Stochastic Excess-of-Loss Pricing within a Financial Framework

Doris Schirmacher, Ph.D., FCAS
Ernesto Schirmacher, Ph.D., FSA
Neeza Thandi, Ph.D., FCAS, MAAA

Abstract

This paper is aimed at the practicing actuary to introduce the theory of extreme values and a financial framework to price excess-of-loss reinsurance treaties. We introduce the reader to extreme value theory via the classical central limit theorem. Two key results in extreme value theory are presented and illustrated with concrete examples. The discussion then moves on to collective risk models, considerations in modeling tail events, and measures of risk. All these concepts are brought together with the modeling of actual losses. In the last section of the paper all previous elements are brought together with a financial framework for the pricing of a layer of reinsurance. The cash flows between the insurance company and its equity holders are modeled.

Keywords. Collective Risk Model, Experience Rating, Extreme Event Modeling, Extreme Values, IRR, Large Loss and Extreme Event Loading, Monte Carlo Valuation, Reinsurance Excess (Non-Proportional), Risk Pricing and Valuation Models, Simulation, Tail-Value-at-Risk.

1 Introduction

The main goal of this paper is to give the practicing actuary some tools (such as extreme value theory, collective risk models, risk measures, and a cash flow model) for the pricing of excess-of-loss reinsurance treaties. In particular, we have in mind the pricing of high layers of reinsurance where empirical data is scarce and reliance on a mathematical model of the tail of the loss distribution is necessary.

We introduce extreme value theory through the central limit theorem. The central limit theorem tells us that the limiting distribution of the sample mean is a normal distribution. The analogous result from extreme value theory is that the limiting distribution of the sample maximum is an extreme value distribution.[1] There are three distinct families of extreme value distributions: the Fréchet, Weibull, and Gumbel

[1]We are not being very precise but the gist of the result is correct.

distributions. But these three families can be represented as a one parameter family of distributions.

The next result in extreme value theory is the key result for pricing excess-of-loss reinsurance treaties. This result shows that under certain circumstances the limiting distribution of the excess portion of a loss approaches the generalized Pareto distribution (as the threshold increases). This result provides the theoretical underpinnings for using the generalized Pareto distribution in reinsurance excess-of-loss pricing.

At this point we have a good theoretical model. The rest of the paper is a hands-on approach to pricing an excess-of-loss treaty within a financial framework. In Section 3 we introduce the collective risk model together with the underlying data necessary for pricing. We guide the reader with the adjustments necessary to get the data ready for use in modeling the tail of the distribution of losses. We discuss graphical techniques and the estimation of the parameters for both loss and claim count distributions.

In Section 4 we introduce the collective risk model [4, 26] and various measures to quantify risk: standard deviation or variance, value at risk, tail value at risk, expected policyholder deficit, and probability of ruin. We also discuss the concept of rented capital and incorporate that into our cash flow model.

In the last section we bring everything together to determine the price of a reinsurance layer. Our methodology revolves around the concept of the implied equity flows [10]. The equity flows represent the transfer of money between the insurance company and its equity holders.[2] Our cash flow model is comprehensive. It includes all relevant components of cash flow for an insurance company: underwriting operations, investment activity, assets (both income and non-income producing), and taxes. Our model does not take a simplistic view of taxes where most actuaries in the past have calculated them as a straight percentage applied to the results of each calendar year. Instead we compute the taxable income according to the Internal Revenue Service tax code.

In Appendix B we provide a full set of exhibits showing all the components of the cash flows and the implied equity flows.

[2]One should not interpret this sentence literally. In most situations these transfers do not actually occur between the equity holders and the company. Rather they occur virtually between the company's surplus account and the various business units that require capital to guard against unexpected events from their operations.

2 Extreme Value Theory

The investigation of extreme events has a long history. Hydrologists, studying floods, were probably the first ones to develop methods of analysis and prediction. The book *Statistics of Extremes* [13] was the first one devoted exclusively to extreme values and is considered now a classic. The author stresses the importance of graphical methods over tedious computations and has illustrated the book with over 90 graphs. Since its publication in 1958 extreme value theory has grown tremendously and there are many deep and relevant results, but for our purposes we will mention only two of them. Both results tell us about the limiting behavior of certain events.

The first result (the Fischer-Tippett theorem) is the analog of the well known central limit theorem. Here the extreme value distributions play the same fundamental role as the normal distribution does in the central limit theorem. The second result (the Pickands and Balkema & de Haan theorem) shows that events above a high enough threshold behave as if they were sampled from a generalized Pareto distribution. This result is directly applicable to excess-of-loss reinsurance modeling and pricing. Two well known modern references with applications in insurance and finance are the books by Embrechts et. al. [8] and Reiss & Thomas [25].

In this section we also introduce a powerful graphical technique: the quantile-quantile (or QQ-) plot [5, chapter 6]. In many situations we need to compare two distributions. For example, is the empirical distribution of losses compatible with the gamma distribution? A quantile-quantile plot will help us answer that question.

2.1 Distribution of normalized sums

Actuaries are well aware of the central limit theorem [7, 15]; namely, if the random variables $X_1, \ldots, X_n$ form a random sample of size n from a unknown distribution with mean μ and variance σ^2 ($0 < \sigma^2 < \infty$), then the distribution of the statistic $(X_1 + X_2 + \cdots + X_n)/n$ will approximately be a normal distribution with mean μ and variance σ^2/n.

An equivalent way to think about the central limit theorem and to introduce extreme value theory is as follows: consider a sequence of random variables $X_1, X_2, X_3, \ldots$ from an unknown distribution with mean μ and finite variance ($0 < \sigma^2 < \infty$). Let $S_n = \sum_{i=1}^{n} X_i$ (for $n = 1, 2, \ldots$) be the sequence of partial sums. Then the central limit theorem says that if we normalize this sequence of partial sums

$$\frac{S_n - b_n}{a_n} \qquad \text{with } a_n = n, \text{ and } b_n = n\mu, \tag{1}$$

then the limiting distribution is a normal distribution.

2.1.1 Understanding QQ-plots

Before proceeding with extreme value theory let us introduce a powerful graphical technique known as a *quantile-quantile plot* or QQ-plot which will help us assess whether a data set is consistent with a known distribution. For this graphical display we will plot the quantiles of one distribution function against the quantiles of another distribution function.

The quantile function Q is the generalized inverse function[3] of the cumulative distribution function F;

$$Q(p) = F^{\leftarrow}(p) \qquad \text{for } p \in (0,1) \tag{2}$$

where the generalized inverse function $F^{\leftarrow}$ is defined as[4] (see [8, page 130])

$$F^{\leftarrow}(p) = \inf\left\{x \in \mathbb{R} \colon F(x) \geq p\right\}, \quad 0 < p < 1. \tag{3}$$

The quantity $x_p = F^{\leftarrow}(p)$ defines the pth quantile of the distribution function F.

Suppose that our data set consists of the points $x_1, x_2, \ldots, x_n$. Let $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$ denote our data sorted in increasing order.[5] We also use the convention [5, page 11] that $x_{(i)}$ is the $p_i = (i - 0.5)/n$ quantile.

To check if the distribution of our empirical data is consistent with the distribution function F we plot the points $(Q(p_i), x_{(i)})$; that is, the quantiles of F against the quantiles of our data set.

If the empirical distribution is a good approximation of the theoretical distribution, then all the points would lie very close to the line $y = x$; departures form this line give us information on how the empirical distribution differs from the theoretical

[3] We denote generalized inverse function with a left arrow as a superscript ($F^{\leftarrow}$) instead of the more traditional -1 superscript (F^{-1}). We cannot use the traditional definition of inverse function because some of our cumulative distribution functions are not one-to-one mappings.

[4] Ignoring some technicalities, the operator inf selects the smallest member of a set.

[5] Some authors would denote this sequence using double subscripts: $x_{n,n} \leq x_{n-1,n} \leq \cdots \leq x_{1,n}$.

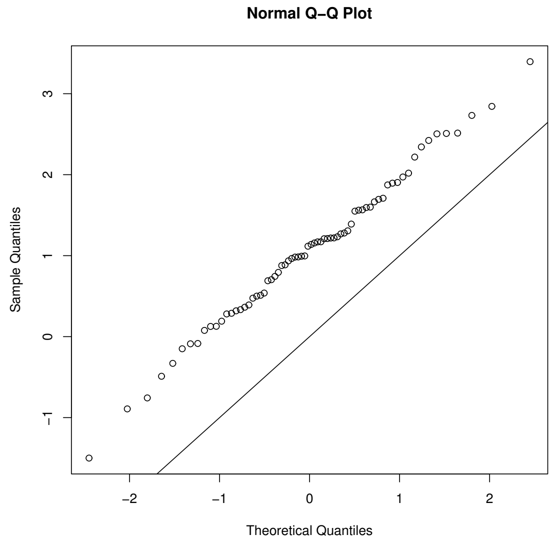

Figure 1: QQ-plot of normal distribution $N(0, 1)$ against $N(1, 1)$. The solid line is $y = x$.

distribution. Figure 1 shows the QQ-plot of a normal distribution $N(1, 1)$ with $\mu = 1$ and $\sigma^2 = 1$ against the standard normal distribution $N(0, 1)$, ($\mu = 0, \sigma^2 = 1$). Note that the points on the graph do not follow the line $y = x$. Rather they follow the line $y = x + 1$. This configuration tells us that we have mis-specified the location parameter. In Figure 2 we have the QQ-plot for a normal distribution with variance equal to 2 against the standard normal distribution. In this case we have mis-specified the variance. This can be readily seen from the graph because the points follow a straight line with slope different from one.

2.1.2 Visualizing the central limit theorem

To visualize the central limit theorem consider a sequence of random numbers from an unknown distribution: $X_1, X_2, X_3, \ldots$. For $n = 1, 2, 3, \ldots$ compute the mean statistic μ_n of the first n terms; that is,

$$\mu_n = \frac{1}{n} \sum_{i=1}^{n} X_i. \tag{4}$$

The central limit theorem tells us that for large enough n the distribution of the mean statistic μ_n is very close to a normal distribution. How can we check that the distribution of μ_n is indeed very close to a normal distribution? Let us draw many

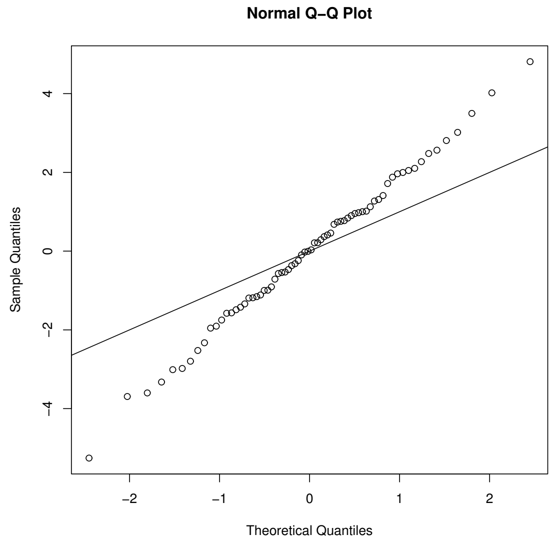

Figure 2: QQ-plot of normal distribution $N(0,1)$ against $N(0,2)$. The solid line is $y = x$.

samples of size n, compute μ_n, and look at the distribution of the mean μ_n.

For an example, take $n = 25$ and calculate 200 means

$$\mu_{25}^{(1)}, \mu_{25}^{(2)}, \mu_{25}^{(3)}, \ldots, \mu_{25}^{(200)} \tag{5}$$

How can we check that the distribution of these sampled means really follows a normal distribution?

We can calculate various numerical summaries: the mean, variance, skewness, kurtosis and others. But relying on numerical summaries alone can be misleading. Rather we should use graphical methods. To assess if our data come from a normal distribution we will show two graphs (see Figure 3). For the first one we will plot the cumulative density function of the sample mean μ_{25} along with the theoretical cumulative density function for the normal distribution. For the second graph we will plot the quantiles of the distribution of μ_{25} against the quantiles of the normal distribution. In this particular example the choice $n = 25$ is large enough so that the central limit theorem applies. Other underlying distributions might require a larger value of n.

