1999 GISG Reinsurance Pricing Working Party

Members of the working party: Sanjiv Chandaria (Co-Chairman)
Fergus Mackie (Co-Chairman)
James Faulkner
Paul Gates
Matthew Maguire
Grant Maxwell
Gregory Overton
Simon Sheaf

Introduction

1 The 1998 Reinsurance Pricing working party produced a paper which covered some of the actuarial and other issues encountered in arriving at a risk premium. A workshop was held at the conference which provided a detailed worked numerical example.

2 This year's working party has attempted to build on this work by describing typical pitfalls which may confront the pricing actuary and some potential solutions. Some numerical examples are included. It should be noted that the list of pitfalls is not intended to be exhaustive.

3 The paper is divided into a number of sections. A basic overview of the pricing process is followed by a commentary on each of the potential pitfalls faced as follows:-

(a) Overview of the pricing process;

(b) Typical pitfalls:-

(i) need to provide all losses that have ever been above data limit;

(ii) developing individual claims;

(iii) underlying policy limits;

(iv) benchmarks for higher layers;

(v) extreme values;

(vi) pricing aggregate features;

(vii) using the wrong aggregate claims distribution.

(c) Appendix – Use of Panjer recursion

4 It should be noted that this paper reflects the majority view of the working party.

Overview of the pricing process

5 Pricing of non-proportional reinsurance treaties is not a standardised process due to the varied and specialised nature of the underlying business and to the often unique features of the contracts themselves.

6 Despite this there are certain steps that are commonly taken when coming up with a price for this business. A typical set of steps is described below. Each is potentially fraught with problems and we discuss some of the more common ones later on in this paper. In practice considerable judgement is required at each step.

How to move from raw data to a final price

7 In the following paragraphs, all our comments relate to a contract incepting 1 January 2000.

8 Check you have the appropriate data. You need the following:

(a) individual claims (ideally) for the last, say, 10 years showing annual evaluations of incurred and paid losses for all claims that have ever been above the predefined census point[1]. (See Pitfall 1 for further details).

(b) exposure information relating to the last 10 years.

(c) the complete treaty document – known as the "slip".

(d) an understanding of the underlying business and the data provided.

9 Decide on an appropriate claims index to bring the historical data up to next year's monetary terms. This will often be related to RPI but is unlikely to match it due to the many other factors that determine claim amounts. Restate historical data based on underwriting[2] year.

10 Determine "Data Limit" by trending the census point of the oldest underwriting year for which we have loss data to next year using the claim trend index. For instance if we have all losses above £50,000 going back to 1990 and we are

[1] Typically 50% of the lowest deductible being priced.

trending claims at 6% per annum the data limit is 50,000 * 1.06^(2000-1990) = 89,542. This Data Limit must be below the lowest deductible.

11 Create a triangle of incurred claim numbers above the Data Limit.

12 Project the claims numbers to ultimate using standard chain ladder or BF techniques.

13 Calculate the ultimate claim numbers for each underwriting year per unit of exposure. Use this information together with market knowledge of past and future likely frequency trends to select a claim count above the Data Limit per unit of exposure for 2000. Multiply this by the expected next year exposure to get the expected next year claim count above the Data Limit[2].

14 Choose an appropriate claims size distribution benchmark curve or alternatively fit one to the data[4]. The latter can be done in Excel or commercially available software packages.

15 Use the curve to extrapolate from the expected next year claims above the Data Limit to the number expected above the treaty deductible. e.g., 10.3 claims expected above the data limit $\Rightarrow$ 10.3 * (1-F(Deductible)) / (1-F(data limit)) = (say) 5.4 It is mathematically possible to extrapolate the claim size distribution curve backwards to estimate a claim count above a deductible that is lower than the Data Limit. This is not recommended, however, unless you have strong independent evidence that the selected loss size distribution is a good fit for values lower than the data limit.

16 Check the 5.4 figure for reasonableness by forming a triangle of incurred claim counts above the deductible and repeating paragraphs 10 and 11. Often there will not be sufficient data to feel happy about the projection to ultimate but it should be possible to get a comfort level for the reasonableness of the figure arrived at in step 13.

---

[2] This will be higher then the expected claim count to the layer being priced as the Data Limit must be lower than the lowest deductible. There is no reason why it should be an integer value as it is a mean expected value not the mode.

17 Use the same claim size distribution curve to estimate the average severity in the layer. This can be calculated using limited expected values (LEV[3]). The average severity for claims above the deductible is:-

$$\frac{(LEV(x; Limit + Deductible) - LEV(x; Deductible))}{(1 - F(Deductible))}$$

This is typically around 50%[4] of the limit for a balanced[5] layer unless the deductible is very high.

18 Multiply the frequency in 13 by the severity in paragraph 15 to arrive at the unadjusted pure premium.

19 Estimate the aggregate loss distribution setting the pure premium as the mean and using past variability of the aggregate loss data or knowledge of benchmark data to help gauge the shape of the distribution.

20 Use this distribution to estimate the expected loss and expected premium[6] after all aggregate features such as annual aggregate deductibles (AADs), annual aggregate limits (AALs), reinstatement premiums, swing rated premiums etc.

21 Determine required profit load and expense load for this treaty. Neither of these tasks is simple if done correctly – see section B of pitfall 6 for further details of the issues surrounding profit criteria.

22 Estimate the discount factor to be applied to losses (LDF). This can either be a benchmark figure (eg 0.8 for claims made liability business 0.7 for claims occurrence liability business) or, if there is sufficient paid losses to the layer, can be estimated directly from the individual paid data.

23 To estimate it from the data:

(a) form a triangle of aggregate trended paid claim amounts to the layer being priced ignoring any aggregate features.

---

[3] LEV(x;X) = The expected value of (x or X, whichever is the lower). LEV(x;X)<=E(x). In Excel the NORMDIST function enables this to be approximated very easily for a lognormal distribution. See appendix in Hogg & Klugman for details of formula for lognormal and other distributions.
[4] Assuming no underlying limits.
[5] A balanced layer is one where limit ≅ deductible.
[6] If the contract contains aggregate features such as paid reinstatements or swing rates then the ultimate premium will depend in part on the aggregate loss distribution.

(b) select a payment pattern based on past payment speeds bearing in mind any likely changes in payment speeds in the future.

(c) apply this pattern to the expected loss calculated in 18.

(d) model the treaty discount factor by applying the selected payment pattern to the expected losses and allowing explicitly for the aggregate features. This should be done on a stochastic basis to capture the non-linear impact on the payout pattern of many aggregate features.

(e) consider the time delay between a claim showing as paid in the data supplied by the broker and when it is physically paid out by the reinsurer and reduce the discount factor accordingly (i.e. apply more discount).

24 Estimate the discount factor to be applied to premiums (PDF1 for up-front premiums and PDF2 for variable premiums), brokerage (BDF), fixed commission (FCDF) and variable commission (VCDF). Terms of trade should be allowed for. This may be a simple matter of looking at the contract wording and adding on agreed terms of trade. On the other hand if premiums and/or commissions are payable based on losses then the premium discount pattern will need to be modelled at the same time as the loss discount factor. It may be necessary to split the premium into an up-front element and a variable element, producing separate PDFs for each element.

Putting it all together

25 To calculate the premium we start with an equation of value. This can be solved in a closed form equation unless the variable premium itself varies based on the value of the up-front premium[7] in which case an iterative approach is simplest. The equation of value is as follows:-

[7] Reinstatement premiums for instance

Present value of income from treaty	=	Present value of outgo from treaty
Up-front Premium*PDF1 + Variable Premium*PDF2	=	E(Loss to layer after aggregate features)*LDF + Brokerage*BDF + Expenses*EDF + Fixed Commission*FCDF + E(Variable Commission)*VCDF + Profit[8]

26 The up-front premium is usually expressed as a percentage of Subject Premium[9]. This has the advantage that if the cedant's estimate of the Subject Premium proves inaccurate the reinsurance premium should adjust accordingly. There are occasions when a change in subject premium does not indicate a corresponding change in exposure. Care needs to be taken not to under or over charge in these circumstances.

27 For example, suppose the original subject premium estimate for 1999 as at 30/9/98 is £100m. Reinsurance premium calculated as 10% of subject premium. During 1999 ceding company unexpectedly cuts rates by 30% to maintain its market share. The exposure to losses has not changed but the subject and hence reinsurance premium has reduced by 30%.

A Simple Example of the Overall Process

28 Cover of £2m xs £2m for 12 months commencing 1st January 2000.

- Subject premium for 2000 expected to be £100m
- Assume 10% interest rate
- Fixed premium payable (for simplicity, assumed payable on expiry)
  - Brokerage 20% - payable at same time as premium
- E(Loss to Layer) estimated to be £1m next year
- Loss discount factor estimated as 0.700
  - Expected profit commission[10] calculated to be £100,000 payable 3 years after expiry of contract. No other aggregate features are present.

---

[8] Here profit is defined as up front profit. If profit is to be recognised at a different rate then a profit discount factor will be required here.

[9] Subject Premium is the amount of premium that the ceding company writes in the year being reinsured. Ideally the reinsurance premium should be expressed as a percentage of an appropriate exposure measure. In practise subject premium is typically used as it is simple to measure and verify.

[10] In practice profit commission is a function of the initial premium so an iterative approach is required here. I have assumed a fixed E(profit commission) of £100,000 for simplicity.

- Expenses 5% of initial premium. Assume these are incurred mid-way through the policy period.
  - Profit target is 8% of up-front premium – Profit is recognised at policy inception. .

Premium = $\dfrac{1{,}000{,}000 \times 0.7 + 100{,}000 \times 1.1^{-4}}{(1.1^{-1}) \times (1 - 0.2) - 0.05 \times (1.1^{-0.5}) - 0.08}$ = 1,281,357 or a rate of 1.28%.

Typical Pitfalls

Pitfall 1 – Need to provide all losses that have ever been above data limit

Introduction

29 When choosing the claims to be included in a pricing analysis of an excess layer, it is important to ensure that information on all relevant claims is provided. Two groups of claims are frequently ignored:-

(a) Claims below the excess point;

(b) Claims that are currently below a given cut-off point but were, at one time, above it.

