# On Predictive Modeling for Claim Severity

Glenn Meyers, FCAS, MAAA, Ph.D.

**Abstract:**

Reinsurers typically face two problems when they want to use insurer claim severity experience to experience rate their liability excess of loss treaties. First, the claim severity data has insufficient volume to make credible projections of excess layer costs. Second, the data they do receive is not fully developed. Most claims that pierce the excess layers can take at least a few years to settle. This paper sets forth a methodology for dealing with these problems. The paper starts with some introductory examples that illustrate how to quantify the inherent uncertainty in fitting claim severity distributions. Then the paper illustrates a Bayesian methodology to estimate the expected cost of excess layers, and shows how to quantify the uncertainty in these estimates. The Bayesian "prior models" are not derived from purely subjective considerations. Instead they are derived after examining the claim severity data of several insurers. Each "prior model" contains claim severity distributions of immature data that are used to calculate the posterior probabilities with comparable immature data submitted by an insurer. Each "prior model" also contains a fully developed claim severity distribution. The estimate of the cost of an excess layer is the average of the fully developed excess layer costs weighted by the posterior probabilities calculated with the immature data submitted by the insurer.

**Keywords:** Loss Distributions, Bayesian Estimation, Excess of Loss Reinsurance

## 1. INTRODUCTION

One of the many jobs an actuary is asked to do is to predict future claim costs in high layers. It is often the case that there are few claims from past experience. When this is the case, an actuary must resort to either one or both of the following.

- Try to discern a pattern in the claims that lie below the layer, and use this pattern to project claim costs in the layer. This is usually done by fitting a parametric probability distribution to these other claims.
- Examine claims from other insurers, or from an industry source, in the hope that these claims are similar to the claims you are trying to project. Often an actuary will make use of a probability distribution that has been fit to these claims.

There are difficulties with each of these approaches. The first approach can have credibility problems if there are not enough claims to get a reliable estimate of the parameters of the parametric probability distribution. And identifying the best distribution can also be a problem. The second approach can have relevance problems if the population that underlies the "industry" is different than the population that the actuary is addressing.

Liability insurance presents yet another problem in fitting claim severity distributions. It can take a considerable amount of time to settle the claims. A changing legal environment will force the actuary to compromise between completeness and relevance of the claim information.

This paper will address each of these problems. Here is a summary of what is to follow.

- I will begin with a description of how to construct a classical (non-Bayesian) confidence interval of parameters of a claim severity distribution using the likelihood ratio test.
- Next I will show how to use Bayes' Theorem to calculate posterior probabilities for a series of selected claim severity distributions. The "selected claim severity distributions" can come from different parameterizations of a selected model, such as a Pareto or a lognormal model. Or the "selected claim severity distributions" can come from different models. This allows us to incorporate what we informally call "model uncertainty" as well as "parameter uncertainty" into our estimation procedure.
- It is generally the case that particular claim severity models, or particular parameterizations of these models, are not of direct interest. What are of interest are functions of the models and their possible parameterizations. An example of such a function would be the expected cost of a particular layer of loss. I will show how to quantify uncertainty in the expected cost of a layer of loss in terms of the posterior probabilities of each of the models and their multiple parameterizations.
- It is possible to associate claim severity distributions developed to their ultimate value, with the immature claim severity distributions representing the data that is currently available. By fitting (i.e., determining the posterior probabilities) the immature data to the immature distributions and then applying the posterior weights to the associated fully developed distributions, it is possible to get estimates of the expected losses for a layer of loss. Furthermore, one can quantify the uncertainty in this estimate.

A major theme of this paper will be the importance of the likelihood function. Loosely stated, the likelihood function is the probability of observing a given set of data as a function of a parametric probability distribution. The likelihood function will play a key role in both the classical and Bayesian methodologies described below.

## 2. CONFIDENCE REGIONS FOR PARAMETERS

I will begin the discussion of confidence regions with a description of hypothesis testing using the likelihood ratio test.

Let:

- $\mathbf{p} = (p_1, p_2, \ldots, p_k)$ be a parameter vector for a given parametric probability distribution;
- $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ be a set of observed losses;
- $\mathbf{p}^{ML}$ be the maximum likelihood estimate of the parameter vector given the data $\mathbf{x}$;
- $\mathbf{p}^{T}$ be the "true" parameter vector underlying the population of interest; and
- $\mathbf{p}^*$ be a parameter vector for a proposed model for a claim severity distribution.

Denote likelihood of a parameter vector, $\mathbf{p}$, given the data, x, by $L(\mathbf{x};\mathbf{p})$.

