ON THE USE OF EXTREME VALUES TO ESTIMATE THE PREMIUM FOR AN EXCESS OF LOSS REINSURANCE

JAN JUNG

Stockholm, Sweden

1. Suppose that the claims experienced by a portfolio could be represented as independent random variables with a distribution function $F(x)$. The net premium per claim for an excess loss cover above an amount of L is then

$$P(L) = \int_L^\infty [1 - F(x)] \, dx. \tag{1}$$

If we have no information about $F(x)$ except a number M of independent claims, we might compute the observed ,,staircase'' distribution function $S_M(x)$ which is for every x an unbiased estimate of $F(x)$, and could thus compute an unbiased estimate for $P(L)$ with the variance [7] [1]

$$\frac{2}{M} \int_L^\infty \int_x^\infty F(x) \, [1 - F(y)] \, dy \, dx.$$

2. In real life we have some qualitative knowledge of $F(x)$ and very limited information about the claims. In his introduction to this subject Beard treats the case where the only information about $F(x)$ consists of the largest claim x_i and the number of claims n_i $(i = 1, 2, \ldots, N)$ observed during N periods (Reference No 2). It is known from the theory of extreme values [3] that for large n_i the distribution of x_i depends mainly on the parameters u_{n_i} and α_{n_i} defined by

$$F(u_n) = 1 - \frac{1}{n} \; ; \; \alpha_n = nF'(u_n). \tag{2}$$

Beard further assumes that $F(x)$ belongs to what is called by

---

[1]) See list of references.

Gumbel ,,the exponential type" of distribution functions, which have an unlimited tail and finite moments. This class is strictly defined by Gnedenko's necessary and sufficient condition [4, p. 68]:

$$\lim_{n \to \infty} n \left[1 - F \left(u_n + \frac{y}{\alpha_n} \right) \right] = e^{-y}. \tag{3}$$

If $F(x)$ satisfies (3), the normed variable

$$y_i = \alpha_{n_1} (x_i - u_{n_1}) \tag{4}$$

has, for $n_i \to \infty$, the asymptotic distribution function

$$\Phi(y) = \exp(-e^{-y}) \tag{5}$$

and the moments

$$E(y_i) \cong C \text{ (Eulers constant} \sim 0.5772) \text{ and Var } (y_i) \cong \frac{\pi^2}{6}. \tag{6}$$

3. If all $n_i \equiv n$ (the case treated by Beard) we may estimate u_n and α_n from the observed x_i-values, either by use of a probability paper, especially constructed from (5) or from the first and second order moments of the x_i-values [1].

Note. There exist distributions with infinite tails and finite moments of all orders which are not of "the exponential type" (Reference No. 4, page 66). If $F(x)$ belongs to these distributions, the above estimation method leads nowhere.

4. If the n_i-values differ but represent ,,equally exposed intervals", we may under certain conditions use a technique of estimation, similar to the above-mentioned. Suppose that the number of claims in the time interval (O, T) constitutes a Poisson process with the intensity λ_t depending only on the time t. Introducing the operational time $S = \int_0^T \lambda_t \, dt$ we get for the largest claim x_S in the operational interval (O, S) the distribution function [5, p. 416]

$$F_S(x) = \sum_{m=0}^{\infty} \frac{S^m}{m!} e^{-S} [F(x)]^m = e^{-S[1-F(x)]} \tag{7}$$

If we define u_S and α_S from (2) it follows from condition (3) that, for large S-values, $y_S = \alpha_S (x_S - u_S)$ has the asymptotic distribution function $\Phi(y)$ given in (5). The assumption ,,equally ex-

posed intervals" implies that we may use for all i a common S-value, tentatively $S^* = \dfrac{1}{N} \displaystyle\sum_{i=1} n_i$, and we may thus estimate u_S and α_S as in the preceding paragraph.