30 We shall discuss these two claim groups separately, along with the impact that their absence has on the pricing process.

Where to set your cut-off point

31 Let us assume that we are rating a £150k xs £150k layer for the year 2000. Let us also assume that we have 10 years worth of claims on which to base our risk price. We might decide to discard claims below £150k as only claims above this level will impact the layer.

32 A flat cut-off at £150k would exclude a lot of relevant claim data. We will need to revalue all claims to current values in order to give a consistent basis for comparison. If claims inflation has been 5% over the last 10 years, then a claim of £92k in 1989 would, in present value terms, breach the £150k layer. As we are rating for the year 2000, a claim of £88k in 1989 would, in year 2000 value terms, breach the £150k layer as illustrated in the table.

33 To avoid confusion, it is simplest to request all claims above a single limit. The lowest claim value in the earliest underwriting year for which information is provided is the highest claims limit that should be set. For our example, a limit of £85k would be reasonable. This is called the census point.

Year	Original value	Year 2000 Value
1989	88k	150k
1990	92k	150k
1991	97k	150k
1992	102k	150k
1993	107k	150k
1994	112k	150k
1995	118k	150k
1996	123k	150k
1997	130k	150k
1998	136k	150k
1999	143k	150k

'Above' the census point?

34 Once a sensible census point has been established, we must make sure that it is interpreted correctly.

35 Claims develop over time. Some claims may not have reached the cut-off point yet but will do so eventually. These claims are allowed for in the "claim count above the data limit" IBNR calculation.

36 Some claims may have been above the cut-off point at one time, but have since reduced to a current value that is less than the cut-off point. It is easy to exclude these claims from the data supplied for pricing. Indeed, reinsureds often deliberately omit to mention these claims as they believe the fewer claims that they report, the lower the premium will be.

37 These claims which have ever been above the census point, however, can have a significant impact on the development factors used to project claims to ultimate. The negative development of these claims acts to reduce the assumed development of the cohort. Their inclusion therefore will result in a lower price. The reinsureds that omit them are not doing themselves any favours.

38 As an example, assume we have 5 claims from the 1996 underwriting year with values as follows:-

Claim no	Dev yr 1	Dev yr 2	Dev yr 3	Dev yr 4	Currently >150k	Ever >150k
1	100	160	200	200	☑	☑
2	100	160	200	100		☑
3	100	100	300	300	☑	☑
4	200	100	100	100		☑
5	0	0	0	200	☑	☑

39 We need to create a triangle of numbers of claims above the census point at each development period. If only details of claims currently greater than £150k are supplied, then the corresponding row in the claim numbers triangle will be:-

No.s of claims above 150k at dev yr:-	**1**	**2**	**3**	**4**
	0	1	2	3

40 If details of claims ever greater than £150k are supplied, then the corresponding row in the claim numbers triangle will be:-

No.s of claims above 150k at dev yr:-	**1**	**2**	**3**	**4**
	1	2	3	3

41 Most actuaries would project the first sequence of claim numbers to a higher ultimate number than the second sequence. Hence it is almost inevitable that a triangle based on claims ever greater than the census point will produce a lower number of pure IBNR claims than a triangle based on claims currently greater than the census point.

Pitfall 2 – Developing individual claims

42 A number of issues arise when developing individual claims to their ultimate positions prior to fitting a claim size distribution. This section discusses some of these problems.

Claim Size

43 Claims of different sizes may be expected to exhibit different rates and patterns of development. Failing to allow for this may lead to distorted results.

44 A potential solution to this is to divide the claims into a number of groupings depending on their size. Separate triangulations can then be compiled of the claims in the different groupings. Development patterns can then be selected and used to project the relevant claims.

45 One question that then arises is how to form the groupings. There are a number of possible solutions to this. One might be to allocate a claim to a grouping based on its size at the end of a given year of development. An alternative might be to allocate a claim based on its average size at the end of two or more years of development. Clearly, if the years of development that are used are not very early, then you will be unable to allocate claims from the more recent accident years. Not only will this reduce the volume of data from which the development patterns can be derived, it will also mean that you will not know what is the appropriate pattern to use to develop the more recent claims. On the other hand, claims are at their most volatile in the early years of development and you are likely to obtain more homogenous groupings if you base them on claim sizes at a later stage of development. The final choice is a matter of judgement taking account of these factors and the particular circumstances of the account that is being considered.

46 A further question is how many groupings to have. The answer to this will depend on the volume of data available for analysis and the level of variability inherent in it. It is a balancing act between having groupings that are as homogenous as possible and having groupings that each contain a sufficiently credible volume of data for the purpose of the analysis.

Variability in Claims Development

47 Having derived a suitable development pattern for a particular grouping of claims, it is inappropriate simply to apply that development to any claim in the group.

In practice, not all claims develop in exactly the same way - they exhibit significant variability and the selected development pattern merely reflects the average development. Failing to allow for this variability would dampen the apparent variations between claims and hence distort your results.

48 For example, assume that we are pricing a layer with a £120,000 excess and that one of our claim groupings is £75,000 to £100,000. Assume further that our estimated future development factor for this grouping is 20%. Then, applying the average factor to each claim would give a zero cost to the layer for claims currently in this grouping. However, in practice the actual development will be variable around the 20% average figure. This means that some of the claims would be expected to exhibit development greater than 20% and may impact on the layer being priced. In this example, just allowing for average development would understate the claims in the layer.

49 It follows that there is a need to allow for a variability component in the development factor applied to each claim. One possible way to do this is as follows.

50 For each development year, the historical data on individual claims can be used to derive an empirical distribution for the development experienced. This can then be used to randomly generate a "variability factor" to apply to the projected development of a particular claim in that development year. Of course, different variability factors would be generated for each claim in each development year.

51 Having applied the variability factors, it is still important to ensure that the aggregate development projected across all claims is equal to that expected. For example, if your development pattern had indicated that the expected development in year four was 30% then the introduction of the variability factor would mean that you had projected individual claims to develop at a variety of rates, some above and some below 30%, during the fourth year of development. However, once you add the projected development of all claims in the fourth year together, it needs to equal 30%. This can be achieved by scaling all the projected developments up or down by a fixed amount to bring the total to the required amount. For example, going back to the earlier situation, if, having applied the variability factors, your aggregate development came to 32%, you would need to apply a scaling factor of 30/32 to the projected development in year four for each claim.

52 Clearly, you would need to follow similar procedures for each development year.

Closed Claims

53 When developing individual claims, you do not want to develop closed claims beyond their current position. However, if you do not project the closed claims and just apply the "average" development to open claims then, in aggregate, the development will be understated.

54 For example, assume that we expect 40% development in year two and that we have ten claims that are one year developed. Assume the details of the ten claims are as follows:

Claim Number	Amount After 1 Year (000's)	Open/Closed
1	20	Open
2	15	Open
3	25	Closed
4	10	Open
5	5	Closed
6	40	Open
7	10	Open
8	5	Open
9	15	Closed
10	35	Closed
Total	180	

55 The results of applying 40% development to the open claims and leaving the closed claims unchanged are shown below. (Note that, in order to simplify the example, we are ignoring the need to allow for variability which was discussed above. In practice you would need to allow for this but the aggregate result would be the same).

Claim Number	**Amount After 1 Year (000's)**	**Open/ Closed**	**Projected Amount After 2 Years (000's)**
1	20	Open	28
2	15	Open	21
3	25	Closed	25
4	10	Open	14
5	5	Closed	5
6	40	Open	56
7	10	Open	14
8	5	Open	7
9	15	Closed	15
10	35	Closed	35
Total	180		220

56 Hence, the aggregate development = (220-180)/180 = 22.2%

57 In this example, the resulting aggregate development is significantly below 40%.

58 In order to rectify this problem, you need to apply a scaling factor to the projected development on open claims in order to ensure that the aggregate development projected is equal to that expected. In other words, you need to scale the development on open claims in order to ensure that the total development projected is the same as would be obtained by applying the "average" development to all claims (both open and closed).

59 In the example above, applying 40% development to all claims gives 180 * 1.4 = 252. The closed claims total 80, which means that the projected open claims must total 172. Prior to scaling, they total 140. Hence the required scaling factor is 172/140. Applying this gives the following results:-

Claim Number	Amount After 1 Year (000's)	Open/ Closed	Projected Amount After 2 Years Before Scaling (000's)	Projected Amount After 2 Years After Scaling (000's)
1	20	Open	28	34
2	15	Open	21	26
3	25	Closed	25	25
4	10	Open	14	17
5	5	Closed	5	5
6	40	Open	56	69
7	10	Open	14	17
8	5	Open	7	9
9	15	Closed	15	15
10	35	Closed	35	35
Total	180		220	252

60 Hence, the aggregate development = (252-180)/180 = 40%

61 For clarity, we have shown these calculations as a two-stage process. In practise, however, we could combine them all into one step by applying a factor equal to the expected aggregate development divided by the total open claims, to each open claim. In this example the factor would be (180*0.4)/100=72%.

62 Clearly, the same procedure would need to be followed for each development year.

63 It is often not obvious which claims are closed. In many cases, they will not be explicitly labelled as such. In such circumstances, a further problem is how to determine which are open and which are closed. Simply assuming that all claims are open will be unsatisfactory as it is highly unlikely to be true and, consequently, such an assumption would have the effect of dampening the true variability. One solution might be to assume that claims are closed if and only if they have no amounts outstanding. Clearly, this will not always be the case but it is likely to be a reasonable assumption for the majority of claims.

Outliers

64 Having developed the individual claims to their ultimate position, a further question is whether to make use of all of them in fitting the distribution. It is necessary to give careful consideration to the treatment of very large claims since they can have a disproportionate effect on the fitted distribution.

65 One solution might be to ignore them altogether but this would leave no claims above a certain point in the sample to be fitted and would result in a fitted distribution that understated the expected number of claims in the tail.

66 A better solution is to cap large claims at an upper limit. The result of this will be that the number of claims above this limit will be included in the fitting process, but their actual sizes will not be directly used. Hence the fitted distribution will allow for a reasonable number of claims in the tail, but will not be unduly influenced by one or two very large claims.