Figure 4 shows how the central limit theorem applies to any underlying distribution. For this figure we have chosen three underlying distributions: uniform, gamma, and log-normal. The first row of the display shows the underlying distribution's prob-

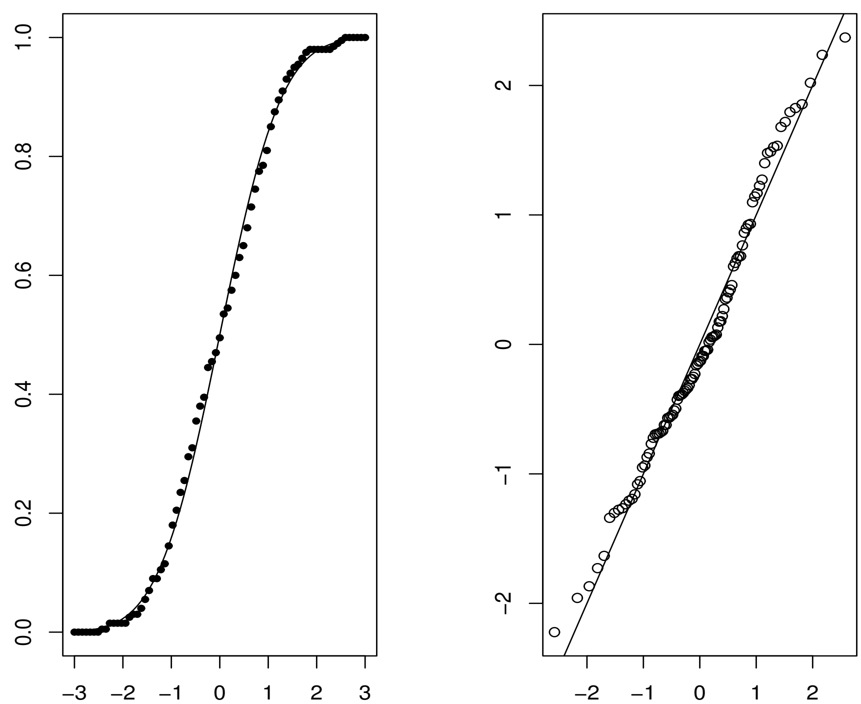

Figure 3: Cumulative density function and quantile-quantile plots for the distribution of the mean μ_{25}.

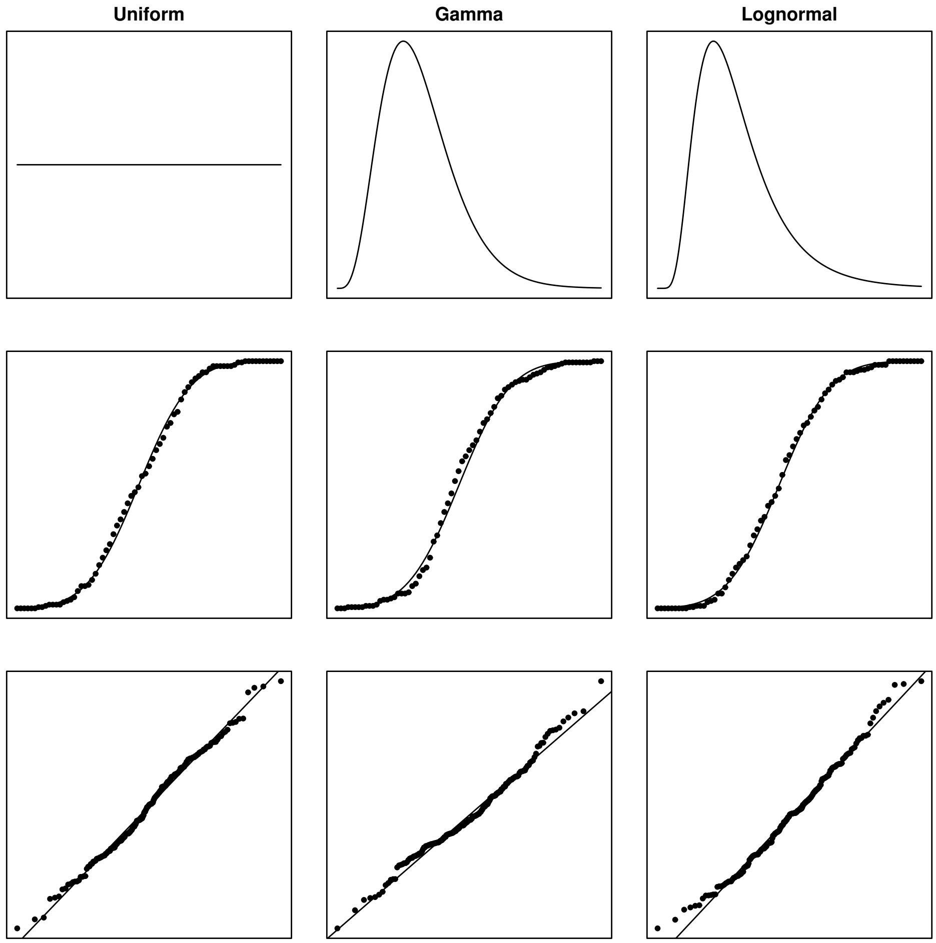

Figure 4: Visualizing the central limit theorem. Top row: underlying density function. Middle row: CDF-plot. (Only 75 of the 200 points were plotted.) Bottom row: QQ-plot.

ability density function. The second row shows the cumulative density function of the mean (dots) along with the cumulative density function for the normal distribution (solid line). Even though the curve does seem to approximate the normal curve fairly close on all three displays of the middle row it is hard for our visual system to distinguish differences from the two curves. The last row of the display shows the QQ-plots. Here it is much easier for us to see that our data (in all three cases) does fall fairly close to the line $y = x$.

Regardless of the underlying distribution (as long it satisfies some mild conditions) the distribution of the mean of a sample follows a normal distribution.

2.1.3 Does the central limit theorem apply to maxima?

While actuaries are interested in the mean severity of claims, they also want to know how large an individual loss might be. Hence, the following question arises naturally: if we replace the mean of a sample with another statistic, say the maximum of the sample, is the limiting distribution (if it exists) still the normal distribution?

As before, consider a sequence of random numbers from an unknown distribution: $X_1, X_2, X_3, \ldots$. For $n = 1, 2, 3, \ldots$ compute the maximum statistic of the first n terms:

$$M_n = \max(X_1, X_2, \ldots, X_n). \tag{6}$$

For n large, does the distribution of M_n converge to a normal distribution?

Using the same experimental procedure as for the mean statistic take $n = 25$ and calculate 200 maxima:

$$M_{25}^{(1)}, M_{25}^{(2)}, M_{25}^{(3)}, \ldots, M_{25}^{(200)}. \tag{7}$$

In Figure 5 we have displayed the cumulative distribution function of the maximum statistic (transformed to have mean zero and unit variance). We have also plotted the standard normal distribution. While we can see that both sets of data do not agree it is hard to know if the departures we see are significant. Our eyes have a hard time distinguishing differences between curved lines. The quantile-quantile plot provides a more powerful graphical technique because we are looking for discrepancies between a straight line and the data. Figure 6 shows clearly that the distribution of the maximum does not follow a normal distribution. If it did the data would fall approximately on a straight line. Rather the points form a concave line. At the upper right-hand corner the data are below the straight line. This implies that the distribution of the maximum is thicker tailed than the normal distribution. The region below the straight line corresponds to points where $Q_t(p) > Q_e(p)$; that is, for a given value of $p \in (0, 1)$ the pth quantile of the theoretical distribution (in our case the normal distribution) is greater than the pth quantile of the distribution. One could argue that our choice of $n = 25$ random numbers is not large enough to show (in our example) that the distribution of the maximum statistic converges to a normal distribution. We performed the same experiment with $n = 100, 1000,$ and 10000 and we still reached the same conclusion: for our example, the distribution of the maximum statistic does not converge to the normal distribution.

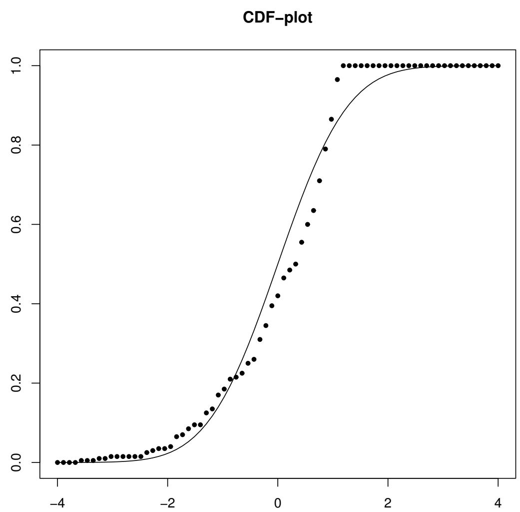

Figure 5: CDF-plot of maximum statistic. The solid line is the standard normal distribution $N(0,1)$.

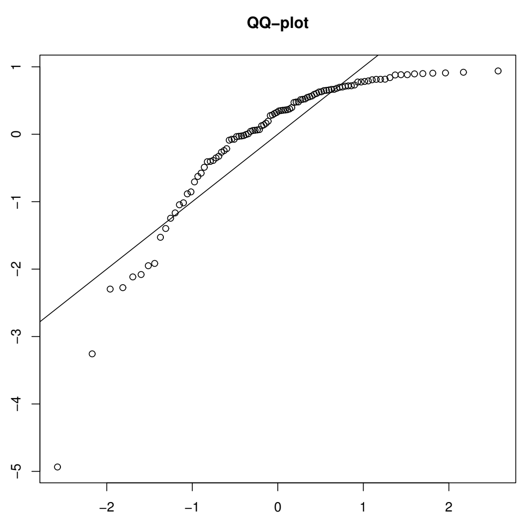

Figure 6: QQ-plot of maximum statistic. The solid line is $y = x$.

2.2 Distribution of normalized maxima

The extreme value theory result analogous to the central limit theorem specifies the form of the limiting distribution for normalized maxima. In place of the partial sums S_n we have the maximum $M_n = \max(X_1, X_2, \ldots, X_n)$.

We know that the distribution of maxima do not follow a normal distribution (see Figure 6). It turns out that the distribution of maxima converges to one of three distributions known as the *extreme value distributions*. The following theorem by Fischer and Tippett [11] explicitly states these three distributions.

Theorem 1 (Fischer-Tippett). *Let X_n be a sequence of independent and identically distributed random variables and let $M_n = \max(X_1, X_2, \ldots, X_n)$ be the maximum of the first n terms. If there exists constants $a_n > 0$ and b_n and some non-degenerate distribution function H such that*[6]

$$\frac{M_n - b_n}{a_n} \xrightarrow{d} H, \tag{8}$$

then H belongs to one of the three standard extreme value distributions:

$$\textit{Fréchet:} \quad \Phi_\alpha(x) = \begin{cases} 0, & x \leq 0, \quad \alpha > 0, \\ \exp(e^{-x^{-\alpha}}), & x > 0, \quad \alpha > 0, \end{cases} \tag{9}$$

$$\textit{Weibull:} \quad \Psi_\alpha(x) = \begin{cases} \exp(-(-x^{\alpha})), & \textit{if } x \leq 0 \textit{ and } \alpha > 0, \\ 1, & \textit{if } x > 0 \textit{ and } \alpha > 0, \end{cases} \tag{10}$$

$$\textit{Gumbel:} \quad \Lambda(x) = \exp(-e^{-x}), \quad \textit{if } x \in \mathbb{R}. \tag{11}$$

A distribution F is said to belong to the *maximum domain of attraction* of the extreme value distribution H if $M_n = \max(X_1, \ldots, X_n)$ satisfies equation (8), where the X_i's are random variables with distribution F.

The Fréchet, Weibull and Gumbel distributions can be written in terms of a one parameter ξ family:

$$H_\xi(x) = \begin{cases} \exp(-(1 + \xi\, x)^{-1/\xi}), & \text{if } \xi \neq 0 \\ \exp(-e^{-x}), & \text{if } \xi = 0 \end{cases} \tag{12}$$

[6] The notation $\xrightarrow{\text{d}}$ refers to convergence in distributions.

Table 1: Maximum likelihood estimates (and their standard errors) for the generalized extreme value distribution $H_\xi([x-\mu]/\sigma)$.

Parameter	Uniform		Gamma		Log-normal	
location (μ)	0.965	(0.003)	9.538	(0.192)	6.521	(0.151)
scale (σ)	0.033	(0.003)	1.707	(0.139)	1.326	(0.111)
shape (ξ)	−0.932	(0.005)	−0.011	(0.073)	0.028	(0.079)

where x is such that $1 + \xi\, x > 0$. This representation is obtained from the Fréchet distribution by setting $\xi = \alpha^{-1}$, from the Weibull distribution by setting $\xi = -\alpha^{-1}$ and by interpreting the Gumbel distribution as the limit case for $\xi = 0$.

