Pricing Excess of Loss Treaty with Loss Sensitive Features: An Exposure Rating Approach

Ana J. Mata, Ph.D*

Brian Fannin, ACAS

Mark A. Verheyen, FCAS

*Correspondence Author: ana.mata@cnare.com

Pricing Excess of Loss Treaty with Loss Sensitive Features: An Exposure Rating Approach

Summary

The main objective of this paper is to show how the use of actuarial models and techniques may make a significant contribution when pricing reinsurance.

We focus our attention on treaty excess of loss reinsurance pricing which is one of the most complex types of reinsurance since several components need to be taken into account: primary policy limits and deductibles, multiple lines of business covered by the same contract and loss sensitive features that vary with the loss experience of the treaty. When the treaty includes loss sensitive features not only are the losses random variables but also the premium and expenses become random variables that depend on the aggregate losses. Therefore the profitability of the risk can only be assessed in terms of expected values of these features.

The objective of the paper is not to develop methods to estimate the expected losses to the treaty but rather to take a selected expected loss (calculated based on exposure and experience methods) and develop an aggregate loss distribution based on a severity distribution that allows us to incorporate all the characteristics of the treaty, i.e. policy limits and deductibles, multiple lines of business and reinsurance layer and attachment. The severity curve is developed based on an exposure rating approach. We compare our approach with other commonly used approaches in practice and show the differences in the results. Worked examples based on practical cases are shown in Section 4.

1 Pricing reinsurance

Reinsurance markets have traditionally been driven by underwriting considerations and therefore actuaries have had very limited involvement in the pricing process. In the USA the role of actuaries in reinsurance pricing has been more clearly defined than in The London Market where only recently actuaries have become more involved in implementing actuarial techniques and models as underwriting tools.

With the development of more sophisticated insurance and reinsurance products the role of actuaries has become a key part of reinsurance pricing, risk management and profitability assessment. It is our objective in this paper to show cases of practical relevance in pricing reinsurance where the use of actuarial techniques can make a significant contribution in pricing and profitability assessments.

What does reinsurance pricing involve? A reinsurance program is often supported by a group of reinsurers each of which undertakes a share of the risk. The "lead" reinsurer is the one that undertakes the largest share of the risk and therefore takes the lead as to which rate must be charged to the ceding company. The lead reinsurer also commonly establishes the terms of the reinsurance contract. Other reinsurers base their pricing analysis on these rates and terms and decide whether they are willing to support the program based on underwriting targets and profitability.

In this paper we focus on reinsurance treaty pricing where the reinsurance contract covers a portfolio of policies each of which has a different policy limit and deductible. We will also allow for treaties that cover multiple lines of business, for example auto and general liability.

What are the components of the reinsurance price? For an extended review of reinsurance pricing and its components see Patrik (1998). The cost of reinsurance is divided into three basic components:

1. Loss cost: this term refers to the expected aggregate losses that would arise from the reinsurance contract during the treaty period.

2. Premium: if the contract is based on a fixed premium rate then this rate must be

determined based on experience, exposure and market benchmarks. In some types of reinsurance treaties the premium varies depending on the total aggregate losses during the treaty period. There are various types of loss sensitive premium. For example, in casualty reinsurance it is common to receive a provisional premium in advance and after a certain period of time the experience is assessed and the premium is adjusted according to the terms of the treaty. In other lines of business there is a limited number of losses or reinstatements covered by the treaty and after each loss an extra premium may be charged in order to reinstate the coverage. For more details on the mathematics of pricing excess of loss with reinstatements see, for example, Mata (2000).

3. Expenses: these could be pre-determined or loss sensitive expenses.

    (a) Pre-determined expenses are usually a fixed percentage of the premium. These include for example ceding commission paid to the ceding company to compensate for their operating expenses. Brokerage may also be a fixed percentage of the premium.

    (b) Loss sensitive expenses include profit based expenses such as profit commissions where reinsurers give back part of the profits to the ceding company after a load for expenses.

Often a profit margin is taken into account when pricing a reinsurance contract. Profits can be added in terms of a fixed load or through modelling of discounted cashflow and investment income. The profit margin added to the treaty depends on company specific underwriting targets.

Table 1 summarises the most common loss sensitive features in treaty reinsurance. This list is not exhaustive, and any combination of features may be included in any given treaty. A more detailed review of loss sensitive premium rating such as margin plus and swing rating see Strain (1987). For the mathematical aspects of pricing excess of loss with reinstatements and aggregate deductibles see Mata (2000).

Feature	Variable component	Description
Margin Plus	Premium	Provisional Premium paid adjusted depending on losses. Premium = Provisional + margin = losses plus loading subject to a maximum and minimum premium.
Swing Rating	Premium	Premium varies with Loss Ratio. Calculated as loss + load subject to maximum and minimum
Profit Commission (PC)	Expenses	A share of the profit is given back to the cedant after allowing for reinsurer's expenses
Loss Corridor	Loss	The ceding company retains part of the risk that starts at a pre-determined value of the loss ratio and for a pre-determined width
Reinstatements	Premium and Losses	Limits the number of total losses covered by the contract. For paid reinstatements then extra premium is payable to reinstate the layer limit.
Annual Aggregate Deductible (AAD)	Loss	The ceding company retains the first D losses in aggregate

Table 1. A summary of loss sensitive features.

If the terms of the treaty include loss sensitive features, it is clear that the premium and expenses become random variables that are functions of the aggregate loss. Hence, at the time of pricing we cannot estimate their value. However based on the aggregate losses we can estimate their expected value. It is therefore very important to be able to estimate an appropriate aggregate loss distribution function that can be used to estimate the expected premium income and expected expenses in order to assess the profitability of the reinsurance contract.

There are several methods based on experience and exposure widely used in order to estimate the expected loss cost of a reinsurance contract. Some commonly used methods are: burning cost, curve fitting, experience rating and exposure rating. Sanders (1995) presents a thorough review of various methods used for pricing in The London Market and Patrik (1998) presents a detailed review of the mathematics of reinsurance pricing.

It is not the objective of this paper to develop a new method of estimating reinsurance losses but rather to develop a methodology to estimate the expected value of those loss sensitive features (premium and commissions) based on the aggregate loss distribution for the layer. The method allows us to model aggregate loss distributions that incorporate all the characteristics of the treaty that are taken into account when estimating the loss cost. Such characteristics could include mixtures of policy limits and deductibles, lines of business covered in the contract and layer size and attachment.

In the rest of the paper, unless otherwise stated, we will follow the notation and definitions given in Klugman, Panjer and Willmot (1998).

Definition 1 **The severity** *can be either the loss or amount paid random variable for a single event. We denote X the severity for a single event and we refer to its expected value and probability distribution as the expected severity and severity distribution respectively.*

Definition 2 **The frequency** *is the number of total random losses. We denote this number of losses N and we refer to its expected value as the expected severity and its probability function as the frequency distribution.*

The various approaches we present in this paper to model aggregate loss distributions are based on the so-called collective risk model:

$$S = X_1 + \cdots + X_N,$$

where X_i is the severity for the ith event and N is the frequency. The X_is are assumed iid and independent of N.

The paper is organised as follows: since our approach is to develop an exposure based severity distribution, Section 2 presents an overview of the ideas behind the exposure rating method. In Section 3 we describe different approaches that can be taken to estimate frequency and severity distributions in order to compute aggregate losses and in Section 3.3 we describe our methodology to estimate an exposure base severity distribution that incorporates all the features of the treaty.

2 A review of the exposure rating method

The exposure rating method is widely used in the reinsurance industry as a method to estimate the expected loss cost of a reinsurance contract. This method requires the use of a severity distribution to model the loss cost per claim. This distribution can be fitted to historical losses for the specific risk. In the absence of adequate data, benchmark severity distributions fitted to industry wide data by line of business might be used. In the USA the ISO uses industry wide losses by line of business to produce benchmark severity distributions.

Industry benchmark loss distributions are particularly useful when there are no historical data available or when the past experience may not be a good representation of the future experience and exposure.

We now introduce some definitions and notation that will be used in the rest of the paper.

Definition 3 *Let X be a random variable with probability density function $f_X(x)$ and cumulative distribution function $F_X(x)$. The Limited Expected Value of X up to a limit m, i.e.* $\min(X,m)$*, is given by*

$$E[X \wedge m] = \int_0^m x f_X(x)dx + m(1 - F_X(m)) = \int_0^m (1 - F_X(x))dx. \qquad (1)$$

See Klugman, Panjer and Willmot (1998). Note that the notation $X \wedge m$ stands for $\min(X,m)$.

Basic ingredients for the exposure method in treaty reinsurance:

1. **Premium:** is the base premium written by the ceding company subject to the treaty. Typically this premium is split by policy limits and deductibles that the ceding company writes. In reinsurance jargon, the distribution of policy limits and deductibles is referred to as "the limits profile".

2. **Ground-up expected loss ratio (FGU ELR):** is the expected loss ratio for the primary insurer (in total or by business segment) during the treaty period. This expected loss ratio is estimated using historical development triangles. The losses are trended to allow for changes in real value of future losses, e.g. the effect of inflation, and then projected to ultimate. Primary rate changes are used to adjust historical written premium so that it is comparable at present rates. This process is called "on-levelling". With ultimate trended losses and "on-level" premium we can estimate the ground-up expected loss ratio. Ideally this process must be carried out for each line of business ceding to the treaty. See McClenahan (1998).

3. **Severity distribution:** risk specific or industry benchmark.

4. **The reinsurance layer:** in the remaining of the paper we consider a generic layer $\ell \; xs \; m$.

The expected losses in the reinsurance layer are calculated for each combination of policy limit and underlying deductible and then added together. The following example illustrates the methodology.

