tag{4.1}$$

$$\text{Var}\,(S_i|\Theta = \theta) = \sigma^2(\theta) = \theta^2. \tag{4.2}$$

Let us assume we know by experience that for a given market and a given branch the average level of $1/\alpha$ is $1/\alpha_0$. For each portfolio we will allow a certain variation from this average level. Then

$$E(\mu(\Theta)) = E(\Theta) = \frac{1}{\alpha_0} \qquad \text{(say)}, \tag{4.3}$$

$$\text{Var}\,(\mu(\Theta)) = \text{Var}\,(\Theta) = \frac{k}{\alpha_0^2} \qquad \text{(say)}. \tag{4.4}$$

with k depending of which variation we will allow from the average level of $1/\alpha_0$. This corresponds to a structure function with mean $1/\alpha$ and variance k/α^2.

Alternatively, if we estimate the α_0 by e.g. the maximum likelihood estimator, we can estimate k by calculating the true variance of $1/\alpha_0$ (see Example 4.1 below).

We get

$$E(\sigma^2(\Theta)) = E(\Theta^2) = (E(\Theta))^2 + \mathrm{Var}\,(\Theta) = \frac{k+1}{\alpha_0^2}\,. \tag{4.5}$$

Furthermore (according to BÜHLMANN (1970)),

$$\begin{aligned}\mathrm{Var}\,(\bar{S}) &= \frac{1}{n}\,E(\sigma^2(\Theta)) + \mathrm{Var}\,(\mu(\Theta))\\ &= \frac{1}{n}\,\frac{k+1}{\alpha_0^2} + \frac{k}{\alpha_0^2}\,.\end{aligned}$$

Using the well-known credibility formula (BÜHLMANN (1970)) we can approximate $E(\mu(\Theta)|S_1, \ldots, S_n)$ by $b\bar{S} + (1-b)\,E(\mu(\Theta))$ where

$$b = \frac{\mathrm{Var}\,(\mu(\Theta))}{\mathrm{Var}\,(\bar{S})}\,.$$

We will now define the "credibility" estimator $\bar{\alpha}$ of α by

$$b\bar{S} + (1-b)\,E(\mu(\Theta)) = \frac{1}{\bar{\alpha}}\,, \tag{4.6}$$

and we get

$$b = \frac{k}{k + (k+1)/n} = \frac{kn}{1 + k(n+1)}\,. \tag{4.7}$$

In other words we will calculate $\bar{\alpha}$ as the inverse of

$$\frac{1}{\bar{\alpha}} = b\,\frac{1}{\hat{\alpha}} + (1-b)\,\frac{1}{\alpha_0}\,,$$

remembering the maximum likelihood estimator $\hat{\alpha}$ in (2.1).

Example 4.1

Consider the losses larger than $c = 1$ mio for five fire portfolios:

Portfolio no i	Number of losses	T as in (2.2)
1	17	10.5
2	12	13.5
3	30	19.5
4	5	3.0
5	10	5.5
Total	74	52.0

Then

$$\frac{1}{\alpha_0} = \frac{52.0}{74} = .703 = \frac{1}{1.422} \Rightarrow \alpha_0 = 1.422.$$

As the grand total of the T's is Γ-distributed we can easily calculate the variance of $1/\alpha_0$:

$$\operatorname{Var}(1/\alpha_0) = \frac{1}{74^2} \operatorname{Var}\left(\sum T\right) = \frac{1}{\alpha^2}.$$

By estimating Var (Θ) by the estimate of the variance of $1/\alpha_0$ and using (4.4) we get $k = 1/74 = 0.0135$ giving us the following results:

Portfolio no i	α^*	b	$\bar{S}$	$\bar{\alpha}$
1	1.524	0.1846	0.618	1.455
2	0.815	0.1378	1.125	1.314
3	1.487	0.2855	0.650	1.454
4	1.333	0.0624	0.600	1.436
5	1.636	0.1175	0.550	1.460

Using this method to estimate α, we obtain that extreme losses will not affect too much the estimate of α.

Example 4.2

When pricing excess of loss treaties in practice the situation is not as is the case in example 2.1 above. The only available loss information is the loss information for the actual portfolio (if we ignore the possibility of picking up information gradually). Let us therefore only consider the portfolio no. 1 in example 2.1.

