STATISTICAL METHODOLOGY FOR LARGE CLAIMS

J. TIAGO DE OLIVEIRA
Center of Applied Mathematics (I.A.C.) Faculty of Sciences, Lisbon

1. INTRODUCTION

The question of large claims in insurance is, evidently, a very important one, chiefly if we consider it in relation with reinsurance. To a statistician it seems that it can be approached, essentially, in two different ways.

The first one can be the study of overpassing of a large bound, considered to be a critical one. If $N(t)$ is the Poisson process of events (claims) of intensity ν, each claim having amounts Y_i, independent and identically distributed with distribution function $F(x)$, the compound Poisson process

$$M(t) = \sum_{1}^{N(t)} h(Y_i, a)$$

where a denotes the *critical level*, can describe the behaviour of some problems connected with the overpassing of the critical level. For instance, if $h(Y, a) = H(Y - a)$, where $H(x)$ denotes the Heavside jump function ($H(x) = 0$ if $x < 0$, $H(x) = 1$ if $x \geq 0$), $M(t)$ is then the number of claims overpassing a; if $h(Y, a) = Y\,H(Y - a)$, $M(t)$ denotes the total amount of claims exceeding the critical level; if $h(Y, a) = (Y - a)\,H(Y - a)$, $M(t)$ denotes the total claims reinsured for some reinsurance policy, etc.

Taking the year as unit of time, the random variables $M(1)$, $M(2) - M(1), \ldots$ are evidently independent and identically distributed; its distribution function is easy to obtain through the computation of the characteristic function of $M(1)$. For details see Parzen (1964) and the papers on *The ASTIN Bulletin* on compound processes; for the use of distribution functions $F(x)$, it seems that the ones developed recently by Pickands III (1975) can be useful, as they are, in some way, pre-asymptotic forms associated with tails, leading easily to the asymptotic distributions of extremes.

The results of Leadbetter (1972) and Lindgren (1975) can also be useful, the last one introducing the notion of *alarm level*, connected with the critical level.

We will not follow this approach, which seems a very interesting one, letting here only this short note.

The second approach, which we will develop, is based in the asymptotic distributions of largest values, largely exposed in Gumbel's (1958) book and used in some papers of Ramachandran (1974) and (1975), for fire losses. A detailed bibliography will appear in the sequel; but we can recall immediately the important paper by de Finetti (1964) and the useful summary by Beard (1963).

2. The Asymptotic Distribution of the Largest Values and m-th largest Values of a Sample

The theory of largest and smallest values of a sample of independent and identically distributed random variables goes as far away as 1920, in a paper by Dodd. Owing to the difficulty of real use of the distributions, in general even dependent of unknown parameters if their analytic forms are known, we resort to the use of asymptotic distributions for (relatively) large samples. This theory began to be developed in the late twenties by Fisher and Tippett and von Mises and was systematized, in a definitive way, by Gnedenko (1943). Gumbel (1935) developed one of the forms of asymptotic distributions of the m-th largest (or smallest) values.

Later the requisites of independence or identical distribution were weakened; we will not refer to them because they do not seem to be very important to the application in insurance theory. In a general way, we can summarize those results by saying that we have the same asymptotic distributions if the marginal distributions are the same and there is a kind of asymptotic independence or if the random variables are independent and their distributions are related in some way.

If $(x_1, \ldots, x_n)$ is a sample of n independent and identically distributed random variables with distribution function $F(x)$, the distribution function of max $(x_1, \ldots, x_n)$ is evidently

$$F^n(x) = \text{Prob}\,(x_1 \leq x, \ldots, x_n \leq x).$$

For some initial distribution functions $F(x)$, there exist constants

https://doi.org/10.1017/S0515036100011338 Published online by Cambridge University Press

λ_n and $\delta_n(>0)$, not uniquely defined, such that there exists a function $L(x)$ such that

$$F^n(\lambda_n + \delta_n x) \xrightarrow{\omega} L(x).$$

The asymptotic distribution function $L(x)$ may be one of the three forms:

$$\Lambda(x) = e^{-e^{-x}} \qquad \text{Gumbel distribution}$$

$$\begin{aligned} \Phi_\alpha(x) &= 0 \text{ if } x < 0 \\ &= e^{-x^{-\alpha}},\ \alpha > 0 \text{ if } x > 0 \end{aligned} \qquad \text{Fréchet distribution}$$

$$\begin{aligned} \Psi_\alpha(x) &= e^{-(-x)^\alpha},\ \alpha > 0 \text{ if } x < 0 \\ &= 1 \text{ if } x > 0 \end{aligned} \qquad \text{Weibull distribution}$$

As the asymptotic distributions are continuous, the convergence is uniform to that, for large n, $L\ (x - \lambda/\delta)$ can be taken as an approximation of $F^n(x)$. This asymptotic distribution contains the two parameters λ (location) and δ (dispersion) and eventually the shape parameter α. Fig. 1 shows the reduced Gumbel density (without location and dispersion parameters) and Fig. 2 and Fig. 3 show how Fréchet and Weibull densities, without location and dispersion parameters, behave with the change of α.