Example 1: Assume the ceding company writes two types of policies:

Deductible	**Limit**	**Total Premium**	**% Premium**
\$10,000	\$250,000	\$250,000	25%
\$25,000	\$500,000	\$750,000	75%

Hence the subject premium is \$1,000,000. Assume that the ground-up loss ratio is expected to be 75% and that the reinsurer underwrites the layer \$350,000 xs \$150,000. Then policies with \$250,000 in limit only cede up to \$100,000 per claim to the layer.

If X if a random variable representing the severity from ground-up, then the total expected aggregate losses to the layer from the first type of policy with \$250,000 in limit are:

$$250,000 * 0.75 * \left(\frac{E[X \wedge (250,000 + 10,000)] - E[X \wedge (150,000 + 10,000)]}{E[X \wedge (250,000 + 10,000)] - E[X \wedge 10,000]} \right)$$

and for policies with limits of \$500,000 are:

$$750,000 * 0.75 * \left(\frac{E[X \wedge (500,000 + 25,000)] - E[X \wedge (150,000 + 25,000)]}{E[X \wedge (500,000 + 25,000)] - E[X \wedge 25,000]} \right).$$

where the expected value is calculated using equation (1) with the corresponding severity distribution function.

The two values calculated above are added and the total represents the aggregate expected losses in the layer during the treaty period.

Generalizing the ideas of the above example, the total losses by combination of policy limit (PL_k) and deductible (d_k) are given by:

$E[Losses_k] =$

$$(\text{SP}_k) * (\text{FGU ELR}) * \left(\frac{E[X \wedge \min(PL_k + d_k, \ell + m + d_k)] - E[X \wedge \min(PL_k + d_k, m + d_k)]}{E[X \wedge (PL_k + d_k)] - E[X \wedge d_k]} \right) \quad (2)$$

where SP_k is the subject premium for each combination of policy limit and deductible. The total expected aggregate losses are:

$$E[Losses] = \sum_k E[Losses_k].$$

The exposure method must be carried out by line of business ceding to the treaty to obtain expected aggregate losses by line of business.

We have assumed so far that the severity distribution represents the distribution function of a single loss before any deductibles. In practice losses used to fit severity distributions may be in excess of deductibles. In this case the distribution would be in excess of primary deductibles and appropriate adjustments should be made to formula (2).

2.1 Using the exposure method to estimate expected frequency

In order to model aggregate loss distributions we are interested in estimating the expected frequency and the expected severity to the layer. The exposure method outlined above gives an expected value of the aggregate losses in the reinsurance layer but it can also be used to estimate an expected frequency to the layer.

The following notation will be used in the rest of the paper: consider the layer $\ell \ xs \ m$ then

$$\text{Loss Cost} = E[S] = E[N_m]E[X_m],$$

where S represents the aggregate losses to the layer, N_m represents the number of losses in excess of the attachment m and X_m represents the non zero losses to the layer, i.e. $X_m = (\min(X - m, \ell) \mid X > m)$. Note that X_m is the severity to the layer conditional on the ground up claim being greater than m.

Result 1 *Let X be a random variable representing the ground-up claim size for a single event with pdf $f_X(x)$ and cdf $F_X(x)$. Let us consider a small layer of size h and attachment m. The loss to the layer given that $X > m$ is given by $X_m = \min(X - m, h)$. Then*

$$E[X_m] = E[\min(h, X - m)] \approx h.$$

Proof. The conditional severity distribution of a loss in the layer has density and distribution functions

$$f_{X_m}(x) = \begin{cases} \frac{f_X(x+m)}{1-F_X(m)} & \text{for } 0 < x < h \\ \frac{1-F_X(m+h)}{1-F_X(m)} & \text{for } x = h \end{cases} \qquad \text{and} \qquad F_{X_m}(x) = \frac{F_X(x+m) - F_X(m)}{1 - F_X(m)}. \tag{3}$$

respectively, see Dickson and Waters (1992). Hence,

$$E[X_m] = \int_0^h \left(\frac{1 - F_X(x+m)}{1 - F_X(m)} \right) dx \approx h \left(\frac{1 - F_X(m)}{1 - F_X(m)} \right) = h$$

using (1) and the fact that in a small interval the integral is approximated by the area of a rectangle of base h and height $(1 - F_X(m))$. See Figure 1.

If we apply the exposure method for the layer 1 xs m we obtain

$$E[S] = E[N_m]E[X_m] \approx E[N_m] \times 1, \tag{4}$$

due to Result 1. Therefore the exposure method applied to the layer 1 xs m gives an expected frequency of losses in excess of the attachment m.

3 Calculating the input to calculate aggregate loss distribution

There are various approaches to calculate an aggregate loss distribution. If we are not interested in incorporating a frequency and severity distribution, then the easiest method

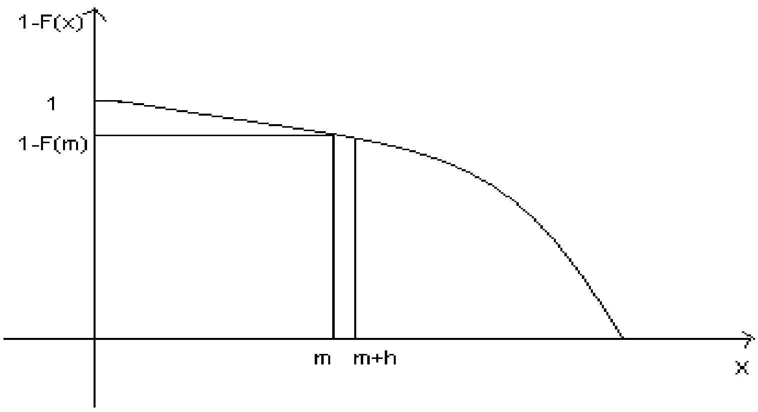

Figure 1: Losses in a small layer $h\ xs\ m$

is to fit a parametric distribution to the aggregate losses. In order to do so it is required to estimate the parameters of such a distribution. One of the easiest methods to estimate these parameters is by the method of matching moments, i.e. matching the expected loss cost and its variance to the mean and variance of the selected distribution.

If we are interested in calculating an aggregate loss distribution that allows for frequency and severity distributions there are various mathematical methods to compute aggregate loss distributions, such as the Panjer recursive algorithm. We present in Appendix A a summary of the most commonly used algorithms to compute aggregate losses given frequency and severity distributions. In order to implement any of these methods it is necessary to estimate suitable frequency and severity distributions.

The loss cost, estimated using a mixture of experience and exposure methods, provides an estimate of the expected aggregate losses to the layer. The loss cost is used as the mean of the aggregate loss distribution, then a severity distribution is chosen to estimate expected cost per claim in the layer (expected severity) and the expected frequency is chosen as the implied parameter between the loss cost and the expected severity.

The expected aggregate loss is not enough information to fit an aggregate loss distribution, more information is needed in order to estimate an aggregate loss distribution. For example, practical considerations to take into account when modelling aggregate loss distribution may be: how to estimate the variance of the expected loss cost, which frequency distribution should be used and, if we have a multi-line treaty which loss distribution should be used? Once these issues are addressed we would have all the input required in order to compute the aggregate loss distribution.

It is desirable that the aggregate loss distribution should also incorporate other features of the treaty under consideration, for example policy limits and deductibles, all lines of business, their distributions and the loss cost selected for each line of business ceding to the treaty.

We discuss below three methods to estimate the input to compute the aggregate loss distribution for a reinsurance treaty. In the rest of the paper we consider a generic layer $\ell \; xs \; m$ and we assume that there are $j = 1, \ldots, n$ lines of business covered by the treaty.

3.1 Method 1: Fitting a parametric curve to the aggregate loss distribution

A very common approach used in practice is to fit a parametric curve to the aggregate losses using the mean and variance of the aggregate losses to the layer. If S represents the aggregate losses to the layer we have:

$$E[S] = \sum_{j=1}^{n} \text{Loss Cost}_j \qquad \text{and} \qquad Var(S) = CV * E[S],$$

where Loss Cost_j is the estimated loss cost for the jth line of business and CV is the coefficient of variation. The CV is a subjective component in this approach and it is typically selected depending on the line of business. However, if a Poisson distribution is used as the frequency distribution no extra assumptions need to be made about the variance of the aggregate losses.

If we use the Poisson distribution with expected frequency λ as the frequency distribution, the mean and variance are directly calculated since

$$E[S] = \lambda E[X_m] \qquad \text{and} \qquad Var(S) = \lambda E[X_m^2],$$

where X_m represents the conditional loss cost for each claim in excess of the attachment m. The expected values above are calculated using a suitable severity distribution function. If there is only one line of business then the expected values can be calculated using the loss distribution for that line of business (benchmark or risk specific). If there are various lines of business then perhaps a predominant loss distribution should be selected, for example the loss distribution for the line of business with the highest exposure.

Finally, once $E[X_m]$ is calculated λ is given by

$$\lambda = \frac{\text{Loss Cost}}{E[X_A]}.$$

Therefore the variance of the aggregate losses $Var(S)$ can also be calculated and a parametric curve can be fitted to the aggregate losses to the layer. Common parametric distributions used to approximate aggregate loss distributions are the lognormal and the gamma distributions.

If we want to use the Negative Binomial distribution as the frequency distribution an additional assumption should be made. We describe in Appendix B an approach to calculating the CV or $Var(S)$ using the Negative Binomial distribution as the frequency distribution for the layer.

A disadvantage of fitting a parametric curve to the aggregate losses to the layer is that it does not take into account the probability of having zero losses in the layer or the probability mass at the layer limit. This problem is overcome by the next approach.

3.2 Method 2: Using benchmark distributions as the severity distributions

We discussed in Section 2 the need of a loss distribution by line of business in order to use the exposure rating method to estimate expected losses to the layer. These loss distributions can be used as the severity distribution of losses in the layer.