We visualize the Fischer-Tippett theorem using the same three underlying distributions (uniform, gamma, log-normal) we used for the central limit theorem. For each underlying distribution we have collected 100 maxima. Each maximum is taken over 25 points chosen at random from the distributions. Table 1 shows the maximum likelihood estimates for each distribution and in Figure 7 we show the corresponding QQ-plots. Note that the shape parameter for the maxima sampled from the uniform distribution is negative. This implies that the uniform distribution is in the maximum domain of attraction of the Weibull distribution. Similarly the shape parameters for the gamma and log-normal are not statistically different from zero. Hence these distributions are in the maximum domain of attraction of the Gumbel distribution.

Distributions that belong to the maximum domain of attraction of the Fréchet distribution include Pareto, Burr, and log-gamma. They are usually categorized as heavy-tailed distributions. Other distributions that actuaries are familiar with include the normal, gamma, exponential, log-normal and Benktander type-I and type-II (see [8, pages 153–7]). These distributions are not as heavy-tailed as the previous examples. They belong to the maximum domain of attraction of the Gumbel distribution. These are medium-tailed distributions. Examples of distributions belonging to the maximum domain of attraction of the Weibull distribution include the beta and uniform distributions. These we shall call thin-tailed distributions.

2.3 Distribution of exceedances

We have seen that the distribution of the maximum does not follow the normal distribution. Rather it follows one of the extreme value distributions: Fréchet, Weibull or Gumbel.

While reinsurance actuaries are interested in the maximum single loss over a given

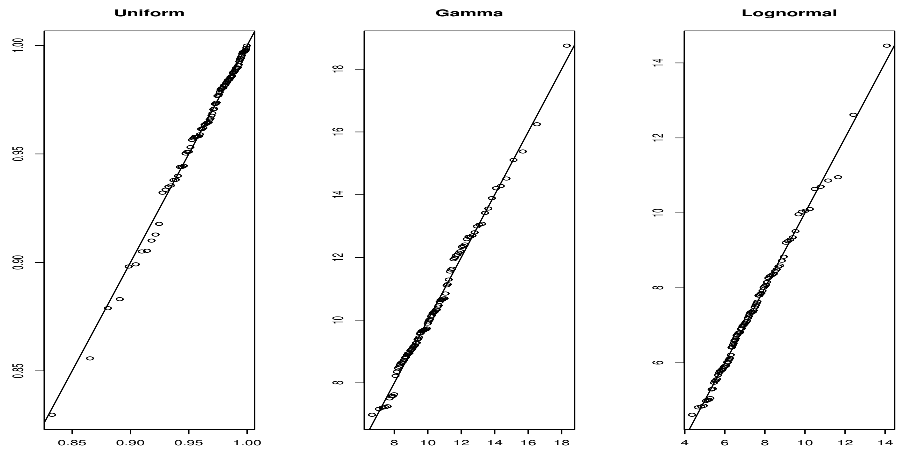

Figure 7: QQ-plots comparing sampled maxima against the fitted generalized extreme value distribution $H_\xi([x-\mu]/\sigma)$.

time period this information is not the area of focus when pricing a contract. The excess of loss reinsurance actuary is concerned about *any* loss that exceeds a predetermined threshold (or attachment point). Suppose that $X_1, X_2, \ldots, X_n$ represent the ground-up losses over a given period. Let u be the predetermined threshold and let

$$Y = [X - u | X \geq u] \tag{13}$$

be the excess of X over u given that the ground-up loss exceeds the threshold. The pricing actuary is interested in the distribution of the exceedances; that is, in the conditional distribution of $Y = X - u$ given that X exceeds the threshold u.

Let F denote the distribution of the random variable X,

$$F(x) = \text{Prob}\,(X < x)\,, \tag{14}$$

and let F_u denote the conditional distribution of the exceedance $Y = X - u$ given that X exceeds the threshold u:[7]

$$F_u(y) = \frac{F(y+u) - F(u)}{1 - F(u)}. \tag{15}$$

Just like the mean statistic converges in distribution to the normal distribution and the maximum statistic converges in distribution to one of the extreme value

[7] The distribution function F_u is also known as the exceedance distribution function, the conditional distribution function, or in reinsurance the excess-of-loss distribution function.

distributions, the exceedances converge in distribution to the generalized Pareto distribution (provided we choose a high enough threshold). The following theorem due to Pickands [24] and Balkema & de Haan [2] shows the result.

Theorem 2 (Pickands, Balkema & de Haan). *For a large class of underlying distribution functions F the conditional excess distribution function $F_u(y)$, for u large, is well approximated by the generalized Pareto distribution $G_{\xi,\sigma}(y)$ where*

$$G_{\xi,\sigma}(y) = \begin{cases} 1 - \left(1 + \frac{\xi}{\sigma} y\right)^{-1/\xi} & \text{if } \xi \neq 0 \\ 1 - \exp(-y/\sigma) & \text{if } \xi = 0 \end{cases} \tag{16}$$

for $y \in [0, (x_F - u)]$ if $\xi \geq 0$ and $y \in [0, -\sigma/\xi]$ if $\xi < 0$.

The point x_F denotes the rightmost point of the distribution function F (which could be finite or infinite).

The class of underlying distribution functions for which the above theorem applies includes most of the standard distribution functions used by actuaries: Pareto, gamma, log-normal, and others (see [16]).

2.3.1 Peaks over threshold method

In order to apply the above theorem we have to choose a threshold. But how do we choose a good threshold? The theorem tells us that if we pick a high enough threshold our data should behave like data that comes from the generalized Pareto distribution. The question is, what characteristics does the generalized Pareto distribution have that we could check against our data? One such characteristic is the mean excess function. The mean excess function for the generalized Pareto distribution $G_{\sigma,\xi}(x)$ is a straight line with positive slope:

$$e(u) = \frac{\sigma + \xi u}{1 - \xi} \tag{17}$$

where $\sigma + \xi u > 0$. The mean excess function[8] for various standard distributions can be found on Table 2.

[8] We use Landau's notation where $o(1)$ stands for an unspecified function of u whose limit is zero as $u \to \infty$.

Table 2: Mean excess functions for some standard distributions.

Distribution	Mean excess function
Pareto	$\frac{\kappa+u}{\alpha-1}, \quad \alpha > 1$
Burr	$\frac{u}{\alpha\tau-1}(1+o(1)), \quad \alpha\tau > 1$
Log-gamma	$\frac{u}{\alpha-1}(1+o(1)), \quad \alpha > 1$
Log-normal	$\frac{\sigma^2 u}{\ln u-\mu}(1+o(1))$
Benktander-type-I	$\frac{u}{\alpha+2\beta\ln u}$
Benktander-type-II	$\frac{u^{1-\beta}}{\alpha}$
Weibull	$\frac{u^{1-\tau}}{c\tau}(1+o(1))$
Exponential	λ^{-1}
Gamma	$\beta^{-1}\left(1+\frac{\alpha-1}{\beta u}+o\left(\frac{1}{u}\right)\right)$
Truncated Normal	$u^{-1}(1+o(1))$

The empirical mean excess function for a sample of data points X_i is given by

$$e_n(u) = \frac{\sum_{i=1}^{n} \max(0, X_i - u)}{\sum_{i=1}^{n} 1_{X_i>u}} \tag{18}$$

where $1_{X>u}$ is the indicator function with value 1 if $X > u$ and zero otherwise.

Figure 8 shows a sample of 2860 general liability losses. Note that most of the losses are very small (say below 500) but there are a few extremely large losses.[9] Figure 9 shows the empirical mean excess plot for these data. Since the mean excess function for the generalized Pareto distribution is a straight line with positive slope, we are looking for the threshold points from which the mean excess plot follows a straight line. There are two regions where the plotted points seem to follow a straight line with positive slope. The first one is from thresholds between 600 and 1000 and the second is between 1000 and 2500. Of course, the second region has very few data points. Based on a threshold $u = 600$ we can fit a generalized Pareto distribution (GPD) (see Figure 10) and check the goodness-of-fit against the data (see Figure 11 for a QQ-plot).

[9] The losses have been scaled so that the largest loss has a value of 10,000.

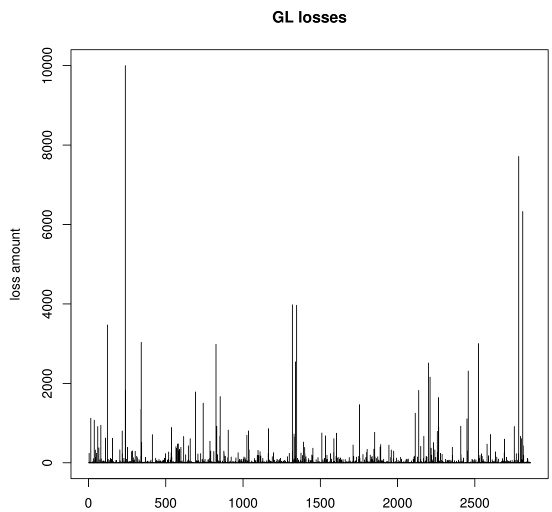

Figure 8: General liability losses. The losses have been normalized so that the maximum loss has a value of 10,000.

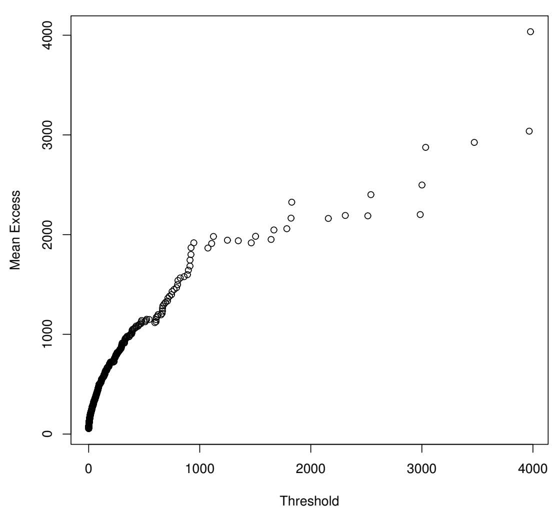

Figure 9: Mean excess plot for general liability losses.

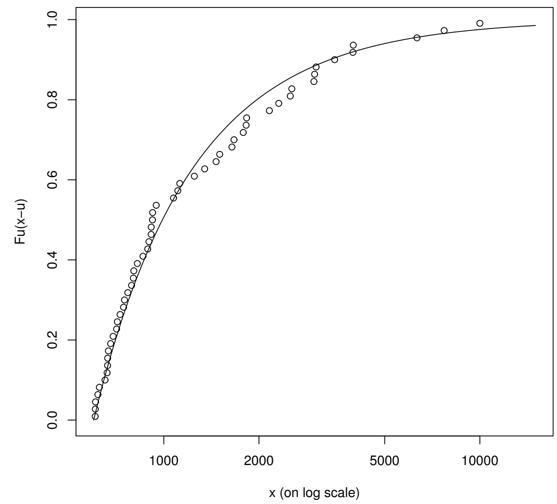

Figure 10: GPD fit (threshold $u = 600$) to the general liability data. The maximum likelihood parameter estimates are $\xi = 0.7871648$ and $\sigma = 423.0858245$

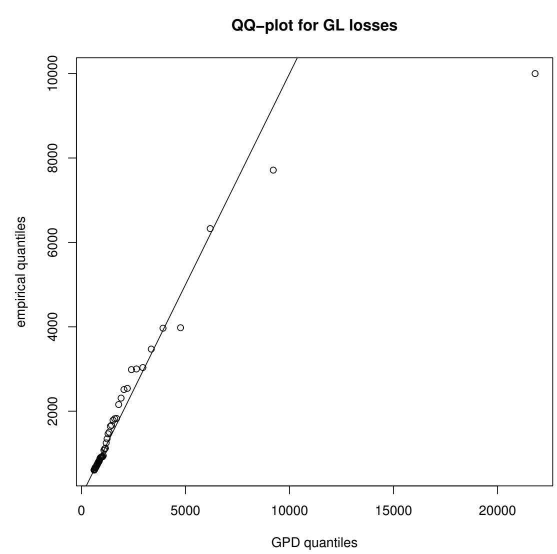

Figure 11: Quantile-quantile plot for general liability losses. The maximum likelihood parameters for the GPD fit are: $u = 600$, $\xi = 0.7871648$ and $\sigma = 423.0858245$.

2.4 Quantile estimation

Estimates of quantiles are important for the actuary and it is easy to calculate them with the generalized Pareto distribution function. To estimate the tail above a threshold u start by re-writing the conditional probability function F_u as follows:

$$F(x) = \text{Prob}\,(X \leq x) = (1 - \text{Prob}\,(X \leq u))\,F_u(x-u) + \text{Prob}\,(X \leq u)\,. \tag{19}$$

From the previous section we know that for large enough threshold u we can approximate $F_u(x-u)$ with the generalized Pareto distribution $G_{\xi,\sigma}(x-u)$. Also using the empirical data we can estimate Prob$(X \leq u)$ with the empirical cumulative density function $F_n(u)$:

$$F_n(u) = \frac{n - N_u}{n} \tag{20}$$

where n is the number of points in the sample and N_u is the number of points in the sample that exceed the threshold u.