As said in the beginning of this chapter we do often have a certain more or less vague expectation of the level of the α-value and then of the level of $1/\alpha$. For a fire portfolio we might expect to find $1/\alpha$ close to $1/1.5 = 0.67 = 1/\alpha_0$.

Now let $k = 0.01$ and 0.02 corresponding to that we allow a coefficient of variation of Θ of 10% respectively 14.1%.

We will then get the following estimate of α for portfolio no. 1:

			$k = 0.01$		$k = 0.02$	
α	α^*	$\bar{S}$	b	$\bar{\alpha}$	b	$\bar{\alpha}$
1.619	1.524	0.618	0.1441	1.516	0.2500	1.528

In this case the parameter k is more or less "politically" set: If we want the individual estimates of α to be close to the "market value" we shall use a small value of k. But of course the weight factor b depends on the number of losses in the acutal portfolio as well. If it is a very small portfolio the individual estimate of α will always be close to the market value.

5. ESTIMATE OF THE RISK PREMIUM FOR AN UNLIMITED COVER IN EXCESS OF x

In excess of loss quotations the important quantity to estimate is not as much the α-value as it is the expected loss amount for the layer.

Consider the unlimited layer in excess of x. The risk premium $P(x)$ of this layer is (see f. ex. BENKTANDER (1978))

$$P(x) = \hat{n}\frac{x}{\alpha-1} = \hat{n}m(x), \tag{5.1}$$

where $\hat{n}$ is the estimated loss frequency of the layer. Thus, the interesting quantity to estimate is not α but rather $m(x)$, or just $m = m(1)$. We have

$$m = \frac{1}{\alpha-1} = (1/\alpha)\frac{1}{1-1/\alpha} = (1/\alpha)\sum_{k=0}^{\infty}\frac{1}{\alpha^k} = \sum_{k=1}^{\infty}\frac{1}{\alpha^k}. \tag{5.2}$$

Let T be defined as in (2.2). Then T is Γ-distributed (n, α) and

$$E(T^k) = \frac{(n+k-1)^{(k)}}{\alpha^k} \tag{5.3}$$

with $(n+k-1)^{(k)} = (n+k-1)\dots(n+1)n$.

As every function $g(T)$ of the T which is sufficient is a minimum-variance unbiased estimator of its mean value (SILVEY (1970), p. 33, or RAO (1973), p. 321), a minimum-variance unbiased estimator of $1/\alpha^k$ is then

$$\widehat{\frac{1}{\alpha^k}} = \frac{1}{(n+k-1)^{(k)}}T^k = c_k T^k. \tag{5.4}$$

Then a minimum-variance unbiased estimator of m is

$$\hat{m} = \frac{\hat{1}}{\alpha - 1} = \sum_{k=1}^{\infty} c_k T^k, \tag{5.5}$$

since $\hat{m}$ is a function of T—which is sufficient—and $E(\hat{\alpha}) = 1/(\alpha - 1)$ which follows from (5.2). As

$$\hat{m}^2 = \left(\sum_{k=1}^{\infty} c_k T^k \right)^2 = \sum_{k=2}^{\infty} T^k \sum_{j=1}^{k-1} c_j c_{k-j}, \tag{5.6}$$

we get

$$E(\hat{m}^2) = \sum_{k=2}^{\infty} A_k \frac{1}{\alpha^k}, \text{ with } A_k = \sum_{j=1}^{k-1} \frac{c_j c_{k-j}}{c_k} = \sum_{j=1}^{k-1} \frac{(n+k-1)^{(j)}}{(n+j-1)^{(j)}}. \tag{5.7}$$

On the other hand,

$$(E(\hat{m}))^2 = \left(\sum_{k=1}^{\infty} \frac{1}{\alpha^k} \right)^2 = \sum_{k=2}^{\infty} \frac{1}{\alpha^k} \sum_{j=1}^{k-1} 1 = \sum_{k=2}^{\infty} (k-1) \frac{1}{\alpha^k}. \tag{5.8}$$