If X represents the ground-up loss per claim, the expected loss to the layer $\ell\ xs\ m$ for the jth line of business can be estimated as:

$$\text{Expected Severity}_j = E^j[X_m] = E^j[\min(X - m, \ell)] = \frac{E^j[X \wedge (m + \ell)] - E^j[X \wedge m]}{1 - F_X^j(m)},$$

where $F_X^j(x)$ is the loss distribution function for the jth line of business and $E^j[]$ is calculated as in (1) with the corresponding pdf. Even though this method does not take into account policy limits and deductibles, it does provide an easy method to estimate the expected severity to the layer by line of business.

Expected frequency in excess of the attachment m by line of business can be calculated as the implied factor between the expected loss cost and the expected severity:

$$\text{Expected frequency}_j = E^j[N_m] = \frac{\text{Loss Cost}_j}{E^j[X_m]}. \tag{5}$$

To compute the aggregate loss distribution we also need to estimate a frequency distribution by line of business. The easiest approach is to use the Poisson distribution by line of business with parameter given by $\lambda_m^j = E^j[N_m]$ calculated as in (5).

Fitting a Poisson distribution to the number of claims only requires estimation of one parameter, however this distribution may not be suitable for all lines of business. In Appendix B we explain how we could fit a Negative Binomial distribution as the frequency distribution.

For a multi-line treaty we need to mix the severity distributions by line of business to obtain an overall claim size distribution for losses in the layer. If $f_m^j(x)$ is the conditional probability density function for losses in excess of m as in (3) for the jth line of business and λ_m^j is the implied expected frequency as in (5), then assuming independence between lines of business the overall probability density function is obtained as:

$$f_{X_m}(x) = \sum_{j=1}^{n} \frac{\lambda_m^j}{\lambda_m} f_{X_m}^j(x) \qquad \text{for } 0 < x \leq \ell,$$

where $\lambda_m = \sum_{j=1}^{n} \lambda_m^j$ and n is the number of lines considered in the treaty.

Most of the algorithms described in Appendix A to compute aggregate loss distributions require the probability function to be discretised. There are various methods to do so and we explain in Appendix A the method of matching the mean to discretise distribution functions.

Although the method to estimate the input distributions to compute aggregate loss distributions overcomes the problem of the probability mass at zero and at layer limit, it does not take into account the distribution of policy limits and deductibles that the primary insurer writes. In other words, it assumes that all losses in the layer could reach the maximum limit and that the impact on the severity of any deductible is negligible. This may overestimate the expected severity to the layer, in particular when primary policies have large deductibles. Considering policy limits and deductibles is of vital importance in treaty reinsurance, however this is not a problem in facultative excess of loss where the reinsurer only covers one primary risk. Therefore, the method presented in this section could be applicable when modelling aggregate losses for facultative excess of loss contracts.

The method we propose in the next section overcomes the problem of taking into account the mixture of policy limits and deductibles.

3.3 Method 3: The exposure based severity curve

In this section we use the exposure method described in Section 2.1 to develop a severity loss distribution for losses in the reinsurance layer that takes into account all combinations of policy limits, deductibles and lines of business.

We develop a discrete probability function for the severity distribution of losses in the layer $\ell \; xs \; m$ which can then be used directly to compute the aggregate loss distribution. **Warning:** for convenience in this section we use the letter "λ" to represent the estimated frequency at various attachment points. This should be confused with the expected frequency for a Poisson distribution.

Our approach is based on the following property: If λ represents the expected number of losses from ground-up and $F_X(x)$ represents the distribution function of the ground up losses, then the expected number of losses in excess of the attachment m is given by:

$$\lambda_m = \lambda(1 - F_X(m)) \tag{6}$$

see, for example, Dickson and Waters (1992). Therefore it follows that $\lambda = \frac{\lambda_m}{1-F_X(m)}$.

From formula (6) it is clear that if we can estimate the expected frequency at various points along the possible values of losses in the layer then we can estimate the distribution function.

Since we are interested in the distribution function of losses in excess of the attachment m we divide the layer in small intervals of size h as follows:

$$\begin{array}{ccc} h & xs & m \\ h & xs & m+h \\ h & xs & m+2h \\ \vdots & & \vdots \\ h & xs & \ell+m-h \end{array}$$

Note that there are $\frac{\ell}{h}$ sub-layers.

Using the exposure method to calculate expected frequency in a layer as described in Section 2.1, we calculate the expected frequency at each attachment point for each sub-layer for the jth line of business as follows:

$$\begin{array}{cc} \text{Attachment} & \text{Expected frequency} \\ m & \lambda^j_m \\ m+h & \lambda^j_{m+h} \\ m+2h & \lambda^j_{m+2h} \\ \vdots & \vdots \\ \ell+m-h & \lambda^j_{\ell+m-h} \end{array} \tag{7}$$

Since the expected frequency in the above table is calculated using the exposure method it takes into account all combinations of deductibles and policy limits as described in Section 2.

From the expected severity at various attachment points we can calculate the treaty loss distribution by line of business due to the following property:

$$\lambda^j_{m+rh} = \lambda(1 - G^j(m + rh)) = \frac{\lambda^j_m}{(1 - G^j(m))}(1 - G^j(m + rh)). \qquad \text{for } r = 1, 2, \dots, \frac{\ell}{h},$$

where $G^j(x)$ is a "blended" loss distribution for the jth line of business that takes into account all policy limits and deductibles. Therefore $G^j(x)$ is not the same distribution as the benchmark loss distribution used to estimate the expected frequencies at each attachment in (7).

Since we are interested in estimating a conditional distribution of losses in excess of the attachment m we use as our base the expected frequency in excess of m to obtain the following result:

$$P^j(X - m > rh \mid X > m) = \frac{\lambda^j_{m+rh}}{\lambda^j_m} = \frac{1 - G^j_{X_m}(m + rh)}{1 - G^j_{X_m}(m)} \qquad \text{for } r = 1, 2, \dots, \frac{\ell}{h}.$$

Before reading further, if the formulae below look technical at first sight we recommend that the reader skip the details and see the summary presented in (10) for a quick overview of the methods and then return to the mathematical details.

From (7) we can obtain the "blended" distribution of losses in excess of the layer for the jth line of business as follows:

$$G^j_{X_m}(rh) = P^j(X - m \le rh \mid X > m) = \begin{cases} 0 & \text{for } r = 0 \\ 1 - \frac{\lambda^j_{m+rh}}{\lambda^j_m} & \text{for } r = 1, 2, \dots, \frac{\ell}{h} - 1 \\ 1 & \text{for } r = \frac{\ell}{h} \end{cases} \tag{8}$$

and therefore the "blended" probability density function is obtained as follows:

$$p^j_{X_m}(rh) = P^j(X_m = rh) = \begin{cases} 0 & \text{for } r = 0 \\ G^j_m(rh) - G^j_m((r-1)h) & \text{for } r = 1, 2, \ldots, \frac{\ell}{h} - 1 \\ \\ 1 - \displaystyle\sum_{r=0}^{\ell/h-1} p^j_m(rh) & \text{for } r = \frac{\ell}{h} \end{cases} \tag{9}$$

Below is a summary of how to calculate the survival distribution $S^j_{X_m}(x)$, the cumulative distribution $G^j_{X_m}(x)$ and the probability distribution $p^j_{X_m}(x)$ conditional in excess of the attachment m given the frequency at various attachment points as in (7):

x	$S^j_{X_m}(x) = P(X_m > x)$	$G^j_{X_m}(x) = P(X_m \leq x)$	$p^j_{X_m}(x) = P(X_m = x)$
0	1	0	0
h	$\frac{\lambda^j_{m+h}}{\lambda^j_m}$	$1 - \frac{\lambda^j_{m+h}}{\lambda^j_m}$	$G^j_{X_m}(h) - G^j_{X_m}(0)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
rh	$\frac{\lambda^j_{m+rh}}{\lambda^j_m}$	$1 - \frac{\lambda^j_{m+rh}}{\lambda^j_m}$	$G^j_{X_m}(rh) - G^j_{X_m}((r-1)h)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
ℓ	0	1	$1 - \sum_{r=0}^{\ell/h-1} p^j_m(rh)$

(10)

for $r = 1, 2, 3, \ldots, \frac{\ell}{h} - 1$.

Using the severity curve for the jth line of business given in (9), the expected severity in the layer for the jth line of business is calculated as:

$$E^j[X_m] = \sum_{r=0}^{\ell/h} rh p^j_{X_m}(rh),$$

and therefore the implied expected frequency to the layer for the jth line is:

$$E^j[N_m] = \frac{\text{Loss Cost}_j}{E^j[X_m]}.$$

Note that if the selected expected loss cost is different than the expected loss cost given by the exposure method, then the implied frequency would be different than the expected frequency in excess of the attachment m calculated in (7).

Finally we need to mix all the probability density functions by line of business to obtain an overall severity distribution. Assuming independence between lines of business we use the same methodology as in Section 3.2 as follows:

$$p_{X_m}(x) = \sum_{j=1}^{n} \frac{E^j[N_m]}{E[N_m]} p^j_{X_m}(x) \qquad \text{for } x = 0, h, 2h, \ldots, \ell, \tag{11}$$

where $E[N_m] = \sum_{j=1}^{m} E^j[N_m]$ and $p^j_{X_m}(x)$ is as in (9).

For example, using $p_{X_m}(x)$ as in (11) and a Poisson distribution with expected frequency $E[N_m]$ we could use any of the algorithms described in Appendix A. For example, the Panjer recursive algorithm can be easily implemented in this case.

It is worth pointing out that although we have assumed that losses arising from different lines of business ceding to the same layer are independent that this assumption may not be realistic in practice. The assumption of independence makes the mathematics easier. However, further research must be carried out in order to model risk specific dependencies between lines of business since by ignoring dependencies one can underestimate the overall risk.