67 The main problem with this approach is deciding at what point to cap the claims. Clearly, this is a matter for judgement in each individual case after reviewing the distribution of claim sizes and the information on the larger claims. However, in most cases the outliers will be fairly obvious and you would expect to cap relatively few claims.

68 If you are capping claims, then it is important to check that the fitted distribution implies a sensible estimate for the occurrence of claims above the cap. These claims are likely to be of the "low frequency, high severity" type which may not be adequately represented by the standard probability distributions. In these circumstances, techniques such as those discussed in Pitfalls 4 and 5 may help.

Pitfall 3 – Underlying policy limits

69 The aim of this section is to highlight some of the pitfalls associated with underlying policy limits (ie policy limits in the business written by the cedant).

Policy Limits in Exposure to be Priced

70 Policy limits clearly have an impact on the pricing of a particular exposure. For example a layer of £9m xs £1m will have a lower expected loss if a portion of the underlying policies are limited to amounts below £10m. To price this layer using an exposure based approach requires a profile of the underlying policy limits. This would then allow the exposure to be split into that which is limited (perhaps at different levels of policy limit), and that which is unlimited. The layer to be priced can then be split into different sublayers, corresponding to the levels of the different policy limits. Each sublayer will have exposure from the underlying policies which have limits greater than the bottom of the layer. The expected cost of each of these sublayers before allowing for aggregate features can then be calculated normally.

71 Example:

Layer £9m xs £1m
Subject Premium £50m
Policy Limit Profile:

Policy Limits	Proportion of Premium from Policies Subject to Limit
£1m	10%
£2.5m	20%
£5m	30%
£7.5m	15%
£10m	5%
Unlimited	20%

72 Therefore the layer can be divided into sublayers as follows:

0

Sublayer	Subject Premium
£1.5m xs £1m	90% of £50m = £45m
£2.5m xs £2.5m	70% of £50m = £35m
£2.5m xs £5m	40% of £50m = £20m
£2.5m xs £7.5m	25% of £50m = £12.5m

73 The above approach assumes that underlying policy limits cannot be exceeded. This may not necessarily be the case, for example in the case of ECO/XPL (where courts can award greater than underlying policy limits). One simple way to deal with this is to assume that policy limits are exceeded a certain percentage of the time, and therefore increase the exposure for the higher sublayers in the example above.

74 If using a burning cost approach, then ignoring the impact of underlying policy limits may give reasonable results, if the effect of these policy limits has been constant over time. However if the impact has been changing over time (eg due to inflation) an allowance will need to be made.

Impact of Policy Limits on Data

75 There are a number of ways in which underlying policy limits can impact on data to be used in the pricing process. Some examples are given below.

Curve fitting

76 The presence of underlying policy limits in data used for curve fitting will distort the resulting fitted distribution, and could give very misleading results when applied to rate a contract. The following example illustrates this point:-

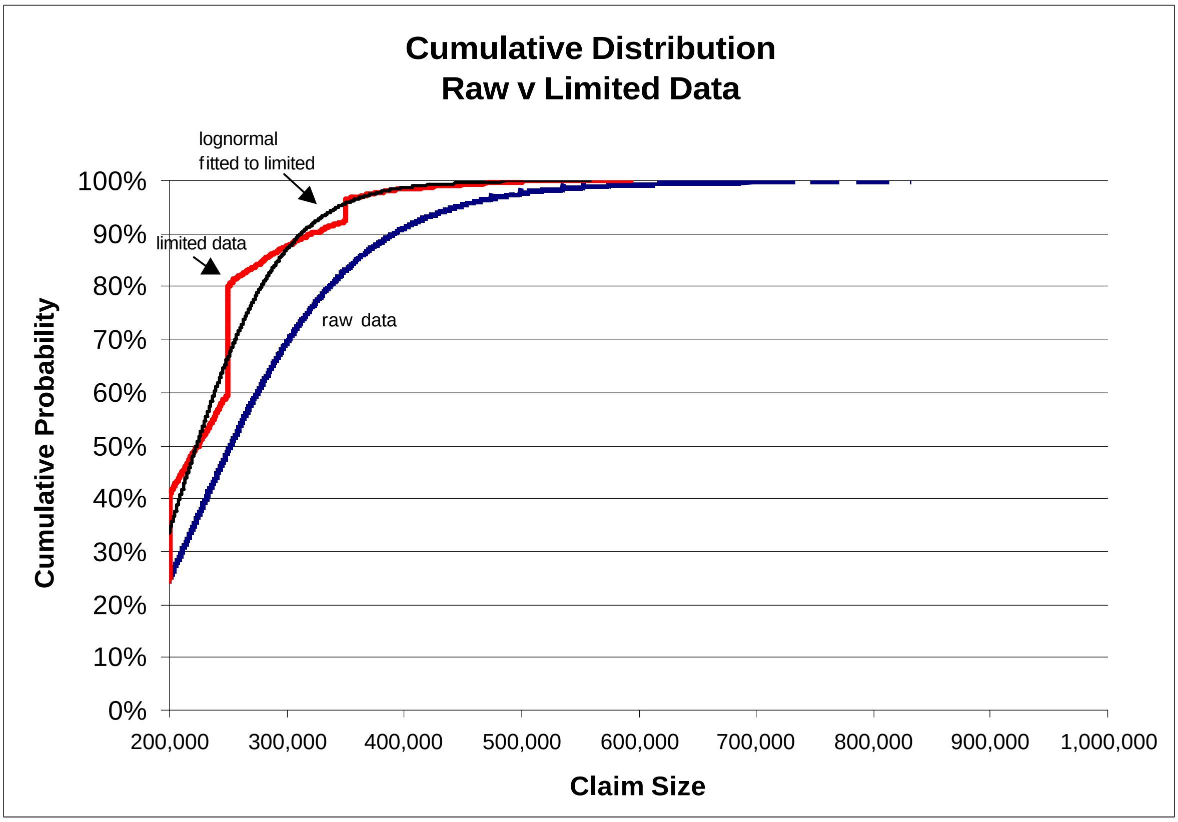
77 We created example benchmark data as follows:

- Simulated a data set of 2,000 claims from a lognormal distribution with mean £266,667 and standard deviation £94,282.

- Randomly overlaid some underlying policy limits on to this data in the following proportions:-

Policy Limit	Proportion of Claims Limited
£200,000	20%
£250,000	40%
£350,000	20%
£500,000	10%
£1,000,000	10%

78 The following graph shows the cumulative distribution of the raw claims, the limited claims, and a lognormal distribution fitted to the limited claims data.



79 The graph shows that the limited claims data is no longer a smooth function, with blocks of claims at each of the policy limits values. The lognormal distribution fitted to this data has a mean of £230,757, and a standard deviation of £61,853. Therefore the policy limits have the effect of reducing the size of the larger claims, which reduces both the mean and the standard deviation of the distribution.

80 As an example of the impact this would have on the pricing of particular contracts, we have calculated the expected cost to two sample layers from both the raw data, and the fitted limited data

Data	£250,000 xs £250,000	£250,000 xs £500,000
Unlimited	£42,537	£1,586
Limited	£14,853	£55

81 Therefore the underlying policy limits have a significant impact on the expected cost to the layer, particularly for higher layers.

82 It is desirable therefore to use a claims size distribution that is free from the impact of underlying policy limits, and then allow explicitly for any policy limits in the exposure to be priced. An algorithm for deriving an unlimited distribution (using both Pareto and lognormal distributions) from data subject to policy limits is described by Patrik[11]. One potential drawback of this approach is that it requires individual claim amount data including the size of the policy limit for each claim.

Impact of Policy Limits on Development Data

83 As well as affecting the claims size distribution, policy limits will impact claims development. Therefore it is important to ensure that any benchmark development factors used to experience rate a contract are derived from data consistent with the exposure to be rated.

[11] PATRIK, G., "Estimating Casualty Incurred Loss Amount Distributions, PCAS

Pitfall 4 – Use benchmarks for higher layers

84 Layers with a high deductible will have relatively few claims. Relying totally on actual claims experience from an individual cedant will produce unstable results. Using the frequency/severity approach described in this paper will help mitigate this to an extent. Nevertheless care needs to be taken especially if a cedant's own claims are used to fit the claim severity curve.

85 Say your book consists of 100 treaties each with high deductibles. If the business was such that you expected a £100m shock loss once in a thousand years for any given cedant then on average only 1 of your treaties would have suffered such a loss in the last 10 years. A once in a thousand years loss is such a freak event that it is pure luck rather than bad management whether or not a given cedant suffers one.

86 If each treaty was individually priced using its own data to fit a loss curve then 99 of the treaties would have been priced assuming 0% chance of a £100m loss. The unlucky cedant who did have a £100m loss in the past 10 years would have been quoted a price roughly £10m higher than if he had not had the loss.

87 In the real world such a cedant would have found another reinsurer who understood the unlucky nature of his £100m hit and would not load the full £10m. Consequently that cedant's treaty would not have been written.

88 Now consider the new portfolio of 99 treaties. There is still a 99/1000 chance of a £100m loss this year yet none of the treaties has been loaded to allow for this. On average therefore this book will be £9.9m underpriced.

89 A potential solution is to use a benchmark loss curve that allows for the one in a thousand year £100m loss for all the treaties. Each cedant would have a load of around £100,000 to allow for this possibility and the book would be correctly priced.

90 In practice it is not clear when a large loss is bad luck or bad underwriting. In general though the higher a layer the less control an individual cedant has over whether the layer is hit and the greater use should be made of benchmarks when pricing.

Pitfall 5 – Allowing for extreme values

Introduction

91 Low frequency, high severity events, such as losses to a high excess of loss contract, are notoriously difficult to model, especially when selecting an appropriate claims size distribution given the sparsity of claims data.

92 Extreme Value Theory ("EVT"), which considers the modelling of extreme events, has been extensively used in hydrology and climatology to model such phenomena as high river levels and high temperatures. EVT is increasingly being used in financial theory and could be used to model extreme insurance losses in the context of a frequency and severity model. In this section we highlight some of the recent research done in this area, and attempt to give an entry point for actuaries into this potentially valuable theory. We do not present the results of any new work, but instead refer extensively to other papers that are likely to be of use to the pricing actuary.