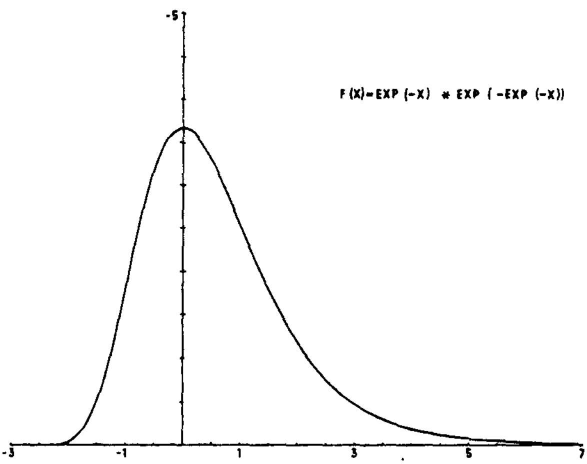

Fig. 1

https://doi.org/10.1017/S0515036100011338 Published online by Cambridge University Press

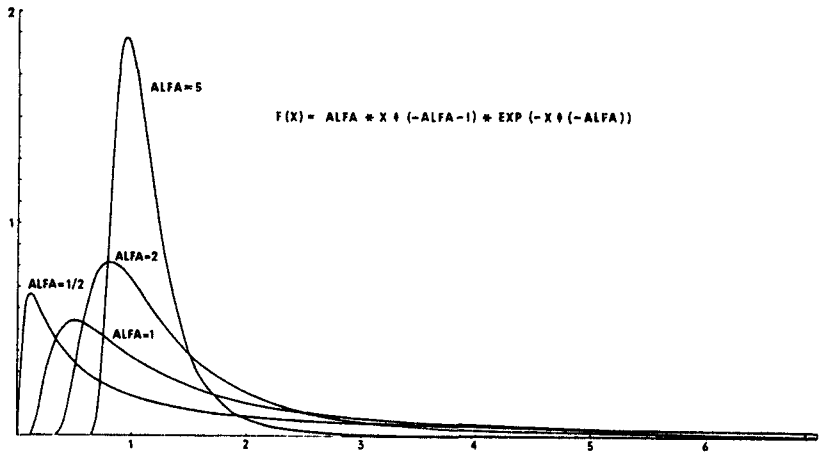

Fig. 2

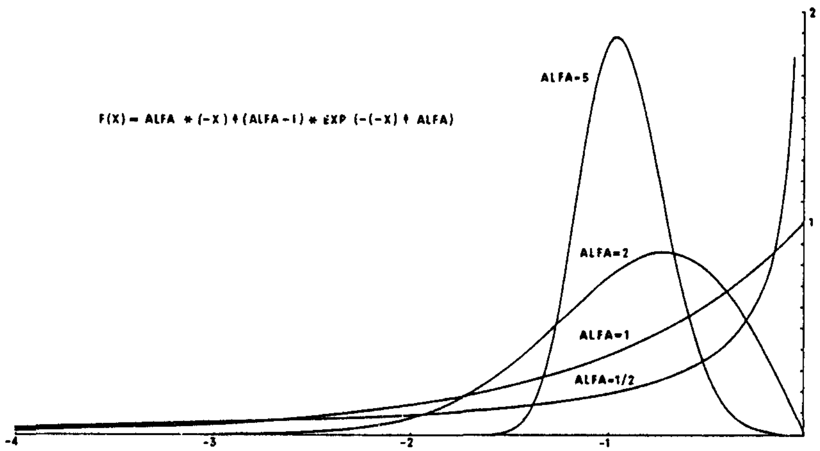

Fig. 3

An important tool to evaluate the use of the distributions of largest values is the behaviour of the force of mortality. The force of mortality for Gumbel distribution is an increasing function as well as for Weibull distributions, and has an U form for Fréchet distributions.

The extension for m-th largest values is immediate. If $x'_1 \leq \ldots \leq x'_n$ denotes the ordered sample, the m-th largest values is the order statistics x'_{n-m+1}; for $m = 1$ we have $x'_n = \max(x_1, \ldots, x_n)$.