4 Worked examples

4.1 Casualty example: Professional Liability

Assume a ceding company wants to cede part of their risk that consist of two lines of business Lawyers liability and Errors and Omissions (*E&O*). For each line of business they write the following limits and deductibles:

	Lawyers		E&O	
Deductible	\$10,000	\$25,000	\$50,000	\$50,000
Limit	\$750,000	\$1,000,000	\$1,500,000	\$2,000,000
Premium	\$1,000,000	\$2,000,000	\$2,000,000	\$3,000,000
FGU LR	65%	65%	75%	75%

Table 2. Limits Profile

We assume that the ground up loss distribution for the lawyers business is a lognormal with parameters $\mu = 8$ and $\sigma = 2.5$ and the E&O business follows a lognormal with parameters $\mu = 9$ and $\sigma = 3$. The reinsurer underwrites the layers \$500,000 xs \$500,000 and \$1m xs \$1m. Note that the lawyers policies only expose the first layer- the first policy up to \$250,000 per claim. The E&O policies expose both layers- the first layers up to maximum limit per claim and the \$1.5m policy limit only exposes the second layers up to \$500,000.

The total subject premium is \$8m and of this we assume that \$7.2m is for the first million in limits and \$800,000 for the second million. This split is based on the theory of Increased Limits Factors (ILF) which depend on the ceding company's rating plan. A good review of how to use increase limits factors in reinsurance pricing is giving in Patrik (1998).

The terms for this layers are:

1. The first layer is margin plus rated with a provisional premium of 12.5% of the premium for the first million, a minimum premium of 7% and a maximum premium

of 18% with a load of 107.5%. See Section 1 and Strain (1987). The treaty also includes a Profit Commission of 15% after 20% for the reinsurer's expenses. Brokerage is assumed 10% on the provisional premium.

2. The second layer is cessions rated, i.e. all of the premium allocated to this layer is given to reinsurers since they are taking all the risk in excess of $1m. The reinsurer pays 15% ceding commission to the ceding company to compensate them for their operating expenses plus a profit commission of 15% after 20% for reinsurer's expenses. Brokerage 10% on gross.

Note that for the first layer the premium is loss dependent since depending on the experience of the treaty the premium would vary between 7% and 18%. In both layers the profit commission payable after expiration is also a random variable that depends on the aggregate losses during the treaty year.

We have selected a loss cost of $750,000 and $375,000 for each layer respectively. We do not discuss how to estimate the loss cost for the layer. This would depend on data available and experience of the risk being priced, see Patrik (1998) and Sanders (1995).

We divided the layer in units of $2,500 and used methods 2 and 3 described in Section 3 to calculate the discretised severity distribution for losses in the layers. Figures 2 and 3 show the severity cumulative distributions for losses in excess of $500,000 and $1m respectively. Note in Figure 2 that the cdf using method 3 has a probability jump at $250,000 due to the policy limit of $750,000 for the lawyers business and a probability mass at the layer limit while the cdf estimated using method 2 only has the probability mass at the limit loss since it does not take into account the policy limits and therefore assumes each policy fully exposes the layer. In Figure 3 the cdf is calculated only using the E&O policies since the lawyers polices do not have limits higher than $1m. Note in Figure 3 that method 3 has a probability jump at $500,000 caused by the policy limit of $1.5m which is not taken into account when we use method 2.

Using this severity distribution, we calculate the expected severity of losses to the layers and the implied expected frequency. Table 2 shows the Loss Cost, expected severity and

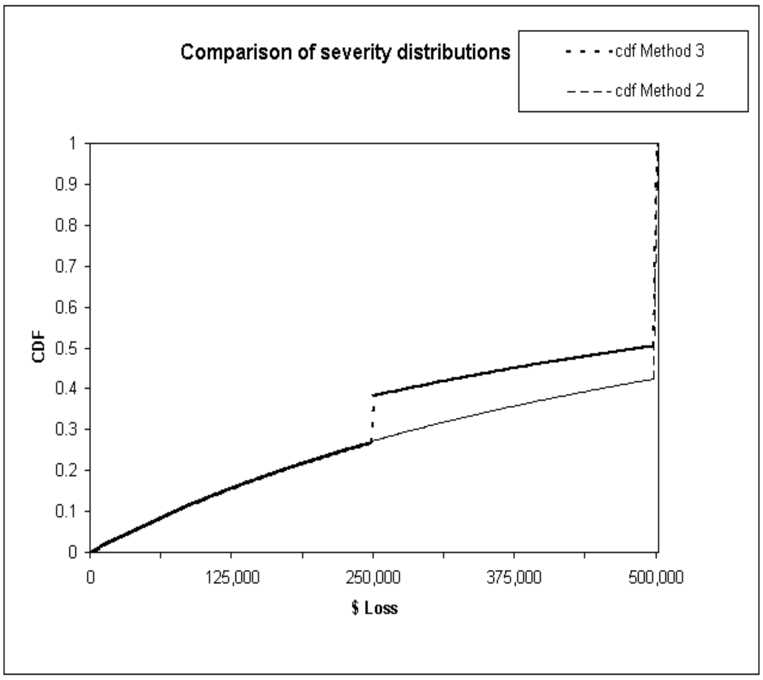

Figure 2: Severity distribution function \$500,000 *xs* \$500,000

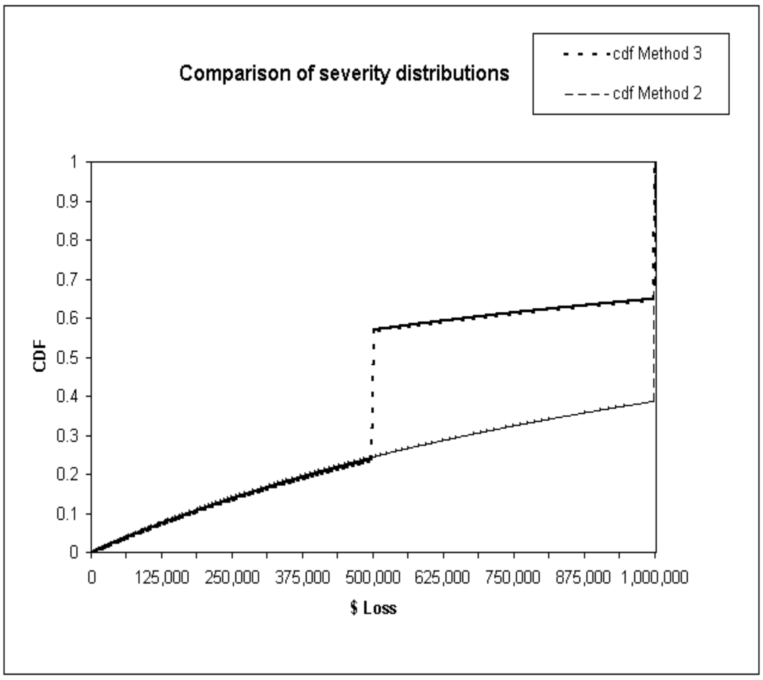

Figure 3: Severity distribution function \$1m *xs* \$1m

frequency for both layers using methods 2 and 3 as in Section 3.

		\$500,000 *xs* \$500,000	\$1m *xs* \$1m
Loss Cost		\$750,000	\$375,000
Method 2	Severity	\$373,134	\$771,549
"benchmark severity"	Frequency	2.01	0.49
Method 3	Severity	351,063	\$628,809
"exposure severity"	Frequency	2.14	0.60

Table 3. Expected Severity and Frequency using Methods 2 and 3 in Section 3

Note that expected severities are overestimated by using method 2 since policy limits are not taken into account and it is assumed that every underlying policy has a potential of a full loss to the layer. Using the expected frequency and a Variance Multiplier of 2 we fitted a Negative Binomial using the methodology described in Appendix B for both methods 2 and 3. Using the discretised severity distributions and the Negative Binomial distributions we used Panjer's recursive algorithm to compute the aggregate loss distribution at multiples of 2,500. For comparative purposes we also fitted a lognormal distribution using the method of moments and the following relationships

$$E[S] = E[N_m]E[X_m] \qquad \text{and} \qquad Var(S) = E[N_m]Var(X_m) + (E[X_m])^2\, Var(N_m),$$

where S represents the aggregate losses, X_m the severity and N_m the frequency to the layer as defined above. In our example above $E[X_m]$ is calculated using the discretised distributions and $E[N_m]$ is the expected value of the Negative Binomial. See Dickson and Waters (1992).

Figures 4 and 5 show the probability density function and the cumulative distribution function of the aggregate losses to the first layer using all three methods described in Section 3. Note in Figure 4 the "spikes" of probability mass at multiples of the policy limits and layer limits. As we discussed above, using method 2 only shows probability

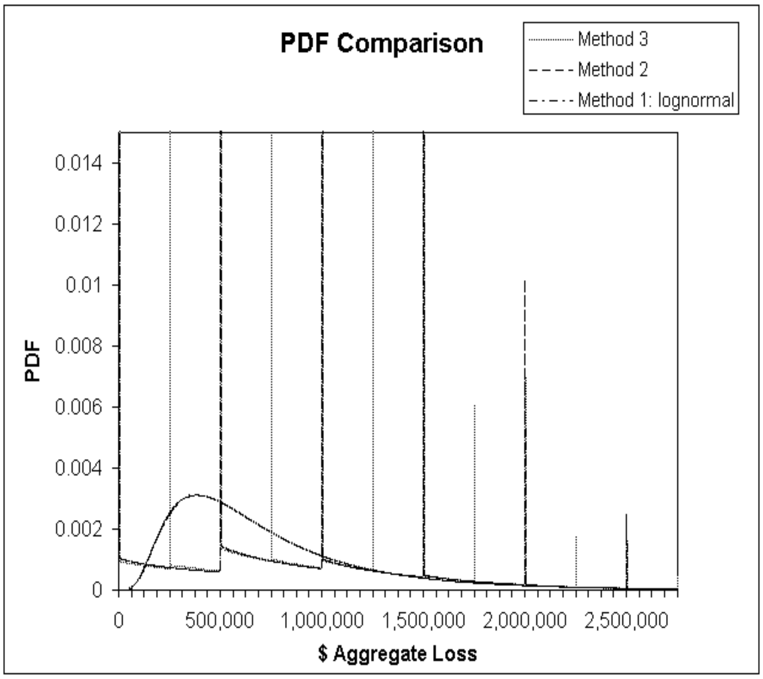

Figure 4: Probability density function of aggregate losses $500,000 *xs* $500,000