Extreme Value Theory

93 The key result of EVT for the purposes of reinsurance pricing, is the Pickands-Balkema-de Haan theorem which says that:

"For a wide class of loss distributions, the distribution of losses that exceed a high enough excess attachment point is a Generalised Pareto Distribution ("GPD")."

94 This result is similar to the well known central limit theorem, which concerns the limiting distribution of the standardised sample mean. The Pickands-Balkema-de Haan theorem applies to the right hand tail of a distribution, rather than the sample mean, and convergence improves when higher tails are considered, rather than when the number of data points increases.

95 As McNeil[12] points out, the wide class of loss distributions include standard loss distributions such as the Pareto, Burr, loggamma, Cauchy, normal, exponential, gamma, lognormal, Weibull, and beta distributions. McNeil gives some further discussion of this result, as does Embrechts et al[13], Patrik & Guiahi[14] and Reiss &

---

[12] MCNEIL, A.J., "Estimating the Tails of Loss Severity Distributions Using Extreme Value Theory", ASTIN Bulletin, Vol. 27, No. 1, pp.117-137, 1997

[13] EMBRECHTS, P., KLÜPPELBERG, C and MIKOSCH.T, "Modelling Extremal Events", Springer Verlag, Berlin 1997.

Thomas[15]. All four also give details of the GPD, though we recommend McNeil as the most lucid account.

Reinsurance Pricing

96 The simple conclusion from this EVT result is that the pricing actuary should consider using the GPD for modelling the low frequency, high severity losses found in the extreme right hand tails of empirical loss severity distributions.

97 There is also some empirical evidence that the EVT approach, and the GPD in particular, may provide good models for insurance losses. Patrik & Guiahi discuss the Insurance Services Office US liability claim severity model and very large claims of a large global reinsurance company. Reiss & Thomas provide many examples, including large Norwegian fire claim data. McNeil extensively analyses Danish data on major fire insurance losses. Various papers at the 1998 GISG and ASTIN Colloquium discuss the use of EVT in pricing Swiss Re's "beta" high-excess property and casualty layers[16].

98 Reiss & Thomas describe various methods of fitting the GPD to the data, including some non-parametric methods. McNeil and Patrik & Guiahi provide a summary of some of these methods.

Pitfalls

99 Whilst the EVT can be used to assist in the severity distribution fitting process, it should not be regarded as the solution to all the problems of modelling low frequency, high severity events. Other sections of this paper consider the general pitfalls of claims frequency and severity methods (such as the estimation of IBNR claims, or restatement of historical data at current values), and those comments are equally applicable here. In particular, the claims frequency and severity model assumes that the underlying claims are independent and identically distributed. For very large losses there may be clustering, trends, seasonal effects and other reasons to doubt this assumption.

100 The data used in a pricing exercise is often incomplete. Losses may only be reported to the reinsurer if they exceed an agreed limit such as half the excess point. McNeil discusses the impact of such truncated or censored data. To a certain extent

---

[14] PATRICK, G. and GUIAHI, F., "An Extremely Important Application of Extreme Value Theory to Reinsurance Pricing", 1998 Casualty Actuarial Society Spring Meeting, Marco Island, Florida,1998.
[15] REISS, R.D., and THOMAS, M., "Statistical Analysis of Extreme Values", Birkhäuser, Basel, Boston, Berlin, 1997.
[16] LIST, H.F., and ZILCH, R., "Extreme Value Techniques Part I: Pricing High Excess Property and Casualty Layers", ASTIN, 1998.

low level losses are less significant when modelling the extremes of the tail. Sensitivity tests can be performed to gauge the impact of excluding low level loss data on the resulting risk premium estimate.

101 When fitting a GPD to claims severity data, the GPD is usually fitted to the distribution of claims above a threshold that is, of course, below the excess attachment point of the contract we wish to price. Choosing this threshold involves considerable practical difficulty. The Pickands-Balkema-de Haan theorem implies that the convergence of our claims distribution above the excess attachment point to a GPD improves as the threshold increases. However, choosing a high threshold results in fewer losses breaching that threshold, giving fewer data points from which to estimate the parameters of the distribution. McNeil describes in his example how plots of the fitted shape parameter, high quantiles, and the estimated risk premium against the chosen threshold can be used to select an appropriate threshold.

102 Examining the sensitivity of the risk premium to the selected threshold is only one possible sensitivity analysis in the EVT framework. It is also important to check the sensitivity of the risk premium to the largest loss. This can be achieved by refitting the curve to the data with the largest loss removed or by adding a new hypothetical largest loss. Unsurprisingly, the sensitivity is often very high, and EVT allows us to examine and analyse this sensitivity. Again, McNeil guides us through this process.

103 The results of fitting a GPD should be compared with fitting a standard loss distribution, such as a lognormal, loggamma, Burr, Weibull or Pareto distribution. To some extent the differences in these results can give a measure of part of the model uncertainty.

Other Applications

104 McNeil and Patrik & Guiahi describe a similar theorem to Pickands-Balkema-de Haan by Fisher-Tippett, concerning the convergence of the distribution of the maximum claim to a contract to a generalised extreme value distribution. This result has been extensively used in hydrology, but seems less relevant to actuaries. Reiss & Thomas also describe how EVT can be used to develop premium rates that vary by probable maximum loss in a consistent manner.

105 Reinsurance pricing is but one area where EVT could be of use to actuaries. The estimation of low frequency, high severity losses is also part of ruin theory (see Reiss & Thomas and Embrechts et al) and EVT could potentially be used in capital management, securitisation of insurance and investment return problems. Embrechts

et al also considers the use of EVT in time series problems, and in estimating the size of T-year losses.

Conclusions

106 EVT has an ambitious remit: To predict the occurrence of rare events often outside the range of available data. There will always be uncertainty about such events, but EVT can make the best use of any available data and provide a coherent statistical framework in which to analyse that data.

Pitfall 6 - Being over simplistic in pricing aggregate features

Some definitions

What are aggregate features?

107 Aggregate features in a broad sense relate to either contracts or parts of contracts where the amount of recoveries and/or premiums payable depends in some way on the total of all losses to the layer. The more common aggregate features are described below.

Annual aggregate deductibles

108 In this case, under a risk excess of loss contract, recoveries can only be made once the aggregate amount otherwise recoverable from the reinsurance exceeds the aggregate deductible.

Annual aggregate limits

109 As the name would imply, a risk excess of loss contract may have a cap set on the maximum amount that can be recovered from the contract per underwriting year.

Paid reinstatement provisions

110 Under this feature, if a loss occurs to a risk excess of loss layer, a premium (normally related to the original premium paid) is payable to reinstate that part of the cover exhausted by the loss.

Swing-rated contracts / “contingent” premium contracts

111 The operation of a swing-rated contract is such that an initial premium is payable and, as risk XL layer losses start to emerge, a further premium is payable until a pre-agreed maximum is reached. The extra premium is generally related to the emerging losses (and is over 100% of them in most cases to cover claim expenses).

112 There are similar arrangements for proportional treaties.

Profit sharing arrangements

113 A number of options are possible whereby the aggregate performance of a risk excess of loss or proportional treaty determines some form of premium rebate / bonus payable to the cedant.

Aggregate excess of loss contracts

114 Under this type of contract, the aggregate loss record is considered (possibly net of other reinsurances inuring to the benefit of the policy) and if the sum of all of the individual losses exceeds the agreed retention, a recovery can be made.

What's wrong with using burning cost?

115 The burning cost method is commonly used for the reasons of simplicity, clarity and popularity with underwriters.

116 However, the method exhibits vulnerability when used to price aggregate features. A few simple examples show that using burning cost adjusted for such features will be inappropriate. In the following example, an annual aggregate deductible (AAD) of £500 is being incorporated into the rating of a £500 xs £500 layer.

Ground-Up Losses

Year >>	1994	1995	1996	1997	1998
Loss no					
1	525	152	257	290	252
2	250	663	350	863	888
3	326	782	933	390	319
4		294	764	395	481
5		300		951	407
6				763	

Layer Losses

Year >>	1994	1995	1996	1997	1998
Loss no					
1	25	0	0	0	0
2	0	163	0	363	388
3	0	282	433	0	0
4			264	0	0
5				451	0
6				263	
Total before AAD	25	445	697	1,076	388
Total after AAD	0	0	197	576	0
Avg before AAD	526				
Avg after AAD	155				
Credit for AAD	74.3%				

117 In all our examples we are assuming that the information for the years 1994 to 1998 is complete in that there are no IBNR/IBNER issues. In this first example, we see the type of volatile experience that one might associate with a medium risk excess of loss layer, with one “light” year (1994) and one “heavy” year (1997) and the other years somewhere between. The reduction in the risk premium is less than the full £500 by virtue of the years where the average before the AAD was below £500 and thus the full amount could not be deducted.

118 The credit for AAD displayed in these tables is defined as the reduction in the risk premium caused by the introduction of the annual aggregate deductible divided by the amount of the annual aggregate deductible.

119 In the second example, the claims experience is similar and indeed the average burning cost before the AAD is exactly the same as in the first example.

Ground-up Losses

Year >>	1994	1995	1996	1997	1998
Loss no					
1	495	152	690	290	252
2	225	565	350	944	455
3	326	990	884	390	319
4		294	905	395	481
5		300		890	407
6				763	

Layer Losses

Year >>	1994	1995	1996	1997	1998
Loss no					
1	0	0	190	0	0
2	0	65	0	444	0
3	0	490	384	0	0
4		0	405	0	0
5		0		390	0
6				263	
Total before AAD	0	555	979	1,097	0
Total after AAD	0	55	479	597	0
Avg before AAD	526				
Avg after AAD	226				
Credit for AAD	60.0%				

120 What can be seen in the second example is that, as the large losses are centred on the years 1995 to 1997, the reduction in the risk premium following the application of the AAD is considerably less than in the first example.

121 In the third example, the claims experience is similar and once again the average burning cost before the AAD is the same as in the previous examples.