Thus the variance of $\hat{m}$ is

$$\begin{aligned} \operatorname{Var}(\hat{m}) &= \sum_{k=2}^{\infty} (A_k - (k-1)) \frac{1}{\alpha^k} \\ &= \sum_{k=2}^{\infty} \frac{1}{\alpha^k} \sum_{j=1}^{k-1} \frac{(n+k-1)^{(j)} - (n+j-1)^{(j)}}{(n+j-1)^{(j)}}. \end{aligned} \tag{5.9}$$

It is rather troublesome to calculate the variance of $\hat{m}$ — but we know the CRAMER-RAO lower limit of the variance (SILVEY (1970), p. 35, or RAO (1973), p. 324). In the first place $\hat{m}$ is a unbiased estimator of m. Secondly

$$\frac{d \ln p(x, m)}{dm} = -\left(\frac{n}{\alpha} - T \right) \frac{1}{m^2} = -\left(\frac{n}{\alpha} - T \right) (\alpha - 1)^2,$$

$$\left(\frac{d \ln p(x, m)}{dm} \right)^2 = \left(\frac{n}{\alpha} - T \right)^2 (\alpha - 1)^4 = \left(\frac{n^2}{\alpha^2} - 2 \frac{n}{\alpha} T + T^2 \right) (\alpha - 1)^4.$$

where $p(x, m)$ denotes the simultaneous distribution (2.7)

$$p(x_1, \ldots, x_n) = \alpha^n c^{n\alpha} \left(\prod_{i=1}^{n} X_i \right)^{-\alpha - 1} = \alpha^n c^{-n} e^{-(\alpha+1)T}.$$

Therefore,

$$E\left(\frac{d\ln p(x,m)}{dm}\right)^2 = \left(\frac{n^2}{\alpha^2} - 2\frac{n^2}{\alpha^2} + \frac{n(n+1)}{\alpha^2}\right)(\alpha-1)^4$$

$$= n\frac{(\alpha-1)^4}{\alpha^2} = I_m.$$

The CRAMER-RAO lower limit of the variance is then

$$\text{(5.10)} \qquad \operatorname{Var}(\hat{m}) \geq \frac{1}{n}\frac{\alpha^2}{(\alpha-1)^4}.$$

As

$$\frac{d\ln p(x,m)}{dm} = -\left(\frac{n}{\alpha} - T\right)\frac{\alpha^2}{n(\alpha-1)^2}\frac{(\alpha-1)^4}{\alpha^2}n$$

$$= \left(\frac{T}{n}\frac{\alpha^2}{(\alpha-1)^2} - \frac{\alpha}{(\alpha-1)^2}\right)I_m,$$

no unbiased estimator with variance equal to the CRAMER-RAO lower limit is existing (SILVEY (1970), p. 38).

We can find an upper limit of the variance of $\hat{m}$ too. Using (3.1) we get another estimate of m

$$\text{(5.11)} \qquad m^0 = \frac{1}{\alpha^0 - 1} = \frac{\bar{X}-c}{c}.$$

It is easy to see that m^0 is an unbiased estimator of m. Furthermore,

$$\operatorname{Var}(m^0) = \operatorname{Var}\left(\frac{\bar{X}}{c} - 1\right) = \frac{1}{n^2}\sum_{i=1}^{n}\operatorname{Var}\left(\frac{X}{c}\right)$$

$$= \frac{1}{n}\frac{1}{c^2}\operatorname{Var}(X) = \frac{1}{n}\frac{\alpha}{(\alpha-1)^2(\alpha-2)}.$$

Hence as $\hat{m}$ is a minimum-variance unbiased estimator of m

$$\text{(5.12)} \qquad \frac{1}{n}\frac{\alpha}{(\alpha-1)^2(\alpha-2)} > \operatorname{Var}(\hat{m}) > \frac{1}{n}\frac{\alpha^2}{(\alpha-1)^4}.$$

A third possibility of estimating m is to use the best estimator α^* of α.