If we let $\widehat{F(x)}$ be our approximation to $F(x)$, then for $x \geq u$ we can estimate the tail of the distribution $F(x)$ with

$$\widehat{F(x)} = (1 - F_n(u))\,G_{\xi,\sigma}(x-u) + F_n(u). \tag{21}$$

It is easy to show that $\widehat{F(x)}$ is also a generalized Pareto distribution function with the same ξ parameter but different σ and u parameters. In fact,

$$\widehat{F(x)} = G_{\xi,\tilde{\sigma}}(x - \tilde{u}) \tag{22}$$

where $\tilde{\sigma} = \sigma(1 - F_n(u))^{\xi}$ and $\tilde{u} = u - [\sigma(1 - (1 - F_n(u))^{\xi})/\xi]$. Appendix A shows the derivation of these new parameters.

From equation (21) we can solve for x to obtain our quantile estimator. Let n be the total number of data points and N_u be the number of observations that exceed the threshold u. Then the pth quantile is given by solving the equation

$$p = \left(1 - \frac{n - N_u}{n}\right)\left\{1 - \left(1 + \frac{\xi}{\sigma}(x_p - u)\right)^{-1/\xi}\right\} + \frac{n - N_u}{n} \tag{23}$$

in terms of x_p. This yields the estimator

$$\widehat{x_p} = u + \frac{\sigma}{\xi}\left[\left(\frac{n}{N_u}(1-p)\right)^{-\xi} - 1\right]. \tag{24}$$

2.5 Risk Premium

Once we have estimated the generalized Pareto distribution for our data it is easy to calculate the risk premium (expected losses) in any given layer in excess of our threshold. Let (r, R) (with $R > r > u$) denote the excess-of-loss layer $(R - r)$ xs r. The risk premium in this layer is

$$P = \int_r^R (x - r)\, f_u(x - u)\, dx + (R - r)\,(1 - F_u(R - u)) \tag{25}$$

where $f_u(x - u)$ is the density of the fitted generalized Pareto model. Notice that the price for any layer above the threshold depends only on the excess distribution F_u.

Let

$$G_{\xi,\sigma}(x-u) = \begin{cases} 1 - \left(1 + \frac{\xi}{\sigma}(x-u)\right)^{-1/\xi} & \text{if } \xi \neq 0 \\ 1 - \exp\left(-\frac{x-u}{\sigma}\right) & \text{if } \xi = 0 \end{cases} \tag{26}$$

be the generalized Pareto distribution function including a location parameter u. Using the Pickands and Balkema & de Haan Theorem we can approximate $F_u(x - u)$ with $G_{\xi,\sigma}(x - u)$ and so any questions about a particular layer of reinsurance can be answered by calculating the appropriate moments using the estimated generalized Pareto distribution function.

Calculating the integral (25) to determine the risk premium we have the following explicit formula

$$P = \begin{cases} \frac{\sigma}{\xi}\left[\left(1 + \frac{\xi}{\sigma}(R-u)\right)^{1-1/\xi} - \left(1 + \frac{\xi}{\sigma}(r-u)\right)^{1-1/\xi}\right] & \text{if } \xi \neq 0 \\ \sigma\left[\exp\left(-\frac{r-u}{\sigma}\right) - \exp\left(-\frac{R-u}{\sigma}\right)\right] & \text{if } \xi = 0 \end{cases} \tag{27}$$

3 Collective Risk Models

We shall look at the aggregate losses from a portfolio of risks. Let S_n denote the sum of n individual claim amounts $(X_1, X_2, \ldots, X_n)$, where n is a random number and the claim amounts X_i's are independent and identically distributed random variables.

That is, S_n follows a collective risk model

$$S_n = X_1 + X_2 + \cdots + X_n, \qquad \text{for } n = 0, 1, 2, \ldots \tag{28}$$

with $S_0 = 0$.

In this paper, we focus on experience rating, rather than exposure rating. The next example will be used throughout the paper to illustrate the concepts.

Illustration

Consider pricing an excess-of-loss reinsurance treaty. The treaty covers a small auto liability portfolio with a retention of 3 million, a limit of 12 million, and an annual aggregate deductible of 3 million for accident year 2005. The cedant has provided the following data on large losses

Table 3: Large losses by accident year (I).

Accident Year	1995	1996	1997	1998	1999
Incurred Losses	692,351	902,742	2,314,953	3,183,920	1,168,803
	767,671	2,037,328	702,022	535,590	1,178,212
	1,274,118	1,232,477	1,023,062	742,667	3,722,663
	1,280,334	822,814	3,579,147	922,728	1,830,560
	779,054	684,503	656,957	923,000	509,205
	525,584		1,796,454	831,689	930,300
	1,101,540		589,947	1,622,289	
	980,171		530,295	4,291,141	
	1,268,650		750,693		
	807,076		531,515		
			1,624,021		
			765,879		

Table 4: Large losses by accident year (II).

Accident Year	2000	2001	2002	2003	2004
Incurred Losses	1,172,325	531,500	870,000	1,297,600	851,259
	1,978,249	630,741	592,600	502,776	1,530,050
	512,380	811,327	1,759,111	2,050,000	1,750,000
	1,441,546	989,497	1,856,305	2,350,000	
	925,617	502,603	750,503	9,510,500	
	774,997	566,382			
	1,102,500	1,118,255			
	608,446				
	2,130,454				
	526,483				
	1,600,942				
	1,547,415				

For simplicity, we assume that the loss data are not censored. That is, the given losses are from ground-up losses without capping at the underlying policy limits. This assumption is not crucial and we will discuss briefly in the sections below how one adjusts the modeling procedure when such an assumption is removed. □

3.1 Loss severity distributions

Estimates of the loss severity distribution play an important role in high excess-of-loss reinsurance layers where relevant empirical losses are scarce. It is particularly important that the selected loss distribution fits well the historical large losses and less relevant in explaining the small losses. In practice, when looking at the historical claims one usually ignores the small losses and analyzes only those losses that exceed a threshold.

Before historical losses can be used in any rating procedure they have to be projected to their ultimate values. This is often done by applying loss development factors. Recall that the loss development factors obtained from the usual accident-year triangle analysis contain two parts: development for known claims and development for unreported cases. For loss severity distribution fitting purpose, we need the development factors for known claims.

In addition, because of the well-recognized differences in development between large losses and small losses, we recommend using loss development factors on known claims derived from large claims only. Depending on the treatment for loss adjustment expenses in the reinsurance treaty one must determine whether to include such expenses in the loss data. In this paper, we shall assume that expenses are included in the losses and from this point forward we shall refer to the sum of the two as losses. In addition to the projection to their ultimate values, losses should also be trended to reflect the changes between the experience period and the coverage period. For example loss severity trends may include monetary inflation, increases in jury awards, and increases in medical expenses.

Illustration

In our auto liability example, assume for simplicity that we have a constant inflation of 3% (per annum) and the following loss development factors for known claims:

Table 5: LDF for known claims by accident year.

Year	LDF for known claims
1995	1.001
1996	1.002
1997	1.003
1998	1.009
1999	1.024
2000	1.044
2001	1.050
2002	1.081
2003	1.108
2004	1.172

Losses should be best divided into paid and outstanding and be adjusted/inflated accordingly. Our constant inflation assumption simplifies this process. Depending on the data, certain losses might be closed and should not be developed further. In our simplified illustration, we shall assume that all losses are subject to further development. Then, for example, the first claim in the amount of 692,351 in accident year 1995 would be developed and inflated to its ultimate value of

$$692,351 \cdot 1.001 \cdot 1.03^{(2005-1995)} = 931,392 \qquad (29)$$

In summary, we have the following indexed claim history:

Table 6: Indexed historical large losses.

Accident Year	1995	1996	1997	1998	1999
Indexed Incurred Losses	931,392	1,180,229	2,940,118	3,950,126	1,428,916
	1,032,717	2,663,567	891,607	664,479	1,440,418
	1,714,020	1,611,319	1,299,345	921,389	4,551,127
	1,722,382	1,075,733	4,545,715	1,144,781	2,237,944
	1,048,030	894,907	834,372	1,145,119	622,527
	707,047		2,281,596	1,031,834	1,137,335
	1,481,858		749,265	2,012,690	
	1,318,585		673,504	5,323,798	
	1,706,664		953,422		
	1,085,727		675,053		
			2,062,597		
			972,709		

Accident Year	2000	2001	2002	2003	2004
Indexed Incurred Losses	1,418,819	628,304	1,027,807	1,525,982	1,027,644
	2,394,198	745,620	700,090	591,266	1,847,084
	620,114	959,097	2,078,191	2,410,806	2,112,608
	1,744,647	1,169,718	2,193,015	2,763,607	
	1,120,238	594,144	886,635	11,184,378	
	937,949	669,540			
	1,334,313	1,321,927			
	736,378				
	2,578,405				
	637,182				
	1,937,558				
	1,872,776				

□

If the losses are censored at the underlying policy limit, without knowing exactly how large the ground-up losses are, then one conservative adjustment is to assign such losses at the appropriate policy limits for the current underwriting standards. For example, suppose that a risk had a policy limit of one million and it generated a loss that was capped at the policy limit. Furthermore, assume that the current underwriting standards would give this risk a policy limit of 1.5 million. Then the *as-if* loss for this risk would be the full policy limit of 1.5 million. On the other hand, if one knows the exact size of the ground-up loss, then one should index the ground-up loss as above and limit it at the appropriate policy limit if necessary.

After the historical losses have been adjusted to an *as-if* basis but before we start the model fitting, it is important to explore the data further to gain better understanding. One way to do so is to plot the empirical mean excess function (18). An upward trend in the mean excess plot suggests a heavy tailed behavior, a horizontal line would be exponentially distributed, and thin tailed distribution usually gives a downward trended mean excess plot (see Table 2).

Illustration

In our auto liability data, we have the following mean excess plot:

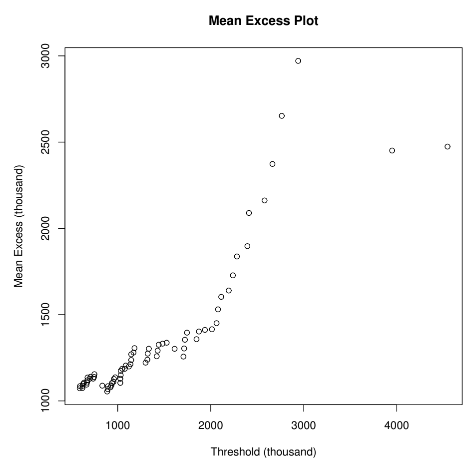

Figure 12: Mean excess loss plot.

The plot shows an upward trend, which suggests that the tail is heavier than an exponentially distributed function. The points above a threshold of 2,000,000 seem to follow a straight line (ignoring the last couple of points which are the average of very few observations). This suggests that a generalized Pareto fit with a threshold of 2 million should provide a good fit. □

Commonly used loss severity distributions in reinsurance pricing include Pareto, log-normal, log-gamma, exponential, gamma, transformed beta, and others. Pareto distributions are particularly popular. Actuaries have recently been applying extreme value theory in estimating the tails of the loss severity distributions [8, 19, 16, 20]. It is particularly useful in pricing high excess of loss layers. The theory suggests that the excess losses above a high threshold are asymptotically distributed according to a generalized Pareto distribution. The loss severity distributions commonly used in reinsurance pricing belong to the class of functions to which the Pickands and Balkema & de Haan theorem 2 applies; showing that the excess loss above a high threshold can be well approximated by the generalized Pareto distribution. This theorem provides the theoretical underpinnings for the popularity of the Pareto distribution in the reinsurance industry when pricing high excess-of-loss layers.

There are various methods of estimating the parameters of the loss severity distributions: method of moments, percentile matching, maximum likelihood, and least squares are among them. The method of moments and percentile matching are easy to implement and convenient but lack of the desirable optimality properties of maximum likelihood and least squares estimators. Maximum likelihood in essence seeks to find the parameters that give the maximum probability to the observed data. Maximum likelihood estimators are asymptotically unbiased and have minimum variance. Unfortunately, it can be heavily biased for small samples. The least squares method seeks to find the parameters estimates that produce the minimum distance between the observed data and the fitted data. The least squares method can be applied more generally than maximum likelihood. However, it is not readily applicable to censored data and is generally considered to have less desirable optimality properties than maximum likelihood. In our model, we will estimate our parameters using maximum likelihood.