The distribution function of x'_{n-m+1} is given by

$$\sum_{k=n-m+1}^{n} \binom{n}{k} F(x)^k (1-F(x))^{n-k} = \sum_{p=0}^{m-1} \binom{n}{p} F(x)^{n-p} (1-F(x))^p$$

and it is easy to show that if $F^n(\lambda_n + \delta_n x) \xrightarrow{\omega} L(x)$, i.e., that the maximum value has the asymptotic distribution $L(x)$ (of one of the three forms Λ, Φ_α, Ψ_α) then the asymptotic distribution of the m-th largest value is given by

$$L(x) \sum_{0}^{m-1} \frac{1}{p!} [-\log L(x)].$$

Note that we have three asymptotic forms and not only the form deriving from $L(x) = \Lambda(x)$, as it is supposed sometimes. For instance, if $F^n(\lambda_n + \delta_n x) \xrightarrow{\omega} \Lambda(x)$, then the reduced asymptotic form for the m-th largest value is

$$e^{-e^{-x}} \sum_{0}^{n-1} \frac{e^{-px}}{p!}.$$

It should be noted that if we take $e^{-x} = m\, e^{-y}$, we obtain the expression given in Gumbel (1958)

$$e^{-m\,e^{-y}} \sum_{0}^{m-1} m^p\, e^{-py} / p!.$$

3. Estimation and Prediction Procedures

As it seems, the two more important problems of statistical decision in actuarial field for the distribution of extremes, are estimation and prediction to be dealt with in this section; it seems that other statistical decision questions are not important in actuarial field.

It must be remarked that we are lacking yet, in many questions, the methodology to obtain the best statistical decision procedures. Its description can be found, in detail, in Tiago de Oliveira (1972) and (1975), not only for Cumbel distribution but also for Fréchet and Weibull ones.

In the sequel we will only describe, as an example, the methodology used when the (supposed) underlying distribution is Gumbel distribution $\Lambda(x)$, sometimes called *the* distribution of extremes.

The maximum likelihood estimators of λ and $\delta(>0)$, the location and dispersion parameters in Gumbel distribution, from a sample $(x_1, \ldots, x_n)$ are given by the equations ($\bar{x}$ denoting the average)

$$\sum_1^n x_i\, e^{-x_i/\hat{\delta}} = (x - \delta) \sum_1^n e^{-x_i/\hat{\delta}}$$

$$\hat{\lambda} = -\hat{\delta} \log \left(\frac{\sum_1^n e^{-x_i/\hat{\delta}}}{n} \right),$$

the first one being solved by iterative methods.

When we take as first approximation $\hat{\delta}_0 = \sqrt{6}\, S/\pi$, where S denotes the standard deviation, the iteration converges numerically, in general, in few steps; the estimate of $\hat{\lambda}$, given by the second equation, is immediate.

As it is well known, the efficiency of those estimators is 1.

Confidence regions can be formed, using the fact that $(\hat{\lambda}, \hat{\delta})$ is asymptotically binormally distributed with mean values λ and δ, asymptotic variances

$$\left[1 + \frac{6}{\pi^2}(1 - \gamma)^2\right] \frac{\delta^2}{n} \text{ and } \frac{6}{\pi^2} \frac{\delta^2}{n}$$

and asymptotic correlation coefficient

$$\rho = \left(1 + \frac{\pi^2}{6(1 - \gamma)^2}\right)^{-1/2}$$

To avoid the iterative procedures we can use linear combinations of order statistics given by the Lieblein-Zellen and Downton statistics, see Tiago de Oliveira (1975), which have, in general, good efficiency, of about 80% or more.

Prediction procedures can be developed from the use of those estimators, for instance, from the use of $\hat{\lambda}$ and $\hat{\delta}$. The prediction of the maximum of N sequent observations is given by $\hat{\lambda} + (\gamma + \log N)\,\hat{\delta}$ and a prediction interval with coefficient $1 - \omega$ (probability $1 - \omega$ that the future observed value will fall in the interval) is given, apart errors of order n^{-1}, by

$$[\hat{\lambda} + (a + \log N)\,\hat{\delta},\ \hat{\lambda} + (b + \log N)\,\hat{\delta}]$$

where a and b are given by the equations

$$e^{-e^{-b}}\, e^{-e^{-a}} = 1 - \omega$$

$$a + e^{-a} = b + e^{-b}.$$

The theory of estimation for the m-th largest value is not yet developed but it can be done in the way of the preceding estimation for the (1st) maximum. For instance, for the 2nd maximum, as the density is

$$e^{-e^{-x}}\, e^{-2x}$$

the estimators are given by

$$\hat{\delta} = 2\left(\bar{x} - \frac{\sum_{1}^{n} x_i\, e^{-x_i/\hat{\delta}}}{\sum_{1}^{n} e^{-x_i/\hat{\delta}}}\right)$$

$$\hat{\lambda} = -\,\hat{\delta}\log\left(\frac{\sum_{1}^{n} e^{-x_i/\hat{\delta}}}{2n}\right)$$

The theory can follow the usual way.

Statistical decision theory of m-th extremes in the case of distribution such that the largest value has a Fréchet or Weibull distribution, not yet developed, will surely have the difficulties found until now for those distributions.