Ground-Up Losses

Year >>	1994	1995	1996	1997	1998
Loss no					
1	778	152	590	290	252
2	515	665	350	744	555
3	656	790	684	390	839
4		294	711	395	481
5		300		890	433
6				713	

Layer Losses

Year >>	1994	1995	1996	1997	1998
Loss no					
1	278	0	90	0	0
2	15	165	0	244	55
3	156	290	184	0	339
4		0	211	0	0
5		0		390	0
6				213	
Total before AAD	449	455	485	847	394
Total after AAD	0	0	0	347	0
Avg before AAD	526				
Avg after AAD	69				
Credit for AAD	91.3%				

122 In this case, the claims are more evenly spread by size across the underwriting years and as a result the impact of the AAD is greater, leading to a much greater reduction in the premium following the AAD.

123 We can see from the above examples that 5 years experience (which is effectively a sample from the many years of possible experience that we could have) is insufficient to judge the “true” effect of the AAD. Techniques that generate the whole distribution of possible outcomes need to be employed.

Solutions

124 Most solutions require the creation of an aggregate claims distribution. An intuitive way of doing this is to consider the constituent elements of frequency and severity, and the fitting of probability distributions to both these elements. Models using these distributions can then be constructed and used to price directly the aggregate features being assessed.

Simulation models

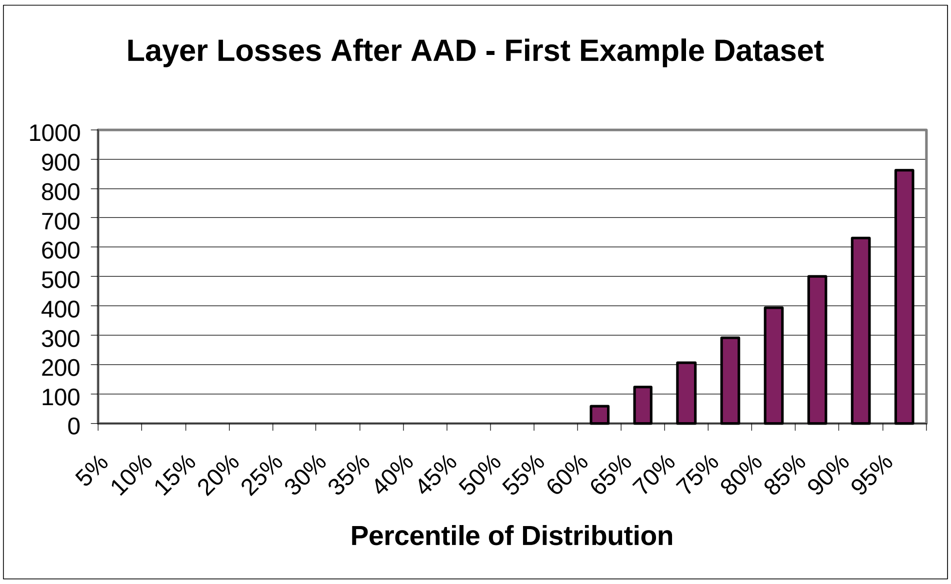
125 Simulation modelling provides an elegant solution to the problems posed by aggregate features, with the ability to construct a model that explicitly takes into account each aggregate feature. The potential downsides relate to the availability of suitable, user-friendly packages and the difficulties involved with convincing many underwriters that the approach is valid.

126 If we return to the sample data sets utilised above, we can employ simulation to determine with more confidence the reduction in the risk premium that would seem appropriate to take account of the AAD of £500 on our layer of £500 xs £500. We use a simple simulation model which samples many times from a Poisson distribution for the number of losses and a lognormal distribution for the severity of losses, and where the parameters are based on the data sets. The results are as follows:-

Example	Average Loss Before AAD	Average Loss After AAD
1	530	188
2	528	186
3	532	180

127 Unsurprisingly, the average loss before the application of the AAD is close in all cases to the 526 that was calculated from the burning costs model (the difference is due to random fluctuation in the simulation process). What is of interest is that the average loss after the AAD is very similar in all three cases, and that the implied credit for the AAD (i.e. the reduction expressed as a percentage of the annual aggregate deductible) is about 69%.

128 One other advantage of constructing a simulation model is that it allows the user to assess the range of outcomes associated with the layer / feature being priced. This information can be useful in answering questions such as “how many years out of 100 is this treaty expected to run clean?” or “what risk premium should be charged so that, based on our projections, the premium should be sufficient 3 out of 4 years?”. On examining the following graph (based on the data set from the first example) it would appear that approximately 55% of the time the treaty will run clean and the premium required for 75% confidence is approximately £300.



Simpler models

129 One may utilise a slightly simpler model that works with the frequency distribution but that simplifies the severity aspect in some way. The advantage is easier explanation to and use by the underwriter. The disadvantage is the possible underestimation of volatility. This latter issue may be resolved if one inserts different options for the severity distribution, which is akin to a simple simulation approach.

Panjer's Recursion Formula

130 This very useful algorithm has clear application in the field of rating aggregate features. This has been covered in the Appendix to this paper.

Direct application

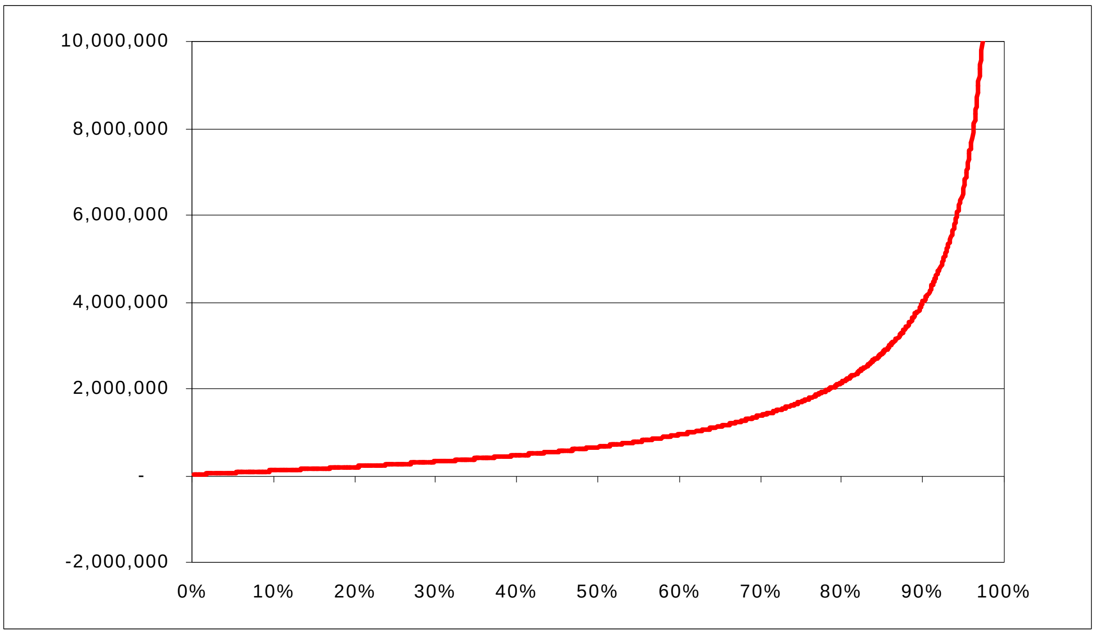
131 A closed form solution can be produced if one makes the simplifying assumption that the aggregate layer claims distribution is lognormal and one makes use of MS Excel's "NORMDIST" built in function. The limited expected value can be calculated based on the cover being rated and the parameters derived, applied to price both the layer loss and most aggregate features directly.

Some further questions relating to Annual Aggregate Deductibles

Section A: Should the risk premium reduce by a constant percentage of the deductible?

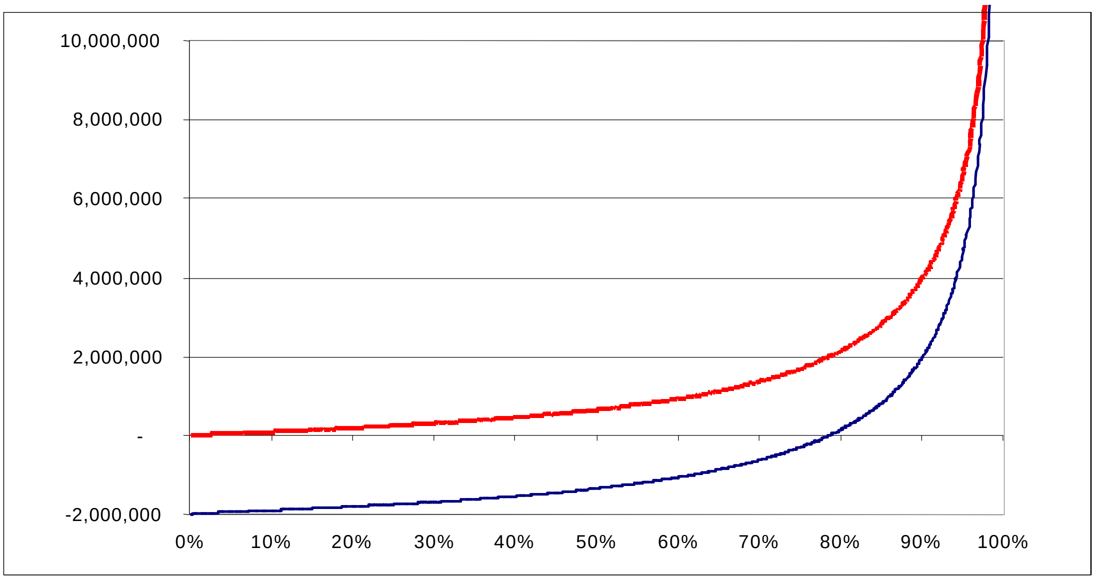
132 When an aggregate deductible is introduced, the risk premium (or the theoretical cost of claims before expenses and profit) is not just reduced by the amount of the deductible.

133 The chart shows the cumulative loss distribution for a line of business. This analysis ignores all inflation or discounting considerations, as well as many other real life issues.