Recall that maximum likelihood method selects the parameters θ's which maximize the likelihood function:

$$L(\theta) = \prod_{i=1}^{n} f(x_i|\theta) \tag{30}$$

or equivalently the log-likelihood function:

$$l(\theta) = ln(L(\theta)) = \sum_{i=1}^{n} ln(f(x_i|\theta)) \tag{31}$$

where $f(x_i|\theta)$ is the probability density function evaluated at x_i given θ.

Illustration

We continue with the auto liability example. To fit a generalized Pareto distribution to the indexed excess loss data, first recall that generalized Pareto distribution probability density function is of the form

$$f_{\xi,\sigma}(x) = \frac{1}{\sigma}\left(1 + \frac{\xi}{\sigma}x\right)^{-(\frac{1}{\xi}+1)} \tag{32}$$

and the log-likelihood function is

$$l(\xi, \sigma) = n\left(\frac{1}{\xi} + 1\right)\ln\sigma - \left(\frac{1}{\xi} + 1\right)\sum_{i=1}^{n}\ln\left(1 + \frac{\xi}{\sigma}x_i\right). \tag{33}$$

With the selected threshold of $2,000,000$, we plug in the excess loss (indexed loss - threshold) values and obtain maximum likelihood estimators of $\xi = 0.66784$ and $\sigma = 591,059.8$.

The graph below shows the cumulative density functions for the generalized Pareto as well as the adjusted empirical distributions (adjusted for inflation and loss development).

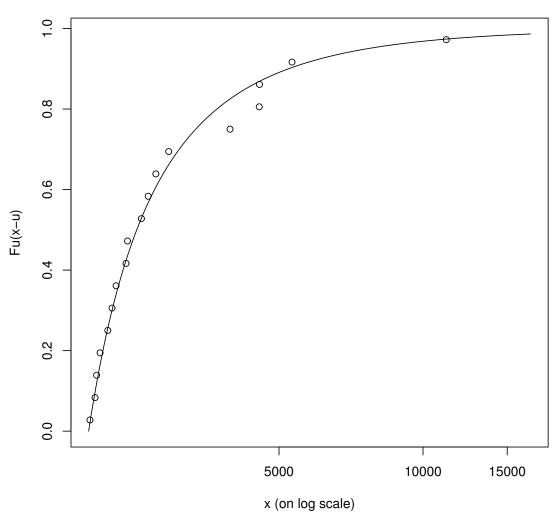

Figure 13: Generalized Pareto cumulative density function.

□

In dealing with censored data, one adjusts the probability function by assigning a mass density at the censor point c (see [17]):

$$\tilde{F}(x) = \begin{cases} F(x), & \text{if } x < c, \\ 1, & \text{if } x \geq c, \end{cases} \tag{34}$$

3.2 Claim frequency distribution

Similar to the loss severity, the claim frequency also needs to be trended and developed to ultimate. One needs to estimate for the not yet reported claims and the possible trends. To estimate the not yet reported claims is easier: one develops the claim number triangle to ultimate. To estimate the possible trends, on the other hand, is certainly not an easy task. Court decisions may influence the liability frequencies; an amendment in the governing law can change the reporting of the WC claims. Both legal and social factors need to be considered when identifying the trends. One further adjustment to the historical frequency is to reflect the historical portfolio sizes. This can be done by comparing the historical exposure sizes to the treaty year exposure size.

Illustration

Continuing with our auto liability example, suppose that we have the following exposure information and claim number development factors:

Table 7: Historical exposure size and claim frequency LDF by accident year.

Year	No. of Exposures	Claim Freq Dev. Factor
1995	21,157,000	1.007
1996	19,739,000	1.007
1997	19,448,000	1.007
1998	19,696,000	1.022
1999	19,406,000	1.030
2000	19,543,000	1.037
2001	19,379,000	1.073
2002	21,186,000	1.197
2003	24,425,000	1.467
2004	27,990,000	2.379
2005	28,000,000	

Then, for example, there is one claim reported in 1996 above the chosen threshold of 2,000,000, with the adjustment for unreported claims and exposure sizes, we get

$$1 \cdot 1.007 \cdot 28,000,000/19,739,000 = 1.43 \qquad (35)$$

That is, assuming no other trends are necessary, we expect to see 1.43 claims above 2 million if year 1996 experience were to happen again with the 2005 exposure size. The following table summarizes the combined adjustments:

Table 8: Indexed claim experience.

Year	No. of claims	No. of *as-if* claims
1995	0	0
1996	1	1.43
1997	4	5.80
1998	3	4.36
1999	2	2.97
2000	2	2.97
2001	0	0
2002	2	3.16
2003	3	5.04
2004	1	2.38

□

To model the claim frequency distribution, we consider three choices of claim frequency distributions: Poisson, negative binomial, and binomial.

- Poisson

  The Poisson distribution is often used in reinsurance pricing for its simplicity. It has a great advantage: the sum of two independent Poisson variables also follows a Poisson distribution. Another advantage is that if the number of claims in a fixed time period follows a Poisson distribution, then

  1. the number of claims above a fixed retention is also Poisson distributed.
  2. the claim number for a subinterval is also Poisson distributed.

The first advantage is particularly useful in excess-of-loss reinsurance pricing because it provides the theoretical background for assuming that loss events are Poisson distribution while adjusting retentions when fitting the distributions. The second advantage works well, for example, when one removes certain benefits from the current plan. The assumption of Poisson needs not be changed if the frequency distribution under the current plan follows a Poisson distribution.

One disadvantage is that the assumption that the rate at which the claims occur is constant over time. This in particular is not applicable in certain sections of reinsurance (for example, earthquake) where the probability of another loss occurring is a lot higher given that one has already occurred.

- Negative binomial

  The negative binomial is a generalization of the Poisson distribution by mixing a Poisson distribution with a gamma mixing distribution. That is, by assuming the parameter λ of the Poisson distribution to be gamma distributed, the resulting distribution is negative binomial. This is particularly useful when the practitioner incorporates parameter uncertainly into the Poisson parameterization.

  The disadvantage of this distribution lies in the difficulty in solving for its maximum likelihood estimators; there is no closed form for them.

- Binomial

  The process of having claims from m independent risks with each risk having probability q of having a claim follows a binomial distribution with parameters m and q. This distribution has finite support $0, 1, 2, \ldots, m$. That is, at most m claims can happen in the specific period of time. This makes the Binomial distribution less popular for reinsurance pricing.

In estimating the parameters of the claim frequency distribution one can use the same methods as for the loss severity distribution; namely, the method of moments, maximum likelihood, least squares, etc. as discussed in Section 3.1. In general, because of the much smaller volatility involved in the claim frequency versus the loss severity, there is less concern in estimating frequency distribution. In our model, we use the method of moments—due to its simplicity—to estimate the parameters of the distribution.

Illustration

In the auto liability example, the average of the sample frequencies is 2.812 and its variance is 3.821. Since the sample variance is larger than the sample mean, we select the negative binomial as the claim frequency distribution. The parameters s, p of a negative binomial distribution are such that[10]

$$\mu = \frac{s(1-p)}{p} \quad and \quad \sigma^2 = \frac{s(1-p)}{p^2} \tag{36}$$

We input the sample mean and sample variance into the equations and solve for s and p in the system of equations with the restriction that s must be an integer. The approximate solutions are $s = 8$ and $p = 0.73993$. This approximation is unbiased but slightly underestimates the variance. □

3.3 Aggregate loss distribution

As stated earlier in this section, we model the aggregate loss distribution from a collective risk theory point of view. The aggregate loss S_n for a specific period of time (usually one calendar year) is the sum of n individual claim amounts (from ground-up):

$$S_n = X_1 + X_2 + \cdots + X_n, \qquad n = 0, 1, 2, \ldots \tag{37}$$

with $S_0 = 0$. n is a random number following the selected claim frequency distribution and the X_is are independent, identically distributed, and follow the selected loss severity distribution. It is also assumed that n and X_i's are independent.

For the annual aggregate layer losses under the reinsurance program, we modify the above formula by applying the reinsurance coverage:

$$\bar{X}_i = \min(l, \max(X_i - r, 0)), \quad \text{for } i = 1, 2, \ldots, n \tag{38}$$

$$\bar{S}_n = \min(L, \max((\bar{X}_1 + \bar{X}_2 + \cdots + \bar{X}_n) - D, 0)) \tag{39}$$

for limit l, retention r, aggregate limit L, and aggregate deductible D.

In general insurance practice, there are four ways in computing the aggregate loss distribution from the selected claim frequency and the loss severity distributions:

1. Method of moments

[10] We are using the following parametrization of the negative binomial density function: $f(k) = \binom{s+k-1}{k} p^s (1-p)^k$.

This method assumes a selected aggregate loss distribution whose parameters are estimated by the empirical moments of the aggregate losses. Such moments are derived algebraically from the moments of the claim frequency and loss severity distributions:

$$\mathrm{E}(\bar{S}_n) = \mathrm{E}(n) \cdot \mathrm{E}(\bar{X}) \tag{40}$$

$$\mathrm{var}(\bar{S}_n) = \mathrm{E}(n) \cdot \mathrm{var}(\bar{X}) + \mathrm{E}(\bar{X})^2 \cdot \mathrm{var}(n) \tag{41}$$

It has the advantage of simplicity and ease of calculation. However, its main disadvantage is inaccuracy. In general, the fitted loss distribution does not model the true aggregate losses well.

2. Monte Carlo simulation

   This method calculates the aggregate loss distribution directly from simulating the claim frequency and loss severity distributions. First, one samples from the frequency distribution to determine the number of claims n in the period. Then, we pick n claims from the severity distribution at random. The sum of the n random claim amounts (adjusted for the reinsurance coverage in place) gives one outcome for the aggregate losses. We repeat many times this sampling procedure to estimate the distribution of the aggregate losses.

   This method provides easy and accurate aggregate distributions. However, some argue that it takes considerable computing time. For further details see [12].

3. Recursive method

   In general, this method requires a discretization of the loss severity distribution and a selection of a large enough number of points for the claim frequency distribution. It involves inverting the Laplace transform of the aggregate loss distribution (for example, see [14]). Panjer [23] gave a direct recursive formula for a particular family of claim frequency distributions that does not involve the Laplace transformation (see also [26])

   The recursive method is fast and accurate most of the time. The disadvantage is the requirement of discretization of the loss severity distribution. There are two methods for carrying out the discretization of a continuous distribution function: the midpoint method and the unchanged expectation method. With both methods one loses information. For further details see [6].

4. Fast Fourier Transform

   The fast Fourier transformation inverts the characteristic function of the aggregate loss distribution, a procedure similar to the recursive method. It also requires a discretization of the loss severity distribution. See [3, 26] for further details.

   The advantage of this method lies in its efficiency and speed. However, the computation tends to be complicated.

We use Monte Carlo simulation to compute the aggregate loss distribution in our model. It is simple and intuitive.

Illustration

Assume that we have fitted a generalized Pareto distribution as our loss severity distribution with parameters $\xi = 0.66784$, $\sigma = 591,059.8$, and threshold of $2,000,000$ and a negative binomial as our claim frequency distribution with parameters $s = 8$ and $p = 0.73993$. Let's also assume that, in one random iteration, the negative binomial distribution produces 4 claims and we generate the following 4 claims from the generalized Pareto distribution:

Table 9: Sample GPD generated losses.

Loss (from ground up)
2,590,062
3,107,208
2,874,384
7,800,324

Keep in mind that the generalized Pareto generates excess loss above the threshold. To convert excess loss to ground up loss one adds back the threshold.

This represents one possible annual outcome. To evaluate the reinsurance recovery, we apply the coverage: 12 million excess of 3 million with annual aggregate deductible of 3 million. The table bellow shows the result.

Table 10: Sample layer losses.

Loss	Layer Loss
2,590,062	0
3,107,208	107,208
2,874,384	0
7,800,324	4,800,324

The sum of all layer losses is $4,907,532$ and the reinsurance recovery is $1,907,532$. Thousands of iterations are generated to derive all possible outcomes and they give us the aggregate loss distribution for the reinsurance coverage. The following graph shows the result of $5,000$ iterations:

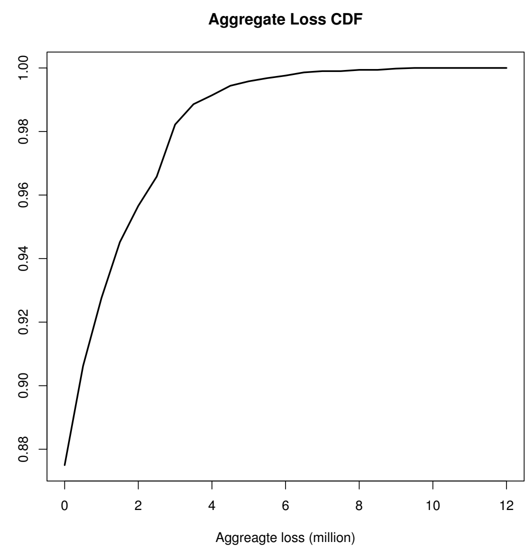

Figure 14: Generated aggregate loss distribution.