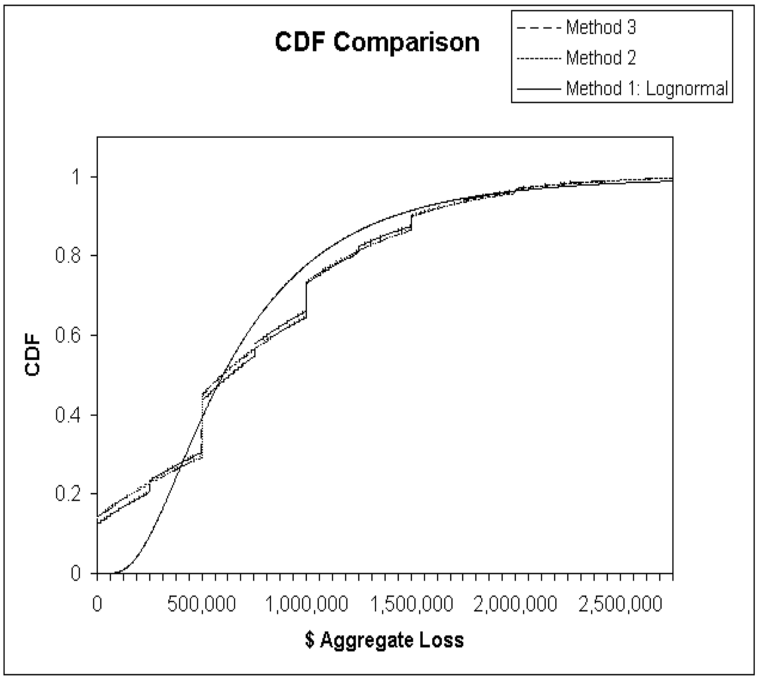

Figure 5: Aggregate loss distribution $500,000 *xs* $500,000

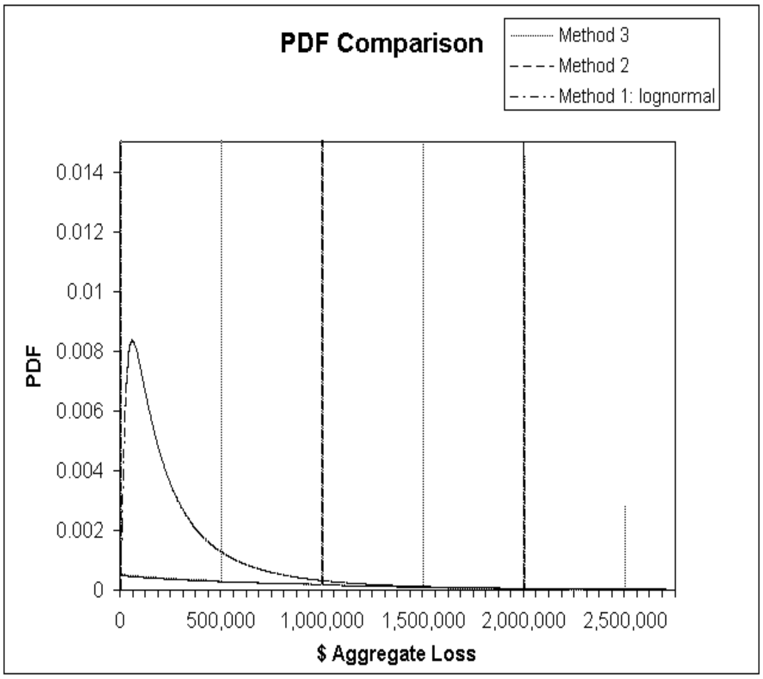

Figure 6: Probability density function of aggregate losses $1m xs $1m

mass at multiples of the layer limit while method 2 shows probability mass at multiples of $250,000 and $500,000 due to the policy limit of $750,000. Also note that the lognormal distribution is a smooth function that does not take into account the probability of having no losses or the full limit loss.

Figures 6 and 7 show the probability density function of the aggregate losses for the second layer using the three methods described above. In Figure 7 the difference between a lognormal and the distribution given by methods 2 and 3 are more noticeable for small losses. The lognormal tails off faster and therefore for large losses all distributions are very similar. We will see below that even though the lognormal does not seem to be a good fit, the overall effect in the expected value calculations is balanced.

We discussed above that the premium and the profit commissions are random variables whose value can not be calculated precisely at the moment of pricing. However, since they are functions of the aggregate loss, given the aggregate loss distribution we can calculate their expected value. For example, for the margin plus rating in the first layer we can estimate the extra premium to be received from the ceding company and therefore estimate a combined ratio for the treaty.

Table 4 below shows the expected value calculations for the first layer. The margin plus premium and the profit commission are expected values calculated using the three aggregate loss distributions shown in Figure 5. It is worth pointing out that these values are not discounted to allow for payment patterns and cashflows.

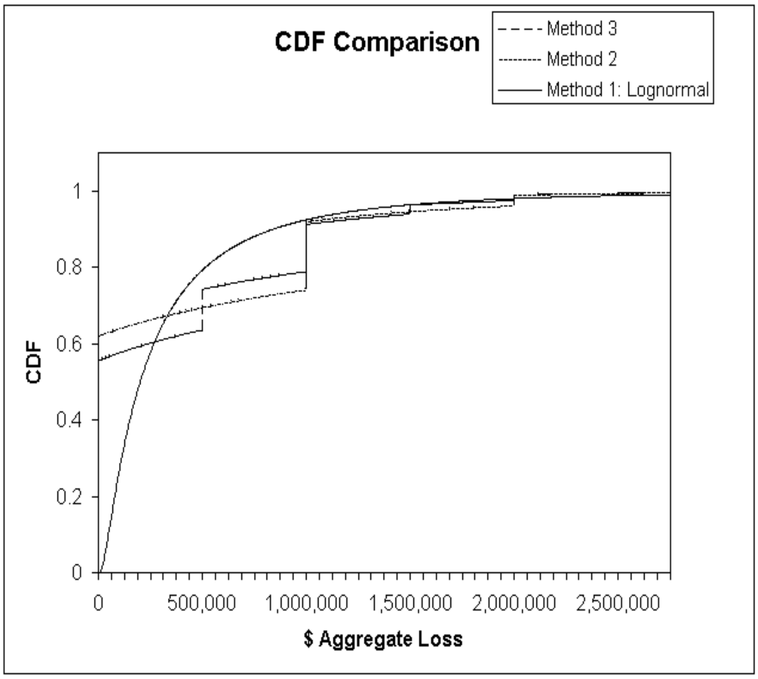

Figure 7: Aggregate loss distribution \$1m xs \$1m

	Method 1:lognormal		Method 2:benchmark		Method 3:exposure	
	Amount	% Premium	Amount	% Premium	Amount	% Premium
Prov. prem.	900,000		900,000		900,000	
Margin Plus prem	222,739		167,989		174,238	
Tot. prem	1,122,739	100%	1,068,041	100%	1,074,238	100%
Tot. Loss	750,000	66.8%	750,000	70.23%	750,000	69.82%
Profit Comm.	25,659	2.29%	32,930	3.07%	32,062	2.98%
Brokerage	90,000	8.02%	90,000	8.43%	90,000	8.38%
Marginal CR	865,659	77.1%	872,930	81.73%	872,062	81.18%

Table 4. Expected results \$500,000 xs \$500,000

Note that the results obtained by using the lognormal approximation are significantly lower than those obtained using the aggregate loss distributions allowing for frequency and severity distributions. The fact that the lognormal distribution does not allow for the probability of zero losses overestimates the extra premium due from the margin plus and it underestimates the profit commission payable to the ceding company. In this example the difference is only 4%, however, for layers with tighter terms, a 4% difference may make the treaty appear unprofitable depending on underwriting targets. We also observe that method 2 slightly underestimates the margin plus premium and overestimates the profit commission. As discussed above, the second method overestimates the expected severity since it does not take into account policy limits that would only partially expose the layer.

In Table 5 we show the Expected Value for the second layer. We note that only small differences are shown between the three methods. We saw in Figure 7 that the lognormal was a poor fit for small values of the aggregate loss. It underestimates the probability of small values but it overestimates the probability of larger values causing a balancing effect. Furthermore, all three distributions are very close in the tail causing a more significant balancing effect. As discussed above, the second method gives a higher estimate since it overestimates the severity to the layer.

	Method 1:lognormal		Method 2:benchmark		Method 3:exposure	
	Amount	% Premium	Amount	% Premium	Amount	% Premium
Tot. prem	800,000	100%	800,000	100%	800,000	100%
Tot. Loss	375,000	46.88%	375,000	46.88%	375,000	46.88%
Ceding Comm.	120,000	15%	120,000	15%	120,000	15%
Profit Comm.	58,951	4.87%	51,473	6.43%%	46,896	5.86%
Brokerage	80,000	10%	80,000	10%	80,000	10%
Marginal CR	661,951	76.74%	626,473	78.31%	621,896	77.74%

Table 5. Expected results \$1m xs \$1m.

4.2 Standard lines example: property and general liability

Assume that an insurance company wants to cede part of its risk consisting of two lines of business: commercial property and general liability. The policy limits profile, deductibles, subject premium and expected ground up loss ratio are given in the following table:

	Commercial Property		General Liability	
Deductible	\$250	\$250	\$10,000	\$10,000
Limit	\$125,000	\$500,000	\$1,000,000	\$2,000,000
Premium	\$4,000,000	\$6,000,000	\$12,000,000	\$8,000,000
FGU LR	50%	50%	75%	75%

Table 6. Limits Profile

We assume the severity distribution of each line of business follows a lognormal distribution with parameters $\mu = 12$ and $\sigma = 0.1$ and $\mu = 11$ and $\sigma = 0.5$ respectively.