4. Some Hints on Applications

The applications of the theory of the m-th largest values has been developed, in the last years, in a series of papers by Ramachandran, for instance, (1974) and (1975), for the case where is supposed that the asymptotic distribution for the largest value is a Gumbel one and, consequently, the reduced asymptotic are

$$\Lambda_m(x) = e^{-e^{-x}} \sum_{0}^{m-1} \frac{1}{p!}\, e^{-px}.$$

https://doi.org/10.1017/S0515036100011338 Published online by Cambridge University Press

A problem which appears in the applications is the fact that, in many cases, when we take k large samples of sizes $n_1, \ldots, n_k$ the attraction coefficients λ_n and δ_n are, in general, different. Under some condition, we can obtain a general relationship between the λ and δ.

Suppose that exist constants α and $\beta(>0)$ such that $e^{\alpha+\beta x}(1 - F(x)) \to 0$ as $x \to \infty$, which corresponds to the asymptotic conditions supposed in Ramachandran (1974).

In that case we can take $\lambda_n = 1/\beta\,(\log n - \alpha)$, $\delta_n = 1/\beta$ so that we get the following relation

$$\lambda_{n'} = \lambda_n + \delta \log n'/n$$
$$\delta_{n'} = \delta_n.$$

Let us, for simplicity, suppose that $Y_j^{(m)}$ is the m-th largest value from a sample of size n_j under the hypothesis made on $F(x)$. Then the random variables

$$z_j = \frac{Y_j^{(m)} - \lambda_{n_j}}{\delta_{n_j}}$$

have the asymptotic distribution $\Lambda_m(x)$.

From the relation given, taking $\lambda_{n_1} = \lambda$, $\delta_{n_1} = \delta$ we see that we can write

$$z_j = \frac{Y_j^{(m)} - \lambda}{\delta} + \log \frac{n_j}{n_1}$$

so that the $Y_j^{(m)}$ have the asymptotic distribution

$$\Lambda_m\left(\frac{x - \lambda}{\delta} + \log \frac{n_j}{n_1}\right)$$

and the estimation of the parameters λ and δ can be made in the usual way, for instance, using the maximum likelihood method.

Another point which is very important in the study of m-th largest values is the choice between one of the forms of asymptotic distributions.

Until now there is no analytic methodology for this choice. A practical suggestion can be the use of graphical methods (for the technique see Tiago de Oliveira (1972)). We can test, graphically, if the largest values follow one of the distributions and, after, *suppose* that the m-th largest values follow the corresponding asymptotic distribution. For that we can build a probability paper for Gumbel distribution and a deck of probability papers, for various values of α, for Fréchet and Weibull distributions. Then the data can be plotted on those probability papers and one of the forms will be accepted when the plotted points fall, approximately, on a straight line. Recall that when $\alpha \to \infty$ both Fréchet and Weibull distributions, with convenient linear changes of the variable, converge to Gumbel distribution; from a practical point of view it means that, for large α, in both cases, data will fit reasonably well in Gumbel probability paper.

References

Beard, R. E., (1963). Some notes on the statistical theory of extreme values, *The ASTIN Bull.*, Vol. III.

Finetti, Bruno de, (1964). La théorie des plus grands valeurs et son application aux problèmes de l'assurance, *The ASTIN Bull.*, Vol. III.

Gnedenko, B. V., (1943). Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.*, Vol. 44.

Gumbel, E. J., (1935). Les valeurs extrêmes des distributions statistiques, *Ann. Inst. H. Poincaré*, Vol. 4.

Gumbel, E. J., (1958). *Statistics of extremes*, Columbia University Press.

Leadbetter, M. R., (1972). Extreme value theory and stochastic processes, *Int. Conf. on Structural Safety and Reliability*, Pergamon Press.

Lindgren, Georg, (1975). Prediction of catastrophes and high level crossings, presented at Warsow 40th meeting of the *International Statistical Institute*.

Parzen, E., (1964). *Stochastic Processes*, Holden-Day Inc.

Pickands III, James, (1975). Statistical inference using extreme order statistics, *Ann. Stat.*, Vol. 3.

Ramachandran, G., (1974). Extreme value theory and large fire losses. *The ASTIN Bull.*, Vol. VII.

Ramachandran, G., (1975). Extreme order statistics in large samples from exponential type distributions and their applications to fire losses, *Statistical Distributions in Scientific Work*, 2, D. Reidel Publ. Cy.

Tiago de Oliveira, J., (1972). Statistics for Gumbel and Fréchet distributions, *Int. Conf. on Structural Safety and Reliability*, Pergamon Press.

Tiago de Oliveira, J., (1975). Statistical decision for univariate extremes, to be publ. in *Trab. Estad. y Invest. Oper.*