The resulting aggregate loss distribution is as expected highly skewed, with 78.1% probability of no losses and a mean of $1,108,974. □

4 Risk Loads and Capital Requirements

An essential job of actuaries is to quantify risks. In this section, we will introduce various risk measures and capital requirements that could be incorporated in our pricing model. We would like to stress that quantifying risks is a complex undertaking. So far no single risk measure can fulfill all of the properties that actuaries would like to have.

For example, the risk measures standard deviation or variance are symmetric. They do not differentiate between losses and gains. In practice, actuaries and management teams are concerned with the management of potential losses. Another risk measure, value at risk, is concerned with potential losses above a threshold. Unfortunately, this measure does not tell us anything about how severe losses could be if they exceed the threshold. Probability of ruin is another risk measure that actuaries have spent considerable time studying. Here the actuary would set the capital requirements of a company so that the probability of ruin is acceptably small. Similar to value at risk, this measure provides no information about the severity of ruin.

4.1 Risk Measures

Various risk measures have appeared in the actuarial literature: standard deviation, variance, probability of ruin, value at risk (VaR), expected policyholder deficit (EPD), and tail VaR are among them.

- Standard deviation or variance

  Standard deviation and variance methods equate more volatility in the loss distribution with more riskiness. These methods set a risk load directly proportional to the standard deviation or variance. They are popular for their simplicity and mathematical tractability. However, they ignore the distinction between the upside and downside risks, which is critical for proper pricing especially when the loss distribution is highly skewed.

- Probability of ruin

  Probability of ruin focuses on the theoretical ruin threshold, the point where the liabilities are greater than the assets. For a probability ϵ, it seeks the capital amount such that

  $$\text{Prob}(X \leq \text{capital} + \text{E}(X)) = 1 - \epsilon \tag{42}$$

  Probability of ruin is easy to understand and to compute. Unfortunately, it considers only the probability of ruin and lacks the consideration of loss severity when ruin occurs.

- Value at risk (VaR)

Value at Risk is generally defined as the capital necessary, in most cases, to cover the losses from a portfolio over a specified holding period. The VaR is defined as the smallest value that is greater than a predetermined percentile of the loss distribution. That is, for a selected probability α,

$$\mathrm{VaR}_\alpha = \inf\{x|\,\mathrm{Prob}(X \le x) > \alpha\} \tag{43}$$

Similar to the probability of ruin risk measure, VaR is easy to understand and to compute but lacks the consideration of loss severity.

- Policyholder deficit (EPD)

  Expected Policyholder Deficit is the expected value of the difference between the amount the insurer is obligated to pay the claimant and the actual amount paid by the insurer, provided that the former is greater. Mathematically, it is

$$\int_{\mathrm{capital}+\mathrm{E}(X)}^{\infty} (x - \mathrm{capital} - \mathrm{E}(X)) \cdot f(x)\,dx \tag{44}$$

  where $\mathrm{E}(X)$ is the expected loss and capital refers to the excess of assets over liabilities. When considering the EPD as the risk measure, one usually uses the ratio of the EPD to the expected loss (called the EPD ratio) to adjust to the scale of different risk element sizes. That is, an EPD ratio of ϵ would set capital to be the amount such that

$$\frac{\int_{\mathrm{capital}+\mathrm{E}(X)}^{\infty} (x - \mathrm{capital} - \mathrm{E}(X)) \cdot f(x)\,dx}{\mathrm{E}(X)} = \epsilon \tag{45}$$

  Expected policyholder deficit considers the severity as well as the probability of the deficit. However, it is highly sensitive to extreme events.

- Tail value at risk (Tail VaR)

  Tail Value at Risk is also called tail conditional expectation or expected shortfall and it is the conditional expected value of losses:

$$\mathrm{TailVaR}_\alpha(X) = \mathrm{E}\left(X|X \ge \mathrm{VaR}_\alpha(X)\right) \tag{46}$$

  That is, for a selected probability α, TailVaR at α is the expected value of those losses greater than or equal to $\mathrm{VaR}_\alpha(X)$. Unlike VaR, TailVaR considers the loss severity. It is also less sensitive to extreme losses than the expected policyholder deficit measure.

Among the risk measures stated so far in this section, TailVaR is the only *coherent* risk measure in the sense discussed in the paper *Coherent Measures of Risk* [1]. A risk measure is said to be coherent if it satisfies four axioms: sub-additivity, monotonicity, positive homogeneity, and translation invariance.

The sub-additivity axiom ensures that the merging of two portfolios of risks does not create extra risk. Monotonicity says that if portfolio X always generates losses smaller than portfolio Y, then the risk measure for X should not be larger than that of Y. Positive homogeneity tells us that merging two *identical* portfolios doubles the risk measure. Finally, translation invariance says that if we add a constant to all losses of a portfolio, then the risk measure of the portfolio should also increase by the same constant (see [1, 21]). For our model we have selected Tail VaR at 99% as our risk measure.

In addition to the risk measures mentioned above, there are many other risk measures such as CAPM (see [9]), marginal cost of capital with and without the application of game theory [18, 22], and worse conditional expectation [1] (this is also a coherent risk measure). One should also take them into consideration when selecting a risk measure.

4.2 Capital Requirements

For insurance companies operating in the United States, the capital requirements are heavily regulated by the NAIC risk-based capital standards. Reinsurance companies (especially non-US reinsurers), on the other hand, are usually not as heavily regulated as primary insurance companies are with respect to capital requirements. In our model we will simplify matters and not consider how the NAIC risk based capital requirements would be affected in the pricing of a single excess-of-loss treaty. Rather we will take the position that we are evaluating the treaty on a stand-alone-basis. Moreover, the capital requirements are directly tied to the selected risk measure.

One main reason to purchase reinsurance is to mitigate large losses and to reduce the volatility of the underwriting results. As a result of the reinsurance purchase, the amount of capital required to guard against unexpected losses and high volatility is reduced. We consider the reduction in capital as *rented* capital from the reinsurer. That is, the reduction in required capital from before reinsurance to after reinsurance is the amount of capital that reinsurer provides. The ceding insurer must pay a fee for renting this capital from the reinsurer. The amount of rented capital depends on the reinsurance coverage and can be computed directly from the simulation results.

Illustration

The simulation results from our auto liability example are as follows:

Table 11: Aggregate loss distribution statistics before and after reinsurance.

Percentile	Gross	Net
5.0 %	0	0
10.0 %	2,034,094	2,034,094
25.0 %	4,105,924	4,105,924
50.0 %	7,578,306	7,459,913
75.0 %	13,198,430	12,000,000
90.0 %	20,148,295	17,003,454
92.0 %	22,175,644	18,257,892
98.0 %	37,985,818	27,054,490
99.0 %	51,212,932	38,456,724
99.5 %	76,985,851	62,732,939
99.8 %	143,818,339	122,821,239
100.0 %	904,128,288	882,361,851

The following graph makes the reinsurance effect clearer:

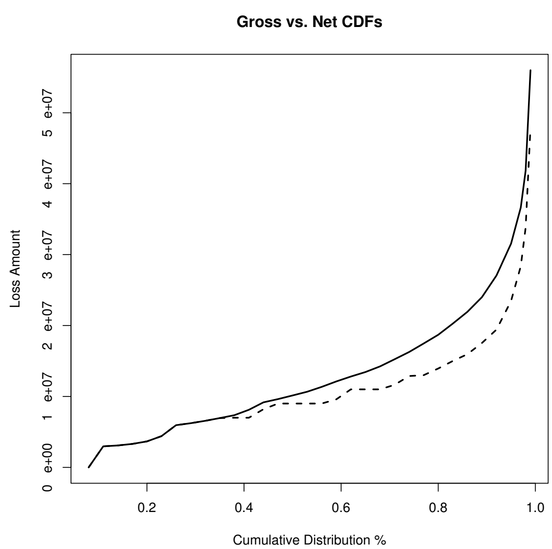

Figure 15: Gross versus net loss distributions.

With the chosen risk measure of Tail VaR at 99%, we first find the VaR at 99% and then compute the conditional expected value of losses given that they are larger than or equal to the 99% VaR. In our simulation we have to take the average of all losses greater than or equal to the 99% VaR value. The following table shows the VaR values and the Tail VaR values at 99% on a gross and net bases.

Table 12: Tail VaR calculation.

	Gross	Net
$\text{VaR}_{99\%}$	51,212,932	38,456,724
$\text{TailVaR}_{99\%}$	128,583,553	115,354,489

Therefore, we have a capital reduction of

$$128,583,553 - 115,354,489 = 13,229,064$$

and this is the amount of *rented* capital that will be incorporated in the premium calculation. □

5 Reinsurance IRR Pricing Model

Our pricing methodology follows closely the paper *Financial Pricing Model for Property and Casualty Insurance Products: Modeling the Equity Flows* [10]. Readers interested in the reasoning and intuitions of the details should refer to the paper.

The goal of IRR pricing is to generate the equity flows (net cash flows) associated with the treaty being priced. The amount of premium is an unknown that must be solved for so that the IRR on the resulting equity flows is equal to the pricing target. In theory an iterative process is used to solve for the premium. In practice the premium is found by running the goal seek algorithm in Excel©.

With the objective of generating the equity flow, the model is designed to calculate the cash flows necessary for the calculation of the equity flow. These cash flows are:

- U/W cash flows
- Investment income cash flows
- Federal income tax flows ("+" denotes a refund; "−" denotes a payment)

- Asset flows

The equity flow is then calculated via the basic relation:[11]

$$\text{Equity Flow} = \text{U/W Flow} + \text{Investment Income Flow} + \text{Tax Flow} - \text{Asset Flow} + \text{DTA Flow}. \quad (47)$$

We use the convention that a positive equity flow denotes a flow of cash from the insurer to the equityholders, and a minus a payment by the equityholders to the insurer.

5.1 Illustration

Recall the illustrative excess-of-loss treaty from Section 3. It is assumed to be effective Jan 1, 2005.

Certain treaty characteristics serve as inputs to the model. These characteristics consist of the following *costs* for the layer being priced:

- amount of expected ultimate loss for the layer ($1,108,974 as detailed in Section 3),
- brokerage expenses as a percentage of base premium (10%),
- LAE as a percentage of base premium (3%),

and the following collection/payment *patterns*:

- premium collection pattern (assumed to be 100% at treaty inception)
- loss payment pattern:

Table 13: Loss payment pattern.

Year	Loss payment	Year	Loss payment
2005	22.2%	2010	4.7%
2006	29.3	2011	4.3
2007	15.9	2012	3.7
2008	7.9	2013	3.5
2009	5.8	2014	2.7

[11] Properly speaking, the change in the deferred tax asset (DTA) is not a cash flow, if by cash one means cash equivalents. Since assets include DTA, the change in assets is also not a cash flow. But, then item (assets − DTA) consists of cash equivalents and hence Δ(assets − DTA) is a cash flow. We have simply expressed Δ(assets − DTA) as the difference Δassets minus ΔDTA.

The model uses annual valuations. With the exception of the written premium and UEPR which incept on Jan 1, our simplifying assumption is that all accounting and cash flow activity occur at year end.

The model also requires certain *parameter* inputs consisting of:

- investment rate of return on invested assets (5.5%),
- effective tax rate for both investments and U/W income (assumed to be 35% for both),
- surplus assumptions (specifics discussed in Section 4.2 above),
- the target return on capital (12%), and
- IRS loss & LAE reserve discount factors:

Table 14: Internal Revenue Service discount factors.

Year	Discount Factor	Year	Discount Factor	Year	Discount Factor
2005	0.7410	2010	0.7583	2015	0.8805
2006	0.7367	2011	0.7554	2016	0.9221
2007	0.7438	2012	0.7823	2017	0.9766
2008	0.7040	2013	0.8117	2018	0.9766
2009	0.7264	2014	0.8441	beyond	0.9766

5.1.1 Assets

Required Surplus

Surplus is held only for the policy term in our illustration. It exists to cover unforeseen contingencies and is determined to maintain an acceptable level of risk. As discussed above in section 4.2 we used a rented capital approach using a 99% TVaR level of risk to give a surplus need of $13,229,064.This surplus we assume is held for the first year only. This assumption reflects the fact that new business writings pose a greater risk than business in reserve run-off which has capital embedded in reserves to support unforeseen contingencies.