The reinsurer underwrites excess of loss for the layers $\$300,000\ xs\ \$200,000$. The initial premium for this layer is 2.04% of the subject premium of \$30m. The treaty includes 2 paid reinstatements at 100%. For details on the mathematics of reinstatement premiums see Mata (2000). Brokerage is 10% on the initial premium and no extra commission is paid for reinstatement premiums.

For this example we used a Poisson distribution and discretised the severity distributions in units of 2,500.

Figure 8 shows the severity distributions calculated using methods 2 and 3 as in Section 3. We note that the severity distribution given by method 2 has a significantly heavier tail, which overestimates the expected severity to the layer. Table 7 shows the expected severity and implied frequency. The assumed expected loss cost is \$350,000.

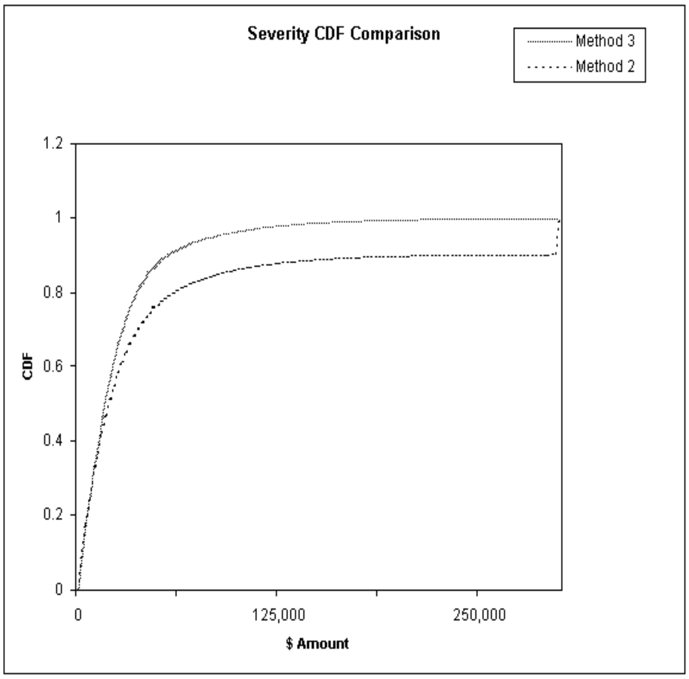

Figure 8: Severity distribution $300,000 xs $200,000

		$300,000 xs $200,000
Loss Cost		$350,000
Method 2 "blended severity"	Severity Frequency	$26,727 13.1
Method 3 "benchmark severity"	Severity Frequency	$54,970 6.37

Table 7. Expected Severity and Frequency using Methods 2 and 3 in Section 3

Note that, as discussed in the casualty example, the second method overestimates the severity since it does not take into account policy limits. In this case it assumes that the first policy may also impact the layer which is not the case.

Figures 9 and 10 below show the probability density function and distribution function of the aggregate losses. In this case we observe that the lognormal and the aggregate loss distribution given by our exposure base severity distribution approach are very similar.

The reason why we do not observe the "spikes" in this distribution function is because the expected frequency is high and therefore due to the Central Limit Theorem the curve results in a smooth curve. We note that the aggregate loss distribution given by the second approach has a significantly higher tail than the lognormal approximation and the exposure base distribution.

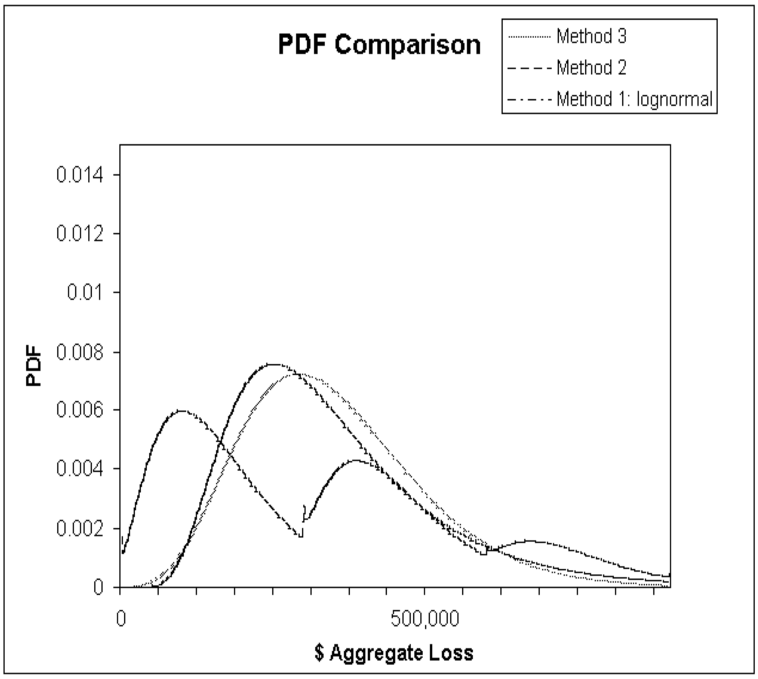

Figure 9: Probability density function of aggregate losses \$300,000 xs \$200,000

Table 8 gives the expected value of the features under this treaty. Note that the only loss sensitive features are the reinstatement premiums.

	Method 1:lognormal		Method 2:benchmark		Method 3:exposure	
	Amount	% Premium	Amount	% Premium	Amount	% Premium
Prov. prem	612,000		612,000		612,000	
Reinst. prem	155,948		208,823		163,251	
Total. prem	768,448	100%	821,323	100%	775,751	100%
Tot. Loss	350,000	45.55%	350,000	45.12%	350,000	42.61%
Brokerage	61,250	7.97%	61,250	7.46%	61,250	7.9%
Marginal CR	411,273	53.52%	411,250	50.07%	457,771	53.01%

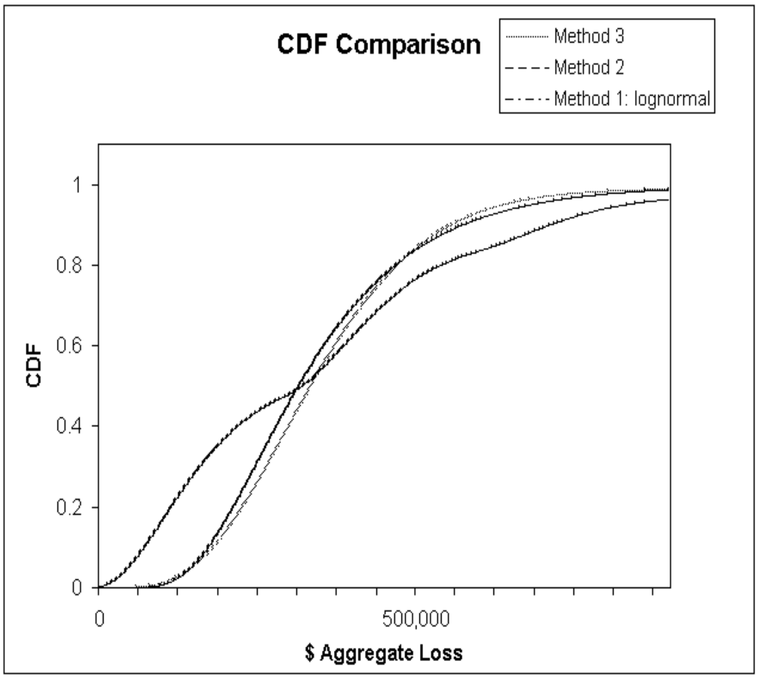

Figure 10: Aggregate loss distribution $300,000 xs $200,000

Table 8. Expected results $300,000 xs $200,000.

We note from Table 8 that in this example the lognormal is a reasonable approach. When there is a large expected number of claims, the Central Limit Theorem dominates and the resulting aggregate distribution should not be too far form a Normal approximation. The overall results obtained under method 2 are underestimated since as discussed above this method overestimates the severity and therefore it overestimates the expected reinstatement premium receivable by reinsurers.

From the examples above we note that when a large number of claims is expected, fitting a parametric distribution seems to be a reasonable approximation to the aggregate loss distribution. However, for high excess layers or lines of business where primary policies include large deductibles and a small number of loss is expected to impact the reinsurance layer, using a lognormal or other parametric distribution would underestimate the expected results making the results look better. This underestimation may have significant impact on underwriting decisions.

5 Practical considerations for implementation of exposure based severity distribution

The method we develop in Section 3.3 may at first appear time consuming to compute because the exposure method must be run several times just to obtain the severity distribution. This severity distribution is then used to compute the aggregate loss distribution using Panjer recursion or any other alternative methods. The computational time for some lines of business may increase significantly, particularly for property lines where the severity depends on each combination of policy limit and deductible. However, we have implemented this example for real accounts and it runs reasonably fast. We discuss below some issues that should be taken into account when using this model for practical purposes.

5.1 How to choose the span h

Most methods available to compute aggregate losses that allow frequency and severity distributions are based on a discrete loss distribution. When the loss distribution is a continuous function, a discrete version can be computed as in Appendix A. In order to do so we must choose the units in which the discrete probability function will be calculated. The chosen units h are called "the span" or units. There is no definite methodology on how to chose the span. If we choose a very small span then we require more iterations to go far enough in the tail of the aggregate loss distribution (say 99.9%), which increases computational time. If we choose a large span we lose resolution and accuracy in the distribution function.

By fitting a parametric distribution, we can estimate the desired percentile of the distribution. We can fix in advance the maximum number of iterations we want to perform and the minimum size of the span, for example $1,000$. Then we can estimate a preliminary span as follows:

$$h^* = \min\left(1,000, \frac{\text{99.9th percentile}}{\text{Max. iterations}}\right).$$

It is desirable that h and $\frac{\ell}{h}$ are integers, where ℓ is the size of the layer. Therefore by trial and error and based on h^* we can estimate an appropriate h.