We get

$$\text{(5.13)} \qquad m^* = \frac{1}{\alpha^*-1} = \sum_{k=1}^{\infty}\frac{1}{\alpha^{*k}} = \sum_{k=1}^{\infty}\frac{1}{(n-1)^k}T^k.$$

Furthermore,

$$(5.14)\quad E(m^*) \sum_{k=1}^{\infty} \frac{1}{(n-1)^k} E(T^k) = \sum_{k=1}^{\infty} \frac{(n+k-1)^{(k)}}{(n-1)^k} \frac{1}{\alpha^k} > \sum_{k=1}^{\infty} \frac{1}{\alpha^k} = m.$$

This result follows directly from Jensen's inequality since the function $g(x) = 1/(x-1)$ is convex.

As

$$m^2 = \sum_{k=2}^{\infty} T^k \sum_{j=1}^{k-1} \frac{1}{(n-1)^j} \frac{1}{(n-1)^{k-j}} = \sum_{k=2}^{\infty} T^k \frac{k-1}{(n-1)^k},$$

we get

$$E(m^{*2}) = \sum_{k=2}^{\infty} \frac{(k-1)(n+k-1)^{(k)}}{(n-1)^k} \frac{1}{\alpha^k},$$

and furthermore,

$$(E(m^*))^2 = \left(\sum_{k=1}^{\infty} \frac{(n+k-1)^{(k)}}{(n-1)^k} \frac{1}{\alpha^k} \right)^2$$
$$= \sum_{k=2}^{\infty} \frac{1}{\alpha^k} \frac{1}{(n-1)^k} \sum_{j=1}^{k-1} (n+j-1)^{(j)} (n+k-j-1)^{(k-j)}.$$

The variance of the estimator m^* then is

$$(5.15)\quad \operatorname{Var}(m^*) = \sum_{k=2}^{\infty} \frac{1}{\alpha^k} \frac{1}{(n-1)^k} \sum_{j=1}^{k-1} ((n+k-1)^{(k)} - (n+j-1)^{(j)} (n+k-j-1)^{(k-j)})$$
$$= \sum_{k=2}^{\infty} \frac{1}{\alpha^k} \frac{1}{(n-1)^k} \sum_{j=1}^{k-1} (n+k-j-1)^{(k-j)} ((n+k-1)^{(j)} - (n+j-1)^{(j)})$$
$$> \sum_{k=2}^{\infty} \frac{1}{\alpha^k} \sum_{j=1}^{k-1} \frac{1}{(n-1)^j} ((n+k-1)^{(j)} - (n+j-1)^{(j)})$$
$$> \sum_{k=2}^{\infty} \frac{1}{\alpha^k} \sum_{j=1}^{k-1} \frac{(n+k-1)^{(j)} - (n+j-1)^{(j)}}{(n+j-1)^{(j)}} = \operatorname{Var}(\hat{m}).$$

Summing up from (5.14)-(5.15) we have the following

$$(5.14a)\quad E(m^*) > E(\hat{m}) = m,$$

$$(5.15a)\quad \operatorname{Var}(m^*) > \operatorname{Var}(\hat{m}).$$

If we use the minimum-variance unbiased estimator of α in the formula (5.1) we will overestimate the risk premium. The best estimator of the risk premium for an unlimited layer with priority x is

$$(5.16)\quad \hat{P} = \hat{n} \hat{m} x.$$

https://doi.org/10.2143/AST.20.2.2005443 Published online by Cambridge University Press

This formula has been developed in the special case where it is possible to estimate the loss frequency directly. If the excess point x is so high that only very few or no losses are exceeding x, we have to use the α in calculating $\hat{n}$ as described in BENKTANDER (1988). In such cases we will get a different "best" estimator of $P(x)$.

The formula (5.16) is only valid for unlimited layers excess x. We get a different and more complicated formula for limited layers.

Therefore, this method to calculate unlimited layer has only a very limited value for practical purposes. It is more of theoretical interest that it is possible to calculate the "best" estimator of the risk premium. The traditional methods to calculate the risk premiums will be more convenient in practice.

6. EXAMPLE

Year 1	Year 2	Year 3	Year 4	Year 5
2,495,000	1,985,000	3,215,000	(no losses)	19,180,000
2,120,000	1,810,000	2,105,000		1,915,000
2,095,000	1,625,000	1,765,000		1,790,000
1,700,000		1,715,000		1,755,000
1,650,000				

Let us consider a motor portfolio. We are interested in finding the risk premium for an unlimited layer with a priority of 1.5 million.