Total Reserve

The total reserve at any point in time is the sum of the unearned premium reserve and the held loss & LAE reserves. We assume no reserve deficiency and so as losses pay out the held loss reserves are taken down dollar for dollar.

Required Assets

The amount of assets the insurance company needs to support the policy is equal to total reserves as defined above, plus the required:

$$\text{Required Assets} = \text{Total Reserves} + \text{Required Surplus} \tag{48}$$

Illustration

On Jan 1, 2005 the UEPR is equal to the WP which is $3,044,605. The contract is yet fully unearned and the loss reserves are $0. With surplus of $13,229,064 put up at treaty inception the total assets are $16,273,669. By year end the UEPR is $0, loss reserves are equal to $862,782 (ultimate losses less paid losses of $246,192), and surplus is $0 for total assets of $862,782.

Table 15: Asset calculation.

	UEPR	Held Reserve	Surplus Capital	Held Asset
1/ 1/2005	3,044,605	0	13,229,064	16,273,669
12/31/2005	0	862,782	0	862,782
12/31/2006	0	537,964	0	537,964
12/31/2007	0	361,694	0	361,694
12/31/2008	0	274,218	0	274,218
12/31/2009	0	209,898	0	209,898
12/31/2010	0	157,776	0	157,776
12/31/2011	0	110,090	0	110,090
12/31/2012	0	69,058	0	69,058
12/31/2013	0	30,244	0	30,244
12/31/2014	0	0	0	0

□

Income Producing Assets

Not all of the assets held by the company to support the policy generate investment income. Both the premium receivable (if any) and the deferred tax asset are non-income producing assets:

$$\text{Income Producing Assets} = \text{Required Assets} - \text{Premium Receivable} - \text{DTA} \tag{49}$$

In our formulation weve assumed all premium is collected up front and consequently the premium receivable asset is zero. The calculation of DTA is discussed below.

Investment Income

The Investment Income earned over the year is simply calculated as the product of the annual effective investment rate of return times the amount of income producing assets held at the beginning of the year:

$$\text{Invest Inc}_{@\text{time } T} = \text{Annual Invest ROR} \cdot \text{Investible Assets}_{@\text{time } T-1} \tag{50}$$

Table 16: Investment income calculation.

	Held Asset	Non-Income Producing	Income Producing	Investment Income
1/ 1/2005	16,273,669	0	16,273,669	0
12/31/2005	862,782	28,656	834,125	895,052
12/31/2006	537,964	17,134	520,830	45,877
12/31/2007	361,694	4,028	357,666	28,646
12/31/2008	274,218	8,307	265,911	19,672
12/31/2009	209,898	6,756	203,142	14,625
12/31/2010	157,776	3,919	153,857	11,173
12/31/2011	110,090	4,163	105,927	8,462
12/31/2012	69,058	3,269	65,789	5,826
12/31/2013	30,244	1,994	28,251	3,618
12/31/2014	0	0	0	1,554

5.1.2 Taxes

IRS Discounted Reserves

The IRS Discounted Reserves are calculated by multiplying the company's Held Reserves by a discount factor. The discount factor varies by line of business, accident year, and by age of the accident year. Our basic formula for IRS discounted reserves is thus

$$\text{IRS Discounted Reserves} = \text{IRS Discount Factor} \cdot \text{Held Reseves} \tag{51}$$

Taxable U/W Income

The IRS defines the taxable U/W income earned over an accounting year as

$$\begin{aligned}\text{Written Premium} - 0.8 \cdot \Delta\text{UEPR} - \text{Paid Expenses}\\ - [\text{Paid Losses} + \Delta\text{IRS Disc Reserves}]\end{aligned} \quad (52)$$

where all activity is over the relevant accounting year. In our illustration the treaty is effective Jan 1 and so the change in the UEPR is identically zero. Table 17 shows the computation of the tax on U/W income.

Tax on Investment Income

This tax is simply calculated as 35% of earned investment income for the year.

Total Tax

The total federal income tax paid each year is equal to the sum of the yearly tax on U/W income and the yearly tax on investment income.

Table 18: Tax calculation.

	Tax on UW Inc	Tax on Inv Inc	Tax Total
1/ 1/2005			
12/31/2005	617,167	313,268	930,436
12/31/2006	−28,656	16,057	−12,599
12/31/2007	−17,134	10,026	−7,108
12/31/2008	−4,028	6,885	2,857
12/31/2009	−8,307	5,119	−3,188
12/31/2010	−6,756	3,910	−2,845
12/31/2011	−3,919	2,962	−957
12/31/2012	−4,163	2,039	−2,124
12/31/2013	−3,269	1,266	−2,003
12/31/2014	−1,994	544	−1,450

Deferred Tax Asset

There are two components to the DTA: the portion due to the Revenue Offset; and the portion due to IRS Discounting of Loss & LAE Reserves.

The DTA due to the Revenue Offset is equal to

$$35\% \cdot 20\% \cdot \text{UEPR} \quad (53)$$

	WP	Expenses	Held Reserve	Paid Loss	IRS Disc Factors	IRS Disc Reserves	Chg in Disc Reserves	Taxable UW Income
1/ 1/2005	3,044,605	395,799	0	0				
12/31/2005			862,782	246,192	0.7410	639,279	639,279	1,763,335
12/31/2006			537,964	324,817	0.7367	396,336	−242,943	−81,874
12/31/2007			361,694	176,270	0.7438	269,021	−127,315	−48,955
12/31/2008			274,218	87,476	0.7040	193,054	−75,968	−11,508
12/31/2009			209,898	64,320	0.7264	152,468	−40,586	−23,735
12/31/2010			157,776	52,122	0.7583	119,648	−32,820	−19,301
12/31/2011			110,090	47,686	0.7554	83,160	−36,488	−11,198
12/31/2012			69,058	41,032	0.7823	54,022	−29,137	−11,895
12/31/2013			30,244	38,814	0.8117	24,548	−29,474	−9,340
12/31/2014			0	30,244	0.8441	0	−24,548	−5,696

For our illustration the year end UEPR is identically equal to zero.

The DTA due to IRS Discounting at the end of Accounting Year T is equal to

$$35\% \cdot [(\text{Held Loss Reserve}_{\text{at time } T} - \text{IRS Loss Reserve}_{\text{at time } T}) - (\text{Held Loss Reserve}_{\text{at time } T+1} - \text{IRS Loss Reserve}_{\text{at time } T+1})] \quad (54)$$

The amount in each square bracket is the amount that reverses in the year (which is all that is statutorily recognized).

Illustration

At year end 2005 the held loss reserve is $862,782 while the IRS discounted reserve is $639,279. If the full DTA were recognized it would be $(862,782-639,279) \cdot 35\% = 78,226$. But only the amount that reverses in one year is recognized. That is, the fully recognized DTA at year end 2006 would be $(537,964 - 396,336) \cdot 35\% = 49,570$. Thus the amount that reverses during 2006 is $28,656 and this is the amount of DTA at year end 2005. □

Table 19: Deferred tax asset calculation.

	Held Reserve	IRS Disc Reserves	DTA (Reserve Disc)
1/ 1/2005			
12/31/2005	862,782	639,279	28,656
12/31/2006	537,964	396,336	17,134
12/31/2007	361,694	269,021	4,028
12/31/2008	274,218	193,054	8,307
12/31/2009	209,898	152,468	6,756
12/31/2010	157,776	119,648	3,919
12/31/2011	110,090	83,160	4,163
12/31/2012	69,058	54,022	3,269
12/31/2013	30,244	24,548	1,994
12/31/2014	0	0	0

5.1.3 Cash Flows

The relevant cash flows for determining the Equity Flow are described below.

U/W Cash Flow

This item is defined as

$$\text{U/W Cash Flow} = \text{WP} - \text{Paid Expenses} - \text{Paid Loss} \tag{55}$$

Investment Income Flow

This item is defined as the yearly investment income earned. The calculation is described above.

Tax Flow

The Tax Cash Flow is defined at the negative (to denote a flow from the company) of the federal income taxes paid that year. The calculation of this flow item is described above.

DTA Flow

The DTA Flow is defined as the change in the DTA asset over a year.

Asset Flow

The asset flow is defined as the change in the required assets. The composition and calculation of the required assets are described above.

Equity Flow

To compute the Equity Flow at each year we use the cash flow definition:

$$\begin{aligned}\text{Equity Flow} = {} & -\text{Asset Flow} + \text{U/W Flow} \\ & + \text{Investment Income Flow} + \text{FIT Flow} + \text{DTA Flow}\end{aligned} \tag{56}$$

Recall that we use the convention that a positive equity flow denotes a flow of cash from the insurer to the equityholders, and a negative a payment by the equityholders to the insurer. The relevant cash flows for our illustration are summarized in Table 20.

The IRR on the resulting equity flows is 12%. The premium of $3,044,605 was iteratively determined with this goal.

	Cash UW	Inv Income	Tax	DTA	Asset	Equity Flow
1/ 1/2005	2,648,806	0	16,273,669	0	0	−13,624,863
12/31/2005	−246,192	895,052	−15,410,887	−930,436	28,656	15,157,968
12/31/2006	−324,817	45,877	−324,817	12,599	−11,522	46,954
12/31/2007	−176,270	28,646	−176,270	7,108	−13,107	22,648
12/31/2008	−87,476	19,672	−87,476	−2,857	4,279	21,094
12/31/2009	−64,320	14,625	−64,320	3,188	−1,552	16,262
12/31/2010	−52,122	11,173	−52,122	2,845	−2,836	11,182
12/31/2011	−47,686	8,462	−47,686	957	244	9,663
12/31/2012	−41,032	5,826	−41,032	2,124	−894	7,056
12/31/2013	−38,814	3,618	−38,814	2,003	−1,275	4,346
12/31/2014	−30,244	1,554	−30,244	1,450	−1,994	1,010

6 Summary

The pricing of high layers of reinsurance is a difficult task primarily because of the nature of extreme events. The practicing actuary requires a diverse toolbox to tackle this pricing problem. As part of the toolbox he needs well grounded statistical methods for analyzing the data at hand, a good understanding of the modeling techniques and risk assessment, and a comprehensive pricing model that does not sweep under the rug many of the regulatory, tax, and business constraints of the insurance company.

In the past very large losses would be labeled as *outlier observations*, rationalized as extremely improbable, and sometimes even removed from the data set. For the reinsurance actuary these observations are likely to be the most important observations in the data set.

In this paper we have introduced results from a well grounded statistical theory to deal with extreme events. The first result tells us that the distribution of the maximum of a sample converges to one of the three extreme value distributions. This result is analogous to the central limit theorem. The second result shows that the distribution of excess losses converges to the generalized Pareto distribution as the threshold increases. This is the key result for pricing very high layers of reinsurance. We also introduce the peaks over threshold method from extreme value theory and a powerful graphical technique, the QQ-plot, to assess distributional assumptions.

The paper also provides a hands-on approach to loss modeling. We present the collective risk model and use it to calculate the aggregate loss distribution for the example that is carried throughout the paper. We also introduced various measures to quantify risk and our treatment of capital requirements. Our discussions on collective risk models and risk measures are by no means complete but the framework we have laid should provide the practicing actuary with a foundation that can be put to practice.

Finally, the cash flow model (IRR pricing model) brings everything together to determine the price of a reinsurance layer. It is designed to calculate the equity flows; that is, the cash flows between the company and its equity holders. This pricing model is comprehensive: it includes all relevant components of cash flow for an insurance company to derive the final price given the risk premium and other parameters.

Appendix A

In this appendix, we will show that the distribution $\widehat{F(x)}$ in equation (21) in section 2.4 is a generalized Pareto distribution by deriving the associate parameters.

$$\begin{aligned}
\widehat{F(x)} &= (1 - F_n(u))\, G_{\xi,\sigma}(x-u) + F_n(u) \\
&= (1 - F_n(u)) \left(1 - \left(1 + \frac{\xi}{\sigma}(x-u)\right)^{-1/\xi}\right) + F_n(u) \\
&= 1 - (1 - F_n(u)) \left(1 + \frac{\xi}{\sigma}(x-u)\right)^{-1/\xi} \\
&= 1 - \left((1 - F_n(u))^{-\xi} \left(1 + \frac{\xi}{\sigma}(x-u)\right)\right)^{-1/\xi} \\
&= 1 - \left((1 - F_n(u))^{-\xi} + \frac{\xi}{(1 - F_n(u))^{\xi} \cdot \sigma}(x-u)\right)^{-1/\xi} \\
&= 1 - \left(1 + \frac{\xi}{(1 - F_n(u))^{\xi}\, \sigma} \left(x - \left\{u - \frac{\sigma}{\xi}\left[1 - (1 - F_n(u))^{\xi}\right]\right\}\right)\right)^{-1/\xi} \\
&= 1 - \left(1 + \frac{\xi}{\tilde{\sigma}}(x - \tilde{u})\right)^{-1/\xi} \\
&= G_{\xi,\tilde{\sigma}}(x - \tilde{u})
\end{aligned}$$

where $\tilde{\sigma} = \sigma(1 - F_n(u))^{\xi}$ and $\tilde{u} = u - [\sigma(1 - (1 - F_n(u))^{\xi})/\xi]$.