5.2 How to include ALAE in the aggregate loss distribution

Apart from the losses incurred there may also be other extra costs covered by the treaty, such as defense costs or Allocated Loss Adjusted Expenses (ALAE). ALAE might be included within the limits, in which case the selected loss cost already includes ALAE, or it might be treated pro-rata, in which case a load must be added to the loss cost. See for example Strain (1987).

If ALAE are treated pro-rata, a quick and easy way to allow for this extra cost in the aggregate loss model described above is to estimate an average ALAE load for the treaty, say θ. Therefore if the reinsurer's total aggregate losses are S, reinsurers would pay $(1+\theta)S$.

When we discretise a loss distribution in units of size h, the aggregate loss S is also calculated in units of size h, therefore the following holds:

$$P\left(S + ALAE = (1+\theta)rh\right) = P(S = rh) \qquad \text{for } r = 0, 1, 2, \ldots$$

So we can expand the units of the aggregate loss distribution from h to $(1+\theta)h$ but keeping the probability distribution unchanged.

6 Conclusions

Reinsurance is a very complex industry that requires sophisticated modelling techniques as well as market and business knowledge. It is perhaps an area where actuaries and underwriters could both make use of their market and technical skills and work together in order to have a better understanding of the risk.

In this paper we have shown how actuarial techniques may make a significant impact when pricing excess of loss treaties with loss sensitive features. By using the models presented in this paper, actuaries and underwriters can help design the best terms and conditions when pricing a risk.

We have shown that by ignoring the probability of zero losses and the mixture of limits and deductibles it is possible to underestimate the expected total combined ratio for the treaty. In the examples shown above, the impact may not appear significant. However for treaties where the terms are tight a difference of 1% or 2% in the combined ratio may make a big difference in the underwriting decision on whether or not to support a reinsurance program. As discussed above, it is of particular importance to use the approach developed in this paper for those risks where a low frequency is expected and a consequently high probability of zero losses. However, when the risk is such that a large number of claims is expected, fitting a parametric distribution to the aggregate loss distribution would be a reasonable approach.

The difficulty of effectively communicating the output of an aggregate model has been discussed, particular when the Panjer recursive algorithm is used to compute the aggregate loss distribution. This is perhaps an area where actuaries may not be required to explain the mathematical details of the model but to explain why the results vary and how to make the best use of these results when pricing a risk and making profitability assessments. The underwriter would then make his decision based not only on the results of the model but also based on their industry insight and experience.

In this paper we have assumed independence between lines of business when a multi-line treaty is priced. However in many practical cases this assumption may not hold. Further research should be carried out in order to incorporate correlations and dependence models between lines of business. Use of frequency dependence or copulas may be helpful in this regard.

Furthermore, when pricing multi-layer reinsurance, layers of the same underlying risk are highly dependent not only through the number of claims but also a claim would only impact higher layers when all lower layers have had a full limits loss. Therefore the dependence structure between layers should be taken into account when assessing the overall risk for multi-layer treaties where each layer is subject to different aggregate conditions. See for example Mata (2000) and Mata (2001) where the effect of the dependence structure between aggregate losses for consecutive layers is studied.

Acknowledgments

The methodology, results and discussion presented in this paper are the personal point of view of the authors and no company or organisation may be responsible for any opinion or comment total or partial expressed in this article. The worked examples provided in Section 4 of this paper were derived from purely hypothetical assumptions and any similarity between these examples and any existing ceding company or insurance company are coincidental only.

References

[1] Dickson, D. and Waters, H. (1992) *Risk Models*. The Institute and The Faculty of Actuaries, London and Edinburgh.

[2] Heckman, P. H. and Meyers, G.G. The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions. *Proceedings of the Casualty Actuarial Society LXX*, pp.22-61.

[3] Klugman, S.A., Panjer, H.H. and Willmot, G.E. (1998) *Loss Models: from data to decisions*. Wiley series in probability and statistics.

[4] Mata, A.J. (2000) Pricing excess of loss reinsurance with reinstatements. *ASTIN Bulletin* **30***(2)*:349-368.

[5] Mata, A.J. (2002) Asymptotic dependence between reinsurance aggregate claim amounts. *ASTIN Bulletin: To appear.*

[6] McClenahan, C. (1998) Ratemaking. In Foundations of Casualty Actuarial Science. *Casualty Actuarial Society.*

[7] Panjer, H.H. (1981) Recursive evaluation of a family of compound distributions. *ASTIN Bulletin,* **12***(1)*:21-26.

[8] Patrik, G.S. (1998) Reinsurance. In Foundations of Casualty Actuarial Science. *Casualty Actuarial Society.*

[9] Robertson, J. (1992) The Computation of Aggregate Loss Distributions. *Proceedings of the Casualty Actuarial Society LXXIX*, pp. 57-133.

[10] Sanders, D. (1995) Pricing in The London Market. *Proceedings of the General Insurance Convention, Bournemouth, October 1995.*

[11] Strain, R.W. (1987) *Reinsurance*. Strain Publishing Incorporated, USA.

[12] Venter, G.G. (1983) Transformed Beta and Gamma distributions and aggregate losses. *Proceedings of the Casualty Actuarial Society LXX*, pp. 156-193.

[13] Wang, S. (1998) Aggregation of correlated risk portfolios: models and algorithms. *Proceedings of the Casualty Actuarial Society LXXXV*, pp. 848-939.

Appendix

A A review of methods to compute aggregate loss distributions

Following is a brief discussion of some of the methods available to compute aggregate loss distributions. The first three methods are fairly simple to understand mathematically and intuitively. The final two methods require a certain level of advanced mathematics and therefore we recommend them only to more adventurous readers. Table A1 provides a summary of advantages and disadvantages of all of the methods presented.

In all cases, unless noted otherwise, we are considering the collective risk model

$$S = X_1 + X_2 + \cdots + X_N \tag{A.1}$$

as defined in Section 1.

A.1 Fit Distribution

Since the desired output of any method is a loss distribution, it makes sense to try and utilise a parametric distribution function which is already well understood, and which accurately represents the aggregate experience under consideration. There are several methods available to do this. If many observations of aggregate loss experience are available, standard statistical estimation procedures such as maximum likelihood may be used. For distributions with only two parameters such as the lognormal or the Pareto, an explicit consideration of the frequency and severity distributions can lead to an aggregate distribution function. This relies on the following relationships between moments of the aggregate losses, severity and frequency distributions:

$$E[S] = E[N]E[X] \qquad \text{and} \qquad Var(S) = E[N]Var(X) + Var(N)\left(E[X]\right)^2,$$

see, for example, Dickson and Waters (1992).

Fitting a parametric distribution to the aggregate losses does not take into account the probability of having zero losses which is a possibility in most excess of loss treaties. This is a clear disadvantage of this method as discussed in Section 3.1. Venter (1983) provides an illustration of how to allocate a probability mass at zero when modelling aggregate loss distributions.

A.2 Simulation

Simulation is an easily understand method to model aggregate losses. This method attempts to recreate the actual insurance process itself by simulating many random periods of experience. The results are collected and presented as an empirical distribution. To simulate aggregate losses, a frequency and severity distribution must be chosen depending on the experience and exposure of the risk. Selecting a benchmark severity distribution and a Poisson distribution may be the easiest approach. Then the following basic algorithm may be followed:

1. Simulate k values from the frequency distribution: $n_1, n_2, \ldots, n_k$.
2. For each value of the n_i given above simulate n_i realisations of the severity distribution: $x_1, \ldots, x_{n_i}$.
3. Calculate the aggregate loss as follows:

$$\begin{aligned}
s_1 &= x_1 + \cdots + x_{n_1} \\
s_2 &= x_1 + \cdots + x_{n_2} \\
&\vdots \\
s_k &= x_1 + \cdots + x_{n_k}
\end{aligned}$$

   These k values of aggregate losses represent k realisations of S in k periods of experience. Using the simulated values the aggregate loss distribution may be estimated.

Simulation presents several shortcomings. First, as a practical consideration, running a simulation model more than once will give different results due to random fluctuations. Depending on how the model is used, this may be a problem. For example, if the model is used for pricing, then for a particular run the risk may appear profitable. However, if the models is run a second time the profitability margin may be different. Second, simulation presents the results of random processes, but does not provide any real insight into how that process behaves. Despite these issues, there are some complex insurance and reinsurance products which may be modelled only by means of a simulation model.

A.3 Panjer Recursive Algorithm

Panjer (1981) presented an algorithm which makes use of a recursive relationship. This algorithm requires a discretised severity distribution and a frequency distribution. The algorithm is specified in terms of a frequency distribution that satisfies the following property:

$$P(N = n) = \left(a + \frac{b}{n}\right) P(N = n - 1) \qquad \text{for } n = 1, 2, 3, \ldots$$

The three distributions that satisfy this property are the Binomial, the Poisson and Negative Binomial distribution.

For the Poisson (λ) distribution, $a = 0$ and $b = \lambda$.

For the Negative Binomial (α, β) with probability function

$$P(N = n) = \frac{\Gamma(\alpha + n)}{\Gamma(\alpha) n!} \left(\frac{1}{1 + \beta}\right)^{\alpha} \left(\frac{\beta}{1 + \beta}\right)^{n} \qquad \text{for } n = 0, 1, 2, \ldots$$

$a = \left(\frac{\beta}{1+\beta}\right)$ and $b = (\alpha - 1)\left(\frac{\beta}{1+\beta}\right)$.

If S represents the aggregate losses, X the loss severity and N the number of claims, then for appropriate units given by the span:

$$P(S=0) = P(N=0)$$
$$P(S=s) = \frac{1}{1-aP(X=0)} \sum_{x=1}^{s} \left(a + \frac{bx}{s}\right) P(X=x)P(S=s-x) \quad \text{for } s = 1,2,3,\ldots$$

This recursive algorithm as well as the rest of the methods discussed below require the severity distribution to be discretised in appropriate units. There are several methods that can be used to discretise the severity distribution depending on the degree of accuracy required for the purpose at hand.