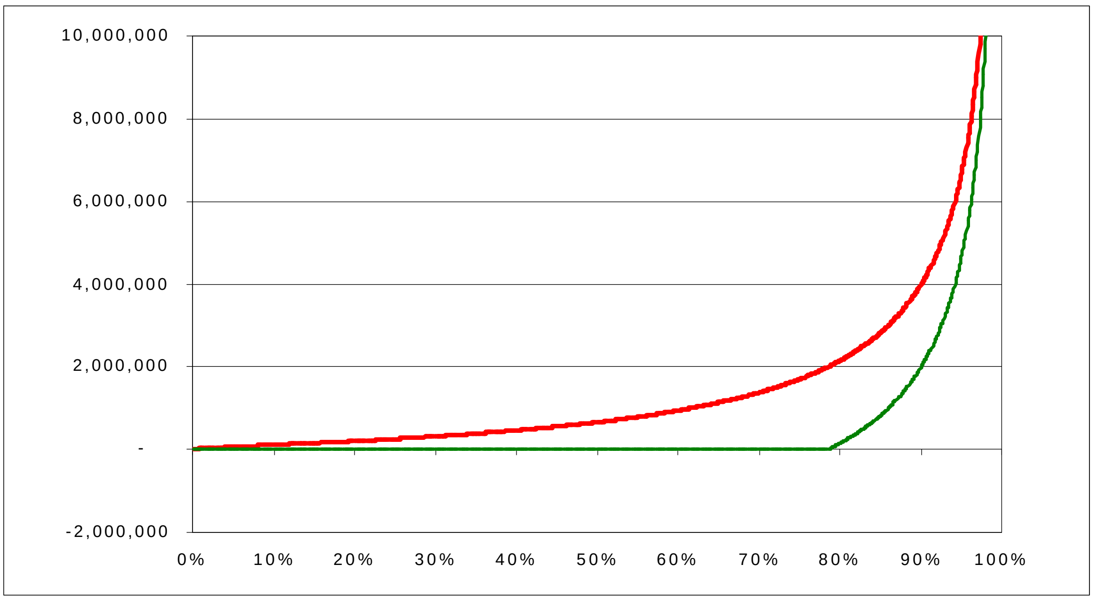
134 The risk premium is simply the area under the chart. In this case it is £1,782,000. What happens when an aggregate deductible of £2m is introduced? It is obvious that the risk premium does not reduce by £2m otherwise a negative premium would result, and from the graph it can be seen that there is risk above £2m, so there must be some premium required.

135 So by how much does it reduce? Is it a constant percentage of the deductible as I have heard it suggested previously, say 50%. So, with a £2m deductible, would the premium reduce by £1m? Taking £2m off every point creates the following graph:



136 As can be seen, there is a fair part of the distribution below zero. If the contract was worded such that if the loss distribution was below zero then the (re)insured made up the difference to the (re)insurer, then the deductible could be simply deducted from the risk premium to give: £1,782,000 - £2,000,000 = -£218,000. A very interesting contract indeed, but not one which is common just yet.

137 Instead, the structure is more like an option pay-off diagram, whereby if the recovery drops below zero then the contract is not exercised. This leaves us with the following chart:



138 By making the same calculation as originally, by determining the area under the line to work out the risk premium, we arrive at a value of £876,000. This is not quite a reduction of 50%, in fact it is closer to 45%. So, does this mean that a figure of 1-45% = 55% is acceptable in every case? Unfortunately not. The reduction in premium as a percentage of the deductible is dependent upon the amount of the deductible, and the distribution of the underlying losses. The following table illustrates this:-

Deductible	Mean	Coefficient of Variation	Reduction	Reduction as a % of the Deductible
0	1,782,228	245%		
10,000	1,772,229	246%	-9,999	100%
50,000	1,732,802	252%	-49,426	99%
100,000	1,685,805	259%	-96,423	96%
250,000	1,561,251	279%	-220,977	88%
500,000	1,396,156	309%	-386,072	77%
1,000,000	1,161,418	365%	-620,810	62%
2,000,000	875,824	465%	-906,404	45%
4,000,000	584,620	643%	-1,197,608	30%
5,000,000	498,535	728%	-1,283,693	26%

139 Note also that while the mean drops slower than the deductible increases, the volatility drops at an even slower rate, and hence the coefficient of variation increases. This will impact on the premium charged as illustrated in Section B.

Section B: How does the profit margin required vary

140 Profit margin is often expressed as a percentage of premium. However, when aggregate features are involved the appropriate percentage of premium required as a profit load varies. By way of illustration, the profit margin required under three different pricing structures is examined.

141 *Method I*: Premium is expected cost of claims plus ½ a standard deviation of the cost of claims.

Deductible	Premiums	Profit Margin	Profit Loading	CofV for Writer
0	3,962,186	2,179,958	122%	200%
250,000	3,735,796	2,174,545	139%	200%
1,000,000	3,282,439	2,121,021	183%	200%
2,000,000	2,910,955	2,035,131	232%	200%
4,000,000	2,465,150	1,880,530	322%	200%
5,000,000	2,312,358	1,813,822	364%	200%

142 *Method II:* Premium is such that a 10% return on capital is required. Capital required is defined as the amount required such that losses do not exceed the premium received plus this capital more than 5% of the time.

Deductible	Premiums	Profit Margin	Profit Loading	CofV for Writer
0	2,220,320	438,092	25%	995%
250,000	1,996,700	435,449	28%	999%
1,000,000	1,565,000	403,582	35%	1051%
2,000,000	1,214,500	338,676	39%	1202%
4,000,000	767,900	183,280	31%	2052%
5,000,000	598,750	100,215	20%	3620%

143 *Method III*: Premium set such that the expected return on capital is 10%. The capital is defined to be the Value at Risk, calculated as the amount that losses exceed the premium multiplied by the probability of this occurring.

Deductible	Premiums	Profit Margin	Profit Loading	CofV for Writer
0	1,872,600	90,372	5.1%	4,722%
250,000	1,651,000	89,749	5.7%	4,743%
1,000,000	1,244,100	82,682	7.1%	5,034%
2,000,000	946,900	71,076	8.1%	5,640%
4,000,000	637,300	52,680	9.0%	7,071%
5,000,000	544,600	46,065	9.2%	7,813%

Points to Note:

144 The choice of 10% as the required rate of return in methods II and III ignores calculations based on diversification benefits, risk free rates of return, and the beta of the policy. The choosing of the required rate of return is an important calculation in itself. This example has simply assumed that 10% was the right rate.

145 In method II the profit loading initially increases then decreases, even though the risk is continually increasing. Intuitively, this seems to be an anomaly. Also, under this approach, eventually the deductible will be greater than the 95th percentile, and so the capital required at this point will be zero. This will imply that no capital is required and hence no profit loading.

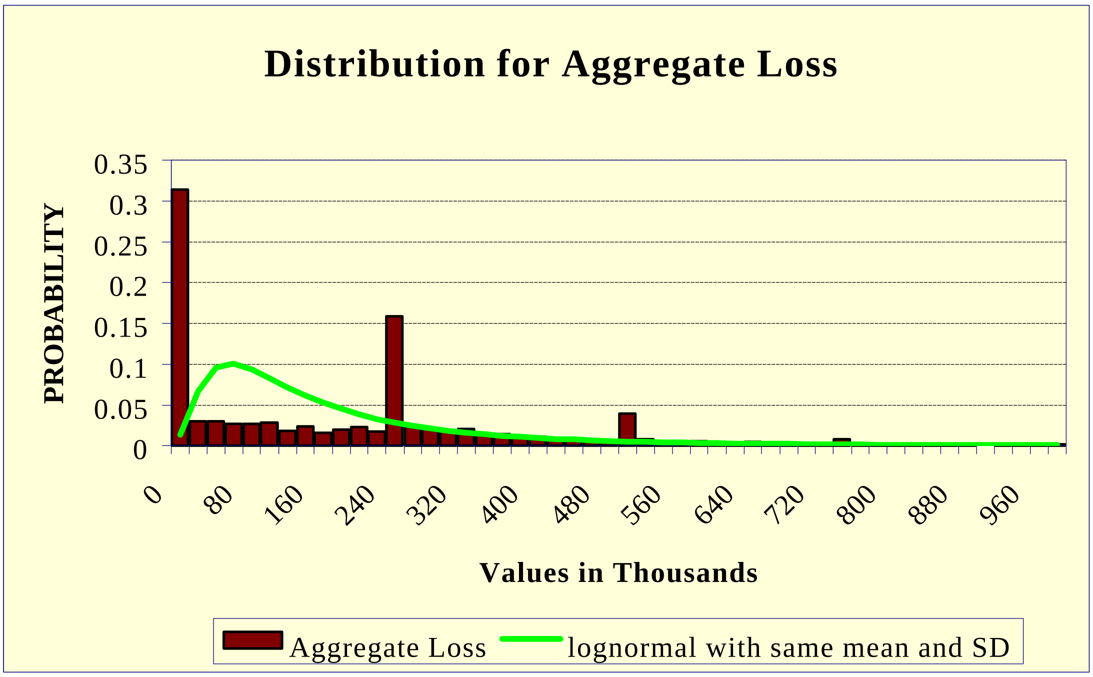
146 The capital required under the Value at Risk approach in method III is quite low, and hence the profit loading is much lower than under the other two methods. Also, the resulting coefficient of variation is particularly high.

Pitfall 7 – Using the wrong Aggregate Claims Distribution

Aggregate Loss Distributions: Log Normal?

147 A common assumption when pricing aggregate feature is to assume the aggregate distribution follows the lognormal or Pareto distributions. These distributions are chosen more for simplicity of calculation than any supposed theoretical justification.

148 Let's assume the insurance book of business produces 1,000 losses each year and each loss is distributed lognormally with mean £10,000 and standard deviation £100,000. We can calculate explicitly using computer simulation packages the aggregate distribution to a £250,000 xs £250,000 layer. The following chart shows the calculated aggregate distribution and compares it to a lognormal curve with the same mean and standard distribution:



149 The calculated distribution shows spikes at integer multiples of the limit (£250,000). These reflect the relative high likelihood of a total loss to the layer compared to a partial loss. The lognormal curve bears little resemblance to the actual distribution.