Note. This somewhat astonishing result, that in this case we may use identical estimation methods with equal or unequal n_i-values, depends on the elastic nature of all asymptotic relations. Thus (3) does *not* imply that $\Phi(y)$ is an acceptable approximation for any realizable n-value. The approach to $\Phi(y)$ is slow for $F(x)$ normal but fast for $F(x)$ exponential (Reference No 5). In the special case of $F(x)=1-e-^{ax-b}$, $\Phi(y)$ happens to be the exact distribution function for y_S, as (3) is then an identity.

5. In the general case with different n_i-values from unequally exposed time intervals, we are forced to introduce a more precise specification, and for reasons appearing later we choose the following model. Let $G(z)$ be an at least numerically known distribution function from which we may calculate the function

$$p(k) = \int_k^\infty [1 - G(z)]\, dz \qquad (8)$$

and the constants v_n and β_n defined by

$$G(v_n) = 1 - \frac{1}{n}; \quad \beta_n = nG'(v_n). \qquad (9)$$

We now introduce the specified assumption

$$F(x) \equiv G(\gamma x + \delta), \qquad (10)$$

where γ and δ are unknown parameters. From (1) and (8) we conclude that

$$P(L) = \frac{1}{\gamma} p(\gamma L + \delta). \qquad (11)$$

From (2), (9) and (10) we may express γ and δ in terms of u_n and α_n (and vice versa) by means of the known constants v_n and β_n:

$$u_n = \frac{1}{\gamma}(v_n - \delta) \qquad \alpha_n = \beta_n \gamma \qquad (12a)$$

$$\gamma = \frac{\alpha_n}{\beta_n} \qquad \delta = v_n - \frac{\alpha_n}{\beta_n} u_n. \qquad (12b)$$

If we introduce (12a) in (4) we have

$$y_i = \beta_{n_i} x_i \, \gamma + \beta_{n_i} \delta - \beta_{n_i} v_{n_i}. \tag{13}$$

As the $y_i : s$ have (asymptotically) the mean C and identical variance, the least squares estimates for γ and δ are those values which minimize

$$Q(\gamma, \delta) = \sum_{i=1}^{N} (\beta_{n_i} x_i \, \gamma + \beta_{n_i} \delta - \beta_{n_i} v_{n_i} - C)^2. \tag{14}$$

Thus to obtain the estimates γ^* and δ^* we have only to solve a system of two linear equations. The resulting value of Q could be roughly compared to the χ^2-distribution with $N-2$ degrees of freedom in order to get at least a vague idea about the applicability of the model.

6. Returning to (11) we get from (12b)

$$P(L) = \frac{\beta_n}{\alpha_n} p \left[\frac{\alpha_n}{\beta_n} (L - u_n) + v_n \right] \tag{15}$$

where the only unknown quantities u_n and α_n may be estimated as hinted in paragraph 3.

Let us apply (15) to $G(z) = 1 - e^{-z}$. From (8) we get $p(k) = e^{-k}$ and from (9) $v_n = lg \, n$ and $\beta_n \equiv 1$. Thus we have in this case

$$P(L) = \frac{1}{n \, \alpha_n} e^{-\alpha_n (L - u_n)}. \tag{16}$$

Formula (16) is exactly the result deduced by Beard in [1 and 2] in quite another way. Beard does not explicitly assume $F(x)$ to be an exponential distribution, but makes use of certain relations between the largest values in one sample, which relations characterize this distribution. However, some of the approximations used cancel out and the practical implication of Beards result could therefore be described thus: ,,Calculate $P(L)$ as if $F(x)$ were equal to $1 - e^{-\gamma x - \delta}$ with those constants γ and δ which make u_n and α_n equal to the values u_n^* and α_n^* estimated from the extremes x_i''.

7. The premium $P^*(L)$ determined by this method involves errors of two kinds. *The estimation errors* $\alpha_n^* - \alpha_n$ and $u_n^* - u_n$ depend on

1) the assumption that $F(x)$ belongs to the exponential type,

2) the sample size n and the rapidity of convergence towards $\Phi(y)$ for the distribution of y_i and

3) the estimation method used.

If 1) is correct, the estimation error could be made small by a careful estimation from many observations.