Appendix B *IRR cash flow model exhibits*

	WP (1)	UEPR (2)	LAE (3)	Brokerage (4)	Paid Loss (5)	Nominal Reserve (6)
11/ 1/2005	3,044,605	3,044,605	91,338	304,461		
12/31/2005					246,192	862,782
12/31/2006					324,817	537,964
12/31/2007					176,270	361,694
12/31/2008					87,476	274,218
12/31/2009					64,320	209,898
12/31/2010					52,122	157,776
12/31/2011					47,686	110,090
12/31/2012					41,032	69,058
12/31/2013					38,814	30,244
12/31/2014					30,244	0

$(3) = (1) \cdot 3.0\%$
$(4) = (1) \cdot 10.0\%$

	Held Reserve (7)	Surplus Capital (8)	Held Asset (9)	Non-Income Producing (10)	Income Producing (11)
1/ 1/2005	0	13,229,064	16,273,669	0	16,273,669
12/31/2005	862,782	0	862,782	28,656	834,125
12/31/2006	537,964	0	537,964	17,134	520,830
12/31/2007	361,694	0	361,694	4,028	357,666
12/31/2008	274,218	0	274,218	8,307	265,911
12/31/2009	209,898	0	209,898	6,756	203,142
12/31/2010	157,776	0	157,776	3,919	153,857
12/31/2011	110,090	0	110,090	4,163	105,927
12/31/2012	69,058	0	69,058	3,269	65,789
12/31/2013	30,244	0	30,244	1,994	28,251
12/31/2014	0	0	0	0	0

$(7) = 100\% \cdot (6)$
$(9) = (2) + (7) + (8)$
$(10) = (21)$
$(11) = (9) - (10)$

	Investment Income (12)	IRS Disc Factors (13)	Taxable UW income (14)	Tax Paid UW (15)	Taxable Invest (16)
1/ 1/2005	0				
12/31/2005	895,052	0.7410	1,763,335	617,167	895,052
12/31/2006	45,877	0.7367	−81,874	−28,656	45,877
12/31/2007	28,646	0.7438	−48,955	−17,134	28,646
12/31/2008	19,672	0.7040	−11,508	−4,028	19,672
12/31/2009	14,625	0.7264	−23,735	−8,307	14,625
12/31/2010	11,173	0.7583	−19,301	−6,756	11,173
12/31/2011	8,462	0.7554	−11,198	−3,919	8,462
12/31/2012	5,826	0.7823	−11,895	−4,163	5,826
12/31/2013	3,618	0.8117	−9,340	−3,269	3,618
12/31/2014	1,554	0.8441	−5,696	−1,994	1,554

$(12)_t = (11)_{t-1} \cdot 5.5\%$
$(14)_t = (1)_t - 80\% \cdot ((2)_t - (2)_{t-1}) - (3)_t - (4)_t - (5)_t - ((7)_t \cdot (13)_t - (7)_{t-1} \cdot (13)_{t-1})$
$(15) = (14) \cdot 35\%$
$(16) = (12)$

	Tax Paid Inv Inc (17)	Total Tax Paid (18)	DTA Revenue Offset (19)	DTA Reserve Disc (20)	Total DTA (21)
1/ 1/2005		0	0	0	0
12/31/2005	313,268	930,436	0	28,656	28,656
12/31/2006	16,057	−12,599	0	17,134	17,134
12/31/2007	10,026	−7,108	0	4,028	4,028
12/31/2008	6,885	2,857	0	8,307	8,307
12/31/2009	5,119	−3,188	0	6,756	6,756
12/31/2010	3,910	−2,845	0	3,919	3,919
12/31/2011	2,962	−957	0	4,163	4,163
12/31/2012	2,039	−2,124	0	3,269	3,269
12/31/2013	1,266	−2,003	0	1,994	1,994
12/31/2014	544	−1,450	0	0	0

$(17) = (16) \cdot 35\%$
$(18) = (15) + (17)$
$(19) = 20\% \cdot (2) \cdot 35\%$
$(20)_t = (((7)_t - (7)_{t-1}) - ((7)_t \cdot (13)_t - (7)_{t-1} \cdot (13)_{t-1}) \cdot 35\%$
$(21) = (19) + (20)$

	Cash UW (22)	Investment Income (23)	Asset (24)	Tax (25)	DTA (26)	Equity Flow (27)
1/ 1/2005	2,648,806	0	16,273,669	0	0	−13,624,863
12/31/2005	−246,192	895,052	−15,410,887	−930,436	28,656	15,157,968
12/31/2006	−324,817	45,877	−324,817	12,599	−11,522	46,954
12/31/2007	−176,270	28,646	−176,270	7,108	−13,107	22,648
12/31/2008	−87,476	19,672	−87,476	−2,857	4,279	21,094
12/31/2009	−64,320	14,625	−64,320	3,188	−1,552	16,262
12/31/2010	−52,122	11,173	−52,122	2,845	−2,836	11,182
12/31/2011	−47,686	8,462	−47,686	957	244	9,663
12/31/2012	−41,032	5,826	−41,032	2,124	−894	7,056
12/31/2013	−38,814	3,618	−38,814	2,003	−1,275	4,346
12/31/2014	−30,244	1,554	−30,244	1,450	−1,994	1,010

$(22) = (1) - (3) - (4) - (5)$
$(23) = (16)$
$(24)_t = (9)_t - (9)_{t-1}$
$(25) = -(18)$
$(26)_t = (21)_t - (21)_{t-1}$.
$(27) = -(24) + (22) + (23) + (25) + (26)$

Appendix C *Distribution Functions*

In Table 21 (on the next page) we present some common distribution functions and their parametrizations.

Name	CDF F or density f	Parameters
Exponential	$F(x) = 1 - e^{-\lambda x}$	$\lambda > 0$
Gamma	$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$	$\alpha, \beta > 0$
Weibull	$F(x) = 1 - e^{-cx^\tau}$	$c > 0, \tau > 0$
Log-normal	$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} e^{-(\ln x - \mu)^2/(2\sigma^2)}$	$\mu \in \mathbb{R}, \sigma > 0$
Pareto	$F(x) = 1 - \left(\frac{\kappa}{\kappa + x}\right)^\alpha$	$\alpha, \kappa > 0$
Benktander-type-I	$F(x) = 1 - (1 + 2(\beta/\alpha)\ln x) e^{-\beta(\ln x)^2 - (\alpha+1)\ln x}$	$\alpha, \beta > 0$
Benktander-type-II	$F(x) 1 - e^{\alpha/\beta} x^{-(1-\beta)} e^{-\alpha x^\beta/\beta}$	$\alpha > 0, 0 < \beta < 1$
Log-gamma	$f(x) = \frac{\alpha^\beta}{\Gamma(\beta)} (\ln x)^{\beta-1} x^{-\alpha-1}$	$\alpha, \beta > 0$

References

[1] Philippe Artzner, Freddy Delbaen, Eber Jean-Marc, and David Heath. Coherent measures of risk. *Math. Finance*, 9(3):203–228, 1999.

[2] A. A. Balkema and L. de Haan. Residual life time at great age. *Annals of Probability*, 2, 1974.

[3] E. O. Brigham. *The Fast Fourier Transform and Its Applications*. Prentice Hall, Englewood, NJ, 1988.

[4] Hans Bühlmann. *Mathematical Methods in Risk Theory*, volume 172 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1996.

[5] John M. Chambers, William S. Cleveland, Beat Kleiner, and Paul A. Tukey. *Graphical methods for data analysis*. The Wadsworth Statistics/Probability Series. Wadsworth International Group, 1983.

[6] Chris D. Daykin, Teivo Pentikäinen, and Martti Pesonen. *Practical Risk Theory for Actuaries*. Number 53 in Monographs on Statistics and Applied Probability. Chapman & Hall, London, 1994.

[7] Morris H. DeGroot. *Probability and Statistics*. Addison-Wesley Series in Statistics. Addison-Wesley Publishing Company, Reading, Massachusetts, second edition, 1986.

[8] Paul Embrechts, Claudia Klüppelberg, and Thomas Mikosch. *Modeling Extremal Events*, volume 33 of *Applications of Mathematics*. Springer-Verlag, Berlin, 1997.

[9] Sholom Feldblum. Risk loads for insurers. *Proceedings of the Casualty Actuarial Society*, 77:160–195, 1990.

[10] Sholom Feldblum and Neeza Thandi. Financial pricing models for property-casualty insurance products: Modeling the equity flows. *Casualty Actuarial Society Forum*, Winter:445–590, 2003.

[11] R. A. Fischer and L. H. C. Tippett. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Cambridge Philos. Soc.*, 24:180–190, 1928.

[12] G. S. Fishman. *Monte Carlo: Concepts, Algorithms, and Applications*. Springer-Verlag, New York, 1996.

[13] Emil Julius Gumbel. *Statistics of Extremes*. Dover Publications, Inc., 2004.

[14] P. E. Heckman and G. G. Meyers. The calculation of aggregate loss distributions from claim severity and claim count distributions. *Proceedings of the Casualty Actuarial Society*, 70:22–61, 1983.

[15] Robert V. Hogg and Allen T. Craig. *Introduction to Mathematical Statistics*. Macmillan Publishing Company, New York, fourth edition, 1978.

[16] McNeil Alexander J. and Thomas Saladin. The peaks over thresholds method for estimating high quantiles of loss distributions. In *Proceedings of the 28th International ASTIN Colloquium*, 1997.

[17] Stuart A. Klugman, Harry H. Panjer, and Gordon E. Willmot. *Loss Models: from data to decisions*. Wiley series in probability and statistics. John Wiley & Sons, New York, 1998.

[18] Jean Lemaire. An application of game theory: Cost allocation. *ASTIN Bulletin*, 14(1):61–82, 1984.

[19] Alexander J. McNeil. Estimating the tails of loss severity distributions using extreme value theory. *ASTIN Bulletin*, 27(1):117–137, 1997.

[20] Alexander J. McNeil and Thomas Saladin. Developing scenarios for future extreme losses using the pot model. preprint, July 1998.

[21] Glenn Meyers. Coherent measures of risk: An exposition for the lay actuary. *Casualty Actuarial Society Forum*, Summer 2000.

[22] Stuart C. Myers and James A. Jr. Read. Capital allocation for insurance companies. *The Journal of Risk and Insurance*, 68(4):545–580, 2001.

[23] H. H. Panjer. Recursive evaluation of a family of compound distributions. *ASTIN Bulletin*, 12:22–26, 1981.

[24] J. Pickands. Statistical inference using extreme order statistics. *Annals of Statistics*, (3):119–131, 1975.

[25] Rolf-Dieter Reiss and Michael Thomas. *Statistical Analysis of Extreme Values: with Applications to Insurance, Finance, Hydrology and Other Fields.* Birkhäuser Verlag, Basel, Switzerland, second edition, 2001.

[26] Shaun S. Wang. Aggregation of correlated risk portfolios: Models and algorithms. *Proceedings of the Casualty Actuarial Society*, 85:848–939, 1998.

Biographies of the Authors

Doris Schirmacher is an Actuary with the Liberty Mutual Group in Boston. Ms. Schirmacher received her Ph.D. in mathematics from the University of Minnesota and is a Fellow of the Casualty Actuarial Society (CAS). Prior to joining the Liberty Mutual Group she worked in the P&C reinsurance division of Zurich RE in Germany.

Ernesto Schirmacher is an Actuary with the Liberty Mutual Group in Boston. Mr. Schirmacher has a Ph.D. in mathematics from the University of Minnesota and is a Fellow of the Society of Actuaries (SOA). Before joining Liberty Mutual he worked in the Life & Health reinsurance division of Zurich RE in Germany.

Neeza Thandi is an Actuary with the Liberty Mutual Group in Boston. Ms. Thandi received her Ph.D. in mathematics from the University of Wisconsin-Madison. She is a Fellow of the Casualty Actuarial Society (CAS) and a member of the American Academy of Actuaries (AAA).