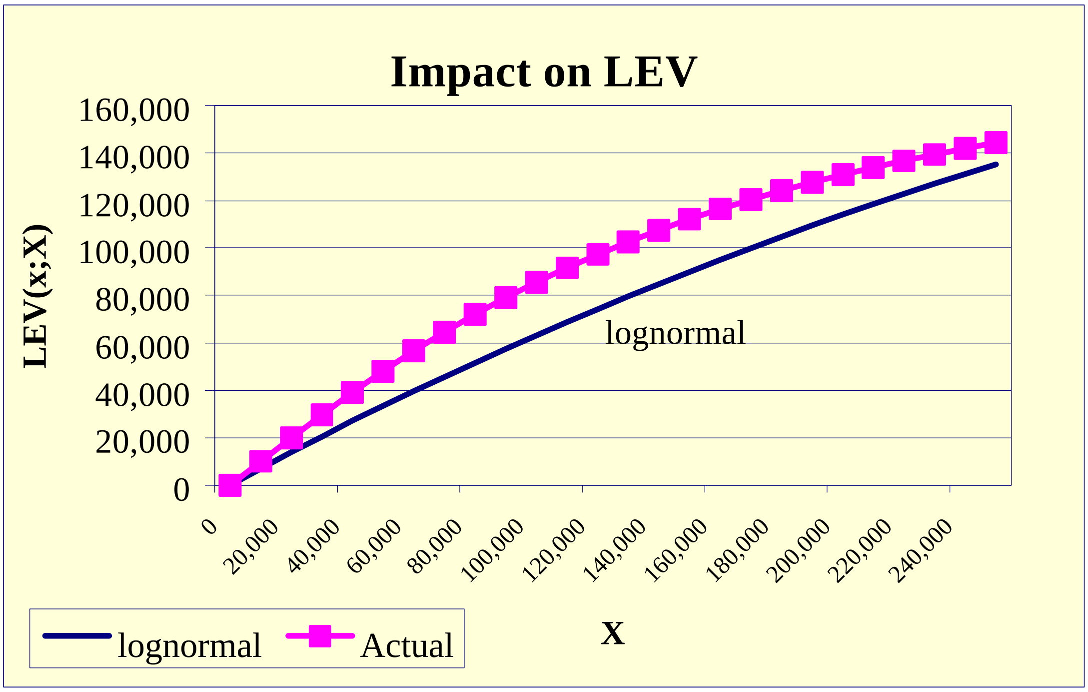
150 Does this mean that any aggregate feature priced using this assumption is wrong? In short the answer is yes, but depending on the actual feature being priced the impact can go from the insignificant to the dramatic.

151 The biggest difference is the probability of a zero claim. The actual probability is 28% but the probability of the lognormal being zero is 0%. You should not be using the lognormal to price such aggregate features as "if the treaty is claim free a no claims bonus will be paid".

152 The lognormal curve deviates significantly from the calculated curve over portions of the x-axis. Overall it averages out as the means of the two curves are the same. Any aggregate feature that makes use of a large part of the x-axis will not therefore be too distorted.

153 For instance E(x;£250,000) is £135,000 for the calculated curve and £144,000 for the lognormal: a difference of some 7%. However E(x;£10,000) is £7,095 for the calculated and £9,996 for the lognormal – a difference of 40%. More importantly E(x;X) is always higher for the actual curve than it is for the lognormal curve. This means that aggregate caps and reinstatement premiums are under priced using a lognormal approximation whilst contracts with AAD's are over priced.

154 The following chart compared the LEVs of the calculated and lognormal curves:



155 In the past making assumptions about the shape of the aggregate loss distribution was a necessary simplification to enable an answer to be reached in a reasonable timeframe. These days faster computers mean that this assumption is increasingly unnecessary and the additional uncertainty introduced into the answer is no longer justifiable.

Appendix 1 – Use of Panjer Recursion

A Description of Panjer Recursion

156 The Panjer Recursion formula is a fast recursive algorithm, which can be used to approximate an aggregate claims distribution. It can be programmed to form part of a reinsurance pricing system.

157 Key Assumptions:

(a) The distributions of the number of claims and claim amounts are known

(b) The claim amounts distribution is a discrete distribution on the non-negative integers

158 Notation:

(a) Let N be a random variable relating to the number of claims and p_r = Prob(N=r)

(b) Let X be a discrete random variable relating to individual claim amounts on the non-negative integers and f_j = Prob(X=j)

(c) Let X_1, X_2, X_3, …. be mutually independent individual claim random variables from the same distribution as X

(d) Let f^n_j = Prob(X_1+X_2+…..+X_n=j)

(e) Let S be a discrete random variable relating to the aggregate claims formed by compounding N and X and g_k = Prob(S=k)

159 To be able to apply Panjer recursion there must exist numbers a and b such that:

$$p_r = (a + b/r) \times p_{r-1} \text{ for } r = 1, 2, 3, \ldots..$$

160 There exist such numbers a and b for many of the common claim number distributions. For example, for Poisson(m), $a = 0$ and $b = m$ and for Negative Binomial (p,k), a = 1-p and b = (1-p)*(k-1).

161 We can generalise the standard Panjer formulae to allow for the possibility of claims of size zero. The recursive formulae can be shown to be:

$$g_0 = \sum_{r=0}^{\infty} p_r f_0^r$$

$$g_r = \frac{\sum_{j=1}^{r} (a + {bj}/{r}) \times f_j g_{r-j}}{(1 - a \times f_0)}$$

162 The formula for g_0 can be simplified to a function of f_0 for the common claim number distributions. For example, for Poisson(m) $g_0 = \exp[m(f_0-1)]$ and for Negative Binomial (p,k) $g_0 = [p/(1-(1-p)*f_0)]^k$.

163 The value for g_r is recursively calculated from the values of g_{r-1}, g_{r-2}, …., g_0. For example, for Poisson(m)

$$g_r = \sum_{j=1}^{r} ({mj}/{r}) * f_j g_{r-j}$$

Use of Panjer Recursion in Reinsurance Pricing

164 Panjer recursion is a useful way of calculating the aggregate claims to a reinsurance treaty and hence, when pricing the treaty, of allowing for features which depend on the amount of aggregate claims.

165 Common aggregate features include:

(a) Annual Aggregate Deductibles

(b) Annual Aggregate Limits

(c) Reinstatement Premiums

(d) Swing Rated Contracts

166 The effect of such aggregate features on the price of a reinsurance treaty can often be properly assessed only with the calculation of the distribution of aggregate claims to the treaty before application of those features.

167 There are several methods of assessing the aggregate claims distribution. Panjer recursion is often used because it is relatively fast to compute. Other methods of allowing for aggregate features, which are commonly used in practice, include:

(a) Simulation

(b) Approximations using simplifying assumptions as to the frequency and/or severity distributions (eg assuming total losses only)

Pitfall with Experience Rating

168 Experience rating models such as Burning Cost models do not ordinarily allow for aggregate features in a theoretically correct way. This can cause both under and over pricing of a layer with aggregate features. Adjustments can be made to the experience rating models to allow for any changes in the amounts of exposure from year to year, but still without an assessment of the aggregate claims distribution the answers are theoretically flawed. For example, in the following more extreme cases:

(a) If the aggregate limit to be priced has never been exceeded in past years for which claims experience exists, then a simple burning cost model would make no deduction in the rate of the layer for the fact that an aggregate limit was being proposed

(b) If an aggregate deductible is being proposed which has never before been exceeded then a zero rate would result from a simple burning cost model, with no allowance for the possibility that claims could exceed the aggregate deductible in the future.

Process Overview

169 We want to calculate the distribution of losses to the reinsurance treaty allowing for any aggregate features. In particular, we want to derive the expected aggregate losses to the treaty and the variance of those losses. First, it is necessary to calculate the distribution of the aggregate losses to the layer ignoring aggregate features before then making allowance for any such features.

170 If, as is usually the case, the individual claim amount distribution is not discrete on the non-negative integers, then a process of discretisation and rescaling is required so that Panjer recursion may be used.

171 The overall process can be summarised as follows:-

(a) Approximate the continuous claim amount distribution by a discrete distribution, and truncate the distribution so that it represents losses to the individual excess of loss layer.

(b) Calculate the aggregate losses to the excess of loss layer (before allowance for any aggregate features) using Panjer recursion and calculate the expected value and variance of those losses.

(c) Calculate the expected value and variance of the aggregate losses to the layer after allowing for any aggregate features.

Process Detail

172 Suppose we wish to price an individual excess of loss treaty with excess point v_1 and "limit" v_2 (i.e. providing v_2–v_1 cover for each individual loss) and including an aggregate deductible h_1 and an aggregate "limit" of h_2 (note that here we use the term "limit" to refer to the limit above zero and not the more usual limit of the cover above the deductible which for the aggregate cover here equals h_2-h_1). We will refer to v_1 to v_2 as the vertical layer and h_1 to h_2 as the horizontal layer.

173 Notation:

(a) Let X represent the continuous distribution of the FGU individual loss amounts with probability density function $f(x)$ and cumulative distribution function $F(x)$

(b) Let Y represent the discrete distribution of individual losses to the vertical layer, which is approximated from X

(c) Let N represent the distribution of the claim frequency adjusted to allow for the expected amount of exposure underlying the proposed reinsurance treaty

(d) Let A represent the discrete distribution of aggregate losses to the vertical layer v_1 to v_2

(e) Let B represent the discrete distribution of aggregate losses to the vertical layer after allowing for the aggregate features h_1 and h_2 (ie. restricted by the horizontal layer)

Discretisation of the continuous FGU Individual Loss distribution

174 The aim is to calculate a discrete distribution of individual losses to the vertical layer such that the expected value and variance of the continuous distribution (when applied to the vertical layer) are preserved.

175 Select an even number of discretisation bands (we denote by n) and hence an odd number of discrete probability mass points (n+1) over which to perform the discretisation. Sensitivity testing on the variability of the results to the number of discretisation bands indicates that for most practical cases the number of bands can be small (less than 100) without greatly effecting the accuracy of the results.

176 We have n+1 points of probability mass, which for Panjer recursion purposes are assumed to represent the non-negative integers (0, 1, 2, …., n). We let the first point 0 of the discrete distribution relate to no loss to the vertical layer and let the last point n relate to a total loss to the vertical layer of amount v_2-v_1. The points in between relate to losses that are spaced evenly at intervals equal to $(v_2-v_1)/n$.