One of the easiest methods available for practical purposes is the method of "matching-mean" which preserves the mean of the continuous version of the severity distribution. Assume X represents the ground up severity and $F_X(x)$ its distribution function. Select the units in which we want to discretise the distribution, say h. We then evaluate the limited expected value as in (1) in Definition 3 at multiples of the units h as:

$$E[X \wedge jh] = \int_0^{jh} (1 - F_X(u))du \qquad \text{for } j = 1,2,\ldots$$

then the probability function is given by:

$$P(X=0) = 1 - \tfrac{E[X \wedge h]}{h}$$
$$P(X=jh) = \tfrac{(2E[X \wedge jh] - E[X \wedge (j-1)h] - E[X \wedge (j+1)h])}{h} \quad \text{for } j = 1,2,3,\ldots$$

See Wang (1998). This method can be easily adapted to discretise the conditional distribution for losses to an excess of loss layer ℓ xs m as:

$$P(X_m = 0) = 0$$
$$P(X_m = jh) = \tfrac{(2E[X \wedge m+jh] - E[X \wedge m+(j-1)h] - E[X \wedge m+(j+1)h])}{h \times (1-F_X(m))} \quad \text{for } j = 1,2,3,\ldots,\tfrac{\ell}{h} - 1$$
$$P(X_m = \ell) = 1 - \sum_{j=0}^{\ell/h-1} P(X_m = jh)$$

A.4 Fourier Transform: Heckman-Meyers (1983)

The following sections require a higher degree of mathematical knowledge. We advise the most practical readers to skip these sections and look at Table A.1 at the end of this appendix where we summarise the practical advantages and disadvantages of all the methods presented to compute aggregate loss distributions.

If X and Y are two independent random variables with distribution functions $F(X)$ and $G(Y)$ the probability density function of the sum $Z = X + Y$ is given by

$$f_z(z) = \int_0^z f(z-y)g(y)dy, \tag{A.2}$$

where $f(x)$ and $g(y)$ are the probability density functions of X and Y respectively. The above function is mathematically known as the convolution of $f(x)$ and $g(y)$.

There is an easy way to calculate the convolution by inverting the characteristic function of a distribution function. The characteristic function of a random variable X (also called the Fourier Transform of the distribution $F(X)$) is defined as:

$$\phi_X(t) = E[e^{itX}] = \int_0^\infty e^{itx} f(x)dx.$$

Note that the characteristic function is equal to the moment generating function evaluated at it. There is a one to one relationship between the characteristic function of a distribution function and the distribution function. Therefore, by inverting the characteristic function we obtain the distribution function.

When interested in computing aggregate loss distributions we can do so by inverting a Fourier Transform. The idea is based on the following relationship:

$$\phi_S(t) = \phi_N\left(\log \phi_X(t)\right), \tag{A.3}$$

where $\phi_S(t)$ is the characteristic function of the aggregate loss in (A.1), $\phi_N(t)$ is the characteristic function of the frequency distribution and $\phi_X(t)$ is the characteristic function of the severity distribution. See Dickson and Waters (1992) or Wang (1998). Then if we invert $\phi_S(t)$ we obtain the distribution function of the aggregate loss S.

There are various algorithms to invert Fourier Transforms. Heckman and Meyers (1983) present an algorithm to invert the Fourier Transform of an aggregate loss distribution for any discrete severity distribution and a negative binomial frequency distribution.

The next section briefly discusses the Fast Fourier Transform which is allows for improved computational speed as compared to the direct inversion method.

A.5 Fast Fourier Transform

The Fast Fourier Transform (FFT) is computationally easier and faster to implement even with the use of a spreadsheet. Excel has an add-in function that allows us to calculate Fourier Transforms and their inverse.

Robertson (1992) presents the mathematical details of the FFT and its inverse IFFT. Robertson also discusses its usefulness to compute aggregate loss distributions and an algorithm that can easily be implemented to compute aggregate loss distributions. Wang (1998) also discusses the advantages of using the FFT to compute aggregate loss distribution. The algorithm requires a discretised severity distribution and a frequency distribution. The severity distribution must be a vector of length 2^k where $k > 0$. Then using the property in (A.3) the Fourier Transform of the aggregate loss can be calculated. Using the inverse FFT we obtain the aggregate loss distribution. The basic steps of the algorithm are:

1. Discretise the severity distribution $f(x)$ in appropriate units, such that we have 2^k units. If there are less units than 2^k the rest of the vector may be filled with zeros.

2. Apply the FFT to the severity distribution in step 1.

3. Evaluate the characteristic function of the frequency distribution in $\log FFT(f(x))$ to obtain the Fourier Transform of S as in (A.3).

4. Apply IFFT to the resulting function in step 3 to obtain the aggregate loss distribution.

The main advantage of this method over all the previous methods described above is that it takes into account frequency and severity distributions and is computationally faster than simulation and the Panjer recursive algorithm.

Another advantage of this method is that it can be readily incorporated into any mathematically capable programming language, and many actual implementations of the algorithm in a variety of languages are available. However this brings up one of the defects of the method. Differences in the various implementations make it hard to reconcile results.

Another practical problem with the fast Fourier transform has to do with the fact that for certain inputs into the model, the output is unusable. The resulting function obtained by inverting the Fourier Transform appears to be a sine function which cannot be used as a distribution function. Though this does not occur often, we have observed that this occurs when the "span" used is too small compared to the mean of the distribution. This problem can be overcome by increasing the size of the span, however at that point some resolution is lost.

Method	Advantages	Disadvantages
Parametric distribution	1) Simple implementation. 2) Easy to understand.	1) Ignores the probability of zero losses. 2) For low frequency layers it mis-estimates the expected value of loss sensitive features.
Simulation	1) Easy to understand and to implement. 2) Allows showing interaction between layers and can be used to easily assess the risk for multi-layer risks.	1) It produces different results every time the model is run. 2) Requires an explicit distribution and it is not possible to incorporate policy limits and deductibles.
Panjer recursion	1) For fixed frequency and severity the same answer is always obtained. 2) It works for any severity and frequency distribution, even empirical distributions. 3) Any number of points may be used when discretising the severity distribution.	1) Implementation and time to run may be a problem depending on the mean of aggregate losses and the expected frequency.
Inverse Fourier Transform or Fast Fourier Transform	1) Fast and efficient to implement and run. 2) Easy to generalise for multi-line aggregate losses when independence is assumed.	1) The number of points has to be a power of 2. 2) Severity and aggregate distribution should be the same length,taking unnecessary memory. 3) Small size of the "span" may cause the inverse Fourier Transform to "cycle", giving a useless output. 4) Different implementation may produce different answers.

Table A.1 Summary of various methods to compute aggregate loss distributions.

B Fitting a Negative Binomial as the frequency distribution

The Poisson distribution is perhaps the most commonly used frequency distribution due to its convenient mathematical properties. It is fairly easy to implement in practice since it only requires the estimation of one parameter.

However it has been largely discussed that in practice some of the characteristics that define the Poisson distribution do not hold. See, for example, Heckman and Meyers (1983).

The next best option as a frequency distribution is the Negative Binomial. There are several possible parameterizations of the Negative Binomial. In the rest of this Appendix we will refer to the Negative Binomial distribution with parameters α and β and probability function

$$P(N=n) = \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)n!}\left(\frac{1}{1+\beta}\right)^{\alpha}\left(\frac{\beta}{1+\beta}\right)^{n} \qquad \text{for } n = 0, 1, 2, \ldots$$

with $E[N] = \alpha\beta$ and $Var(S) = \alpha\beta(1+\beta)$.

It can be shown that if N, the number of losses from ground up follows a Negative Binomial distribution with parameters α and β and $F(x)$ represents the claim size distribution function, then N_m, the number of losses in excess of m, also follows a Negative Binomial distribution with parameters α and $\beta' = \beta(1-F(m))$.

Definition B.1 *If Y is a random variable, the variance multiplier of Y is defined as the ratio between the variance and the mean of a certain distribution, i.e.*

$$VM(Y) = \frac{Var(Y)}{E[Y]}.$$

Note that for the Negative Binomial the variance multiplier is given by:

$$VM(N) = \frac{\alpha\beta(1+\beta)}{\alpha\beta} = (1+\beta).$$

Therefore if it is possible to estimate the variance multiplier for the frequency distribution from ground up, then a Negative Binomial can be fitted to the number of losses as follows:

$$\beta = VM - 1 \qquad \text{and} \qquad \alpha = \frac{E[N]}{\beta}.$$

The variance multiplier depends on the line of business and it can be chosen *ad hoc* depending on the account. There is no established methodology on how to chose the variance multiplier. A variance multiplier of 2 or 3 might be reasonable for most lines of business. If enough claim count data is available one can get an insight on the variability of the number of claims per year.

Given a variance multiplier from ground up, the expected frequency to the layer and the claim size distribution we can fit a Negative Binomial distribution for the number of losses in excess of the attachment m.

It is worth noting that for larger values of the attachment m the variance multiplier decreases, i.e.

$$VM(N_m) = \frac{Var(N_m)}{E[N_m]} = (1 + \beta(1 - F(m))) < (1 + \beta) = VM(N),$$

in other words the variance multiplier of the number of losses in excess of m is lower that the variance multiplier for the number of losses from ground-up. Furthermore, the variance multiplier for large values of the attachment tends to 1, i.e.

$$VM(N_m) = 1 + \beta(1 - F(m)) \longrightarrow 1 \qquad \text{as } m \to \infty.$$

since $F(m) \to 1$ as $m \to \infty$.

Hence, since for large values of the attachment the variance multiplier tends to one, fitting a Poisson distribution for high layers would be a reasonable approximation even when the underlying frequency follows a Negative Binomial distribution.