The structure error $\displaystyle\int_L^\infty [1 - F(x) - e^{-\alpha_n(x-u_n)-lg\,n}]\,dx$

depends on our use of the exponential function for the integral appearing in $P(L)$. This error is completely governed by the behaviour of the unknown $F(x)$.

8. The bare fact is that we cannot get much out of little, and two parameters cannot without further assumptions be sufficient to calculate an infinite integral over an unknown function. In most applications we have a fairly good idea about the general shape of the distribution of the claims. It seems obvious to me that we should use this knowledge to choose a basic function $G(z)$ which describes the distribution of the claims better than $1 - e^{-z}$. If we apply (15) to this function $G(z)$, we should get a reasonable estimate of $P(L)$.

Nevertheless, there remains the difficulty of drawing reliable conclusions about $P(L)$ from two parameters involving the value of $F(x)$ and its derivative at one single point. If the layer to be covered is limited by two limits L_1 and L_2 and if our observed x_i-values give an estimate u_n^* between L_1 and L_2 we may get reliable results. In the case where $L_2 = \infty$, however, the estimated premium will depend mainly on our assumptions about $F(x)$ in the extreme right tail, where we have no or few observations.

The scarcity of claims above L is inherent in the problem, and no statistical method can overcome this fundamental difficulty. It is therefore necessary to use all available data and to avoid unnecessary dissipation of the information. If we have no prior information about $F(x)$, the claims above L together with the number of claims below L form a ,,sufficient statistic'' for $P(L)$. On the other hand the rate of convergence towards the asymptotic

distribution $\Phi(y)$ depends on the behaviour of $F(x)$ when x tends to infinity, about which behaviour we can conclude very little from finite samples. The loss of information when passing to the asymptotic distribution may be considerable. It is possible, that for certain branches of insurance experience will justify the method discussed. Until then we should use the utmost discretion and take the trouble to register all relevant claims. If this is possible and if the limit of self-retention L is independent of the claims, the wellknown and approved method of the first paragraph gives an unbiased estimate of $P(L)$.

Even if our registration is incomplete, there are still conservative methods, which can make use of all our observations, e.g. the general theory of order statistics [4]. But this is outside todays subject.

> "Thus conscience does make cowards of us all
> and thus the native hue of resolution
> is sicklied o'er with the pale cast of thought,
> and enterprises of great pith and moment
> with this regard their current turn awry
> and lose the name of action." [6]

9. *Conclusions.* The use of only extreme values in order to determine the excess of loss premium for an infinite layer above L may be dangerous. Pro primo the results depend largely on our knowledge of the unknown extreme right tail of the distribution of the claims. Pro secundo the uncertainty caused by the use of the asymptotical distribution of the extremes is almost uncontrollable.

There are no good solutions (as in most tail problems) but if we want not too bad results, we must try to registrate all relevant information.

During the winter Beard has spent much time on a stimulating correspondence about this question. As most authors like to quote themselves, I'll finish by quoting a passage in one of my letters: "There is a natural law which states that you can never get more out of a mincing machine than what you have put into it. That is: If the reinsurance people want actuarially sound premiums, they must get a decent information about the claim distributions".

REFERENCES

[1] BEARD, R. E.: "Statistical Theory of Extreme Values and an Application to Excess of Loss Reinsurance." (Mimeographed MS 1962).
[2] BEARD, R. E.: "4th ASTIN Colloquium: Subjects for discussion." The ASTIN Bulletin, Vol. II, part III, p. 313, 1963.
[3] GUMBEL, E. J.: "Statistics of Extremes". Columbia Un. Press. N.Y. 1958.
[4] SARHAN, A. E. and GREENBERG, B. G.: "Contributions to Order Statistics". Wiley, London and N.Y. 1962.
[5] FRANCKX, Ed.: "Sur la fonction de distribution du sinistre le plus elevé". The ASTIN Bulletin, Vol. II, part III, 1963.
[6] HAMLET: Monologue on decision theory. Laggard and Blount, 1623.
[7] VAJDA, S.: "Analytical Studies in Stop-Loss Reinsurance". Skand. Akt. 1951, p. 158.