177 We initially set the probability mass at point 0 equal to the probability that an individual FGU loss is less than v_1 (i.e. $F(v_1)$) and the mass at the last point n equal to the probability that an individual FGU loss is greater than v_2 (i.e. $1-F(v_2)$).

178 For the probability mass of individual FGU losses between v_1 and v_2 we assign the probability to the points of the discrete distribution in such a way as to retain the expected value and variance of the original continuous distribution when applied to the vertical layer.

179 We can discretise the individual loss distribution over successive sets of 3 points (i.e. first between points 0,1 & 2 and then 2,3 & 4 and so on).

180 Let t, t+1 and t+2 be three consecutive points on the discrete distribution relating to individual FGU losses of size $a = v_1 + t*(v_2-v_1)/n$, $b = v_1 + (t+1)*(v_2-v_1)/n$ and $c= v_1 + (t+2)*(v_2-v_1)/n$ respectively. We assign probability masses (p_a, p_b and p_c) to each of these 3 points so as to equate the total mass, expected value and variance of the discrete distribution with that of the continuous distribution over these 3 points. This is equivalent to insisting that every quadratic have the same expectation under the continuous distribution and the discrete approximation.

181 This can be achieved by choosing quadratics that vanish at two out of three sample points and results in the following allocation of probability mass:-

$$p_a = \frac{1}{(c-a)^2}\int_a^c (c-x)(a+c-2x)\,f(x)\,dx$$

$$p_b = \frac{4}{(c-a)^2}\int_a^c (c-x)(x-a)\,f(x)\,dx$$

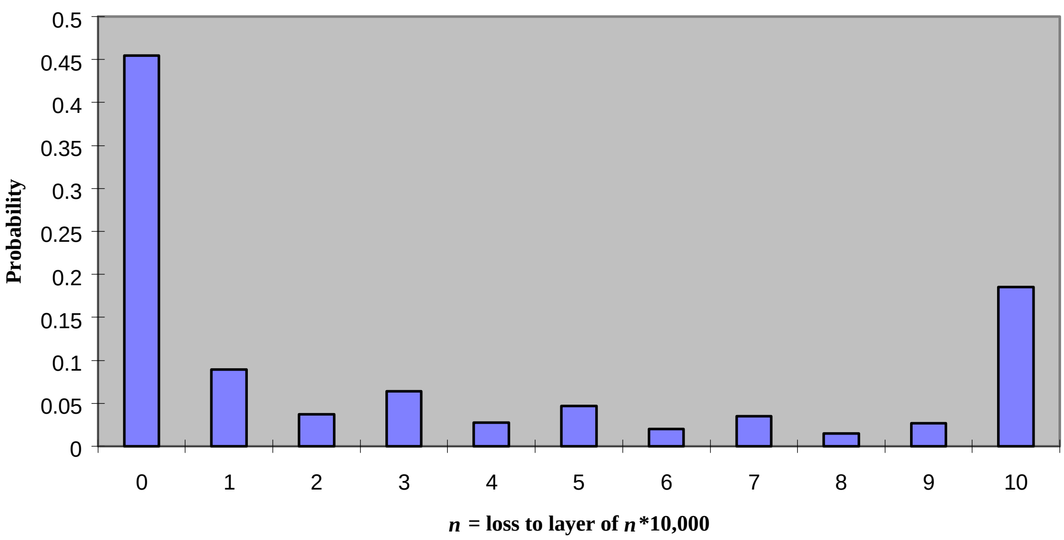
$$p_c = \frac{1}{(c-a)^2}\int_a^c (x-a)(2x-a-c)\,f(x)\,dx$$

182 Formulae that calculate these amounts for common severity distributions (such as Pareto and lognormal distributions) can be readily derived and programmed.

183 Applying this for all successive sets of 3 points and adding up the probability masses at the joins (i.e. at the even points) gives the final discretisation.

184 We then have a discrete distribution of individual losses to the vertical layer represented by points (0, 1, 2, …, *n*).

Discrete Distribution of Individual Loss to Vertical Layer (*Y*)



Panjer Recursion

185 We can then apply Panjer Recursion, using this discrete distribution of individual losses to the vertical layer and our frequency distribution *N*, to calculate the expected value and variance of the aggregate losses to the vertical layer.

186 The points (0, 1, 2,…, n) of each individual loss to the vertical layer represent losses at evenly spaced $(v_2-v_1)/n$ sized intervals. Hence, if an aggregate limit of h_2 is proposed, then we only need to apply Panjer recursion up to a point '*maxpanjer*', where:-

> *maxpanjer* = the nearest integer value to $h_2/((v_2-v_1)/n)$, which relates to an aggregate loss to the vertical layer v_1 to v_2 of h_2

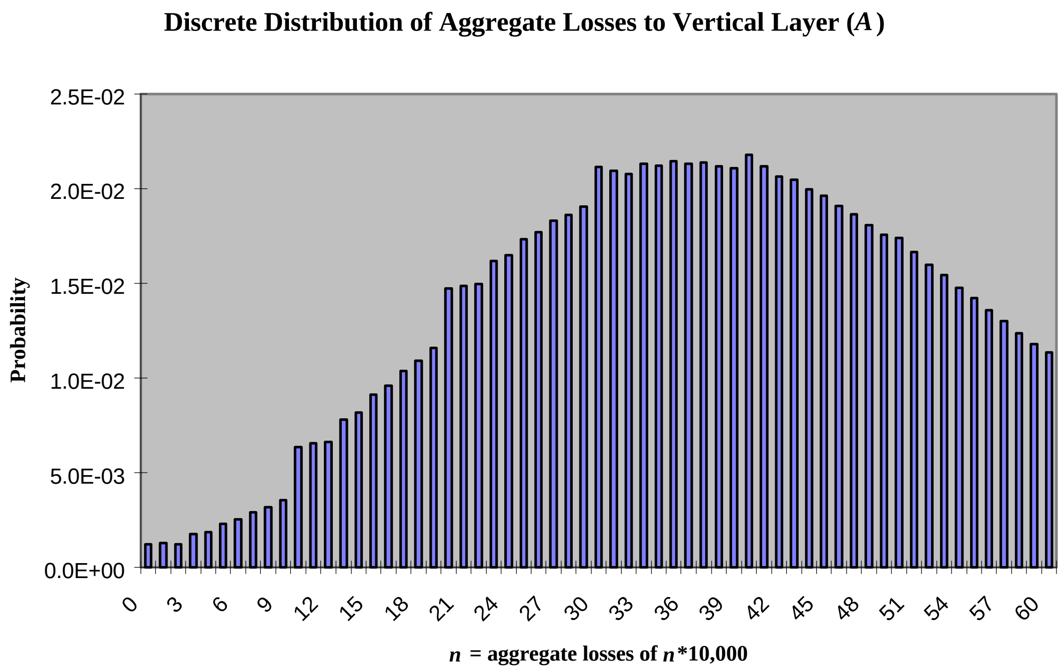
187 The aggregate loss to the vertical layer distribution (A) is thus defined on the *maxpanjer*+1 points (0, 1, 2,…, *maxpanjer*) relating to aggregate losses to the vertical layer, v_1 to v_2, of 0 to h_2 evenly spaced at $(v_2-v_1)/n$ sized intervals.

188 We can also calculate the expected value and the variance of this compound distribution of aggregate losses to the vertical layer. This will allow us to avoid infinite sums when pricing contracts with no annual aggregate limit. We can use the standard formulae:

(a) $E(A) = E(N)*E(Y)$

(b) $Var(A) = E(N)*Var(Y) + Var(N)*E(Y)^2$

189 Where $E(Y)$ and $Var(Y)$ are easily calculated from either the discrete distribution Y or from the continuous FGU distribution applied to the vertical layer.



Expected value and variance of the losses to the treaty allowing for aggregate features

190 We know the expected value E(A) and variance Var(A) of the aggregate losses to the vertical layer, which can be calculated analytically. We also have a discrete approximation to the distribution of A, calculated as high as we want, but not to infinity.

191 The problem is to find expressions for the expected losses, and variance of losses to the vertical layer subject to an aggregate deductible h_1 and aggregate limit h_2. The aggregate losses restricted by the horizontal layer, B, can be calculated as:

$$B = \min\{h_2 - h_1, \max\{A - h_1, 0\}\}$$

192 We must also allow for the possibility that h_2 could be infinite (i.e. there is no aggregate limit).

193 If we define functions $q_1(x)$ and $q_2(x)$ by:

$$q_1(x) = \mathrm{E}(\min\{A, x\})$$
$$q_2(x) = \mathrm{E}(\min\{A^2, x^2\})$$

194 In each case, these can be calculated given the distribution of A on the range (0,x).

195 The tail calculation is not required. Notice that in limiting cases, we have:

$$q_1(0) = 0$$
$$q_2(0) = 0$$
$$q_1(\infty) = \mathrm{E}(A)$$
$$q_2(\infty) = \mathrm{E}(A)^2 + \mathrm{Var}(A)$$

196 Let us now consider the horizontal layer from h_1 to h_2. The aggregate losses restricted by this horizontal layer can be rewritten as:

$$B = \min\{A, h_2\} - \min\{A, h_1\}$$

197 and so the expected losses are readily seen to be:

$$\mathrm{E}(B) = q_1(h_2) - q_1(h_1)$$

198 We can also work out an expression for the squared aggregate losses allowing for the horizontal layer and hence the variance of those losses:

$$B^2 = \min\left\{(h_2 - h_1)^2, \max\{A - h_1, 0\}^2\right\}$$
$$B^2 = \min\{A^2, h_2^{\ 2}\} - \min\{A^2, h_1^{\ 2}\} - 2h_1\left[\min\{A, h_2\} - \min\{A, h_1\}\right]$$

199 and so we have:

$$\mathrm{E}\left(B^2\right) = q_2(h_2) - q_2(h_1) - 2h_1\left[q_1(h_2) - q_1(h_1)\right]$$

200 and obviously Var(B) =E(B^2) – E(B)^2.

201 In each case, of course, we apply the limiting value when h_2 is infinite.

Discrete Distribution of Aggregate Losses to Layer (allowing for aggregate deductible and limit)