We have information about all single loss amounts (from ground up) exceding 1.5 million for the last five years (all the losses have been indexed for inflation). Furthermore, we assume that no such problem as IBNR exists.

The estimated loss frequencies for the layer is $\hat{n} = 3.2$.

In the first case, we calculate the risk premium using m^0 in (5.11), that is the traditional way to calculate the risk premium. We get

$$P^0 = 3.2 \times (3{,}057{,}500 - 1{,}500{,}000) = 4{,}984{,}000\,. \tag{6.1}$$

Secondly, we estimate the α^*, setting $c = 1{,}500{,}000$,

$$\alpha^* = \frac{16-1}{6.482} = 2.314\,.$$

The risk premium is calculated using m^* in (5.13). We get

$$P^* = 3.2\,\frac{1}{2.314-1}\,1{,}500{,}000 = 3{,}652{,}968\,. \tag{6.2}$$

In the third case, we will compute the $\hat{m}(x)$ for the layer.

We get from (2.2) and (5.5.)

$$T = 6.482\,,$$

https://doi.org/10.2143/AST.20.2.2005443 Published online by Cambridge University Press

$$\hat{P} = 3.2 \times 0.6430 \times 1{,}500{,}000 = 3{,}086{,}400. \tag{6.3}$$

Finally, we estimate the α using the credibility formulas (4.6) and (4.7).

We get

$$\frac{1}{\bar{\alpha}} = 0.1368 \times \frac{6.482}{16} + (1-0.1368)\frac{1}{2.5} \Rightarrow \bar{\alpha} = 2.496,$$

using $k = 0.01$ and $\alpha_0 = 2.5$. The risk premium is in this case

$$\bar{P} = 3.2 \times 0.6684 \times 1{,}500{,}000 = 3.208{,}320. \tag{6.4}$$

Summing up, we get the following risk premiums

— P^0 based on moment estimator α^0	4,984,000
— P^* based on best estimator α^*	3,652,968
— $\hat{P}$ based on $\hat{m}$	3,086,400
— $\bar{P}$ based on credibility formula	3,208,200

The risk premiums $\hat{P}$ based on $\hat{m}$ are the best of the three first risk premiums in the table, and they are in this case close to the risk premiums based on the credibility formulas.

The traditional method to calculate the risk premium gives large risk premiums in this example. This is because of the very large loss amount in year 5. If this loss has been 9,180,000, and not 19,180,000, we had the following risk premiums

— P^0 based on moment estimator α^0	2,984,000
— P^* based on best estimator α^*	2,979,419
— $\hat{P}$ based on $\hat{m}$	2,570,015
— $\bar{P}$ based on credibility formula	3,125,999

ACKNOWLEDGEMENT

I should like to thank HENRIK RAMLAU-HANSEN for his useful suggestions and comments to this paper.

REFERENCES

BENKTANDER, G. (1970) Schadenverteilung nach Grösse in der Nicht-lebenversicherung. *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker,* **70**.

BENKTANDER, G. (1978) Largest Claims Reinsurance (LCR). A Quick method to calculate LCR-Risk Rates from Excess of Loss Risk-rates. *ASTIN Bulletin* **10**, 54–58.

BENKTANDER, G. (1988) Security Loadings in excess of loss. *Transactions of the 23rd International Congress of Actuaries, Helsinki* **4**, 9–18.

BÜHLMANN, HANS (1970) *Mathematical Methods in Risk Theory*. Springer-Verlag, Berlin-Heidelberg.

HOGG, ROBERT V. and KLUGMAN, STUART A. (1984) *Loss Distributions*. John Wiley & Sons, New York.
MUKHOPADHYAY, N. and EKWO, M.E. (1987) Sequential Estimation Problems for the Scale Parameter of a Pareto Distribution. *Scandinavian Actuarial Journal*, 83-103.
RAO, C.R. (1973) *Linear Statitistical Inference and Its Applications*. John Wiley & Sons, New York.
SILVEY, S.D. (1970) *Statistical Inference*. Penguin Books, Harmondsworth, Middlesex.

METTE RYTGAARD,
Nordisk Reinsurance Company, Grønningen 25, DK-1270 Copenhagen K, Denmark.