We want to test the null hypothesis:

$$H_0: \mathbf{p}^* = \mathbf{p}^T;$$

against the alternative hypothesis:

$$H_1: \mathbf{p}^* \neq \mathbf{p}^T.$$

**Theorem**

If $H_0$ is true, then the statistic:

$$\ln(LR) \equiv 2 \cdot \left( L\left(\mathbf{x};\mathbf{p}^{ML}\right) - L\left(\mathbf{x};\mathbf{p}^*\right) \right)$$

has a $\chi^2$ distribution with $k$ degrees of freedom.

This theorem is given in Section 13.4.4 of Klugman, Panjer and Willmot (KPW) [2004][1].

Informally, this theorem says that one should accept (or fail to reject) the hypothesis that $\mathbf{p}^*$ is the parameter vector for the population if the likelihood of $\mathbf{p}^*$ is sufficiently "close" to

[1] There are a number of technical conditions placed on the probability distribution for this result to hold. Most of the common distributions used by actuaries (such as Pareto, lognormal and gamma distributions) satisfy these conditions.

the maximum likelihood estimate, $\mathbf{p}^{ML}$, of the sample. More formally, "close" is defined by the above statistic and the critical values of the $\chi^2$ distribution with $k$ degrees of freedom.

To illustrate the likelihood ratio test I took a random sample of 1,000 claims from a Pareto distribution of the form

$$F(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^{\alpha},$$

with $\alpha = 2$ and $\theta = 10{,}000$.

While it is not convenient to list all 1,000 claims in this random sample, here is a grouped summary of these claims.

**Table 1**

| Range | Count |
|---|---|
| $x_i \leq 5{,}000$ | 562 |
| $5{,}000 < x_i \leq 10{,}000$ | 181 |
| $10{,}000 < x_i \leq 20{,}000$ | 134 |
| $20{,}000 < x_i$ | 123 |

Here is the log-likelihood function for the grouped data.

$$l_G(\theta,\alpha) = 562 \cdot \ln\left(1 - \left(\frac{\theta}{\theta+5000}\right)^{\alpha}\right) + 181 \cdot \ln\left(\left(\frac{\theta}{\theta+5000}\right)^{\alpha} - \left(\frac{\theta}{\theta+10000}\right)^{\alpha}\right)$$
$$+134 \cdot \ln\left(\left(\frac{\theta}{\theta+10000}\right)^{\alpha} - \left(\frac{\theta}{\theta+20000}\right)^{\alpha}\right) + 123 \cdot \ln\left(\left(\frac{\theta}{\theta+20000}\right)^{\alpha}\right)$$

Using a general purpose maximizing tool, Excel Solver, I found the maximum likelihood estimate of the Pareto parameters for the grouped data to be equal to $\left(\theta_G^{ML}, \alpha_G^{ML}\right) = (7447.8, 1.6041)$.

Here is the log-likelihood function for the detailed data.

$$l_D(\theta,\alpha) = 1000 \cdot \left(\ln(\alpha) + \alpha \cdot \ln(\theta)\right) - (\alpha+1)\sum_{i=1}^{1000} \ln(x_i + \theta).$$

Using Excel Solver, I found the maximum likelihood estimate of the Pareto parameters for the detailed data to be equal to $\left(\theta_D^{ML}, \alpha_D^{ML}\right) = (9626.8, 1.8079)$.

Note that the parameter estimates in each case are different from the true parameters that I used to generate the simulated data. If I were to generate another simulation, I would get a different parameter estimate. Repeated simulations will yield samples from a bivariate distribution of parameter estimates. There is a formula that describes the bivariate distribution of the maximum likelihood estimates in terms of the parameters that are used to generate the original distribution.[2]

Our problem is different. In practice, we don't know the parameters of the underlying distribution.[3] A question to ask is the following: What are acceptable parameters for the distribution given the data we have? To answer this question one can invoke the likelihood ratio test by defining a *p*% confidence region of the parameters as the set of all parameters that pass the likelihood ratio test at the *p*% level.

Figures 1 and 2 are plots of the parameters that pass the likelihood ratio test at the 5% level for the grouped and detailed data, respectively. These plots were generated by calculating the likelihood ratio statistic for a grid of $(\theta,\alpha)$ points and plotting them if the statistic was less than the 5% critical value, 5.99, of the $\chi^2$ distribution with two degrees of freedom.

It is worth noting that the confidence region is wider for the grouped likelihood data than for the detailed likelihood data. This illustrates the additional information provided by the detailed data.

---

[2] The distribution of the maximum likelihood estimates has an asymptotic normal distribution with parameters given by the Fisher Information Matrix. See Section 12.3 of Klugman, Panjer and Willmot [2004].

[3] In practice, we don't even know the underlying distribution itself. I will get to that below.

Figure 1
Confidence Region
Grouped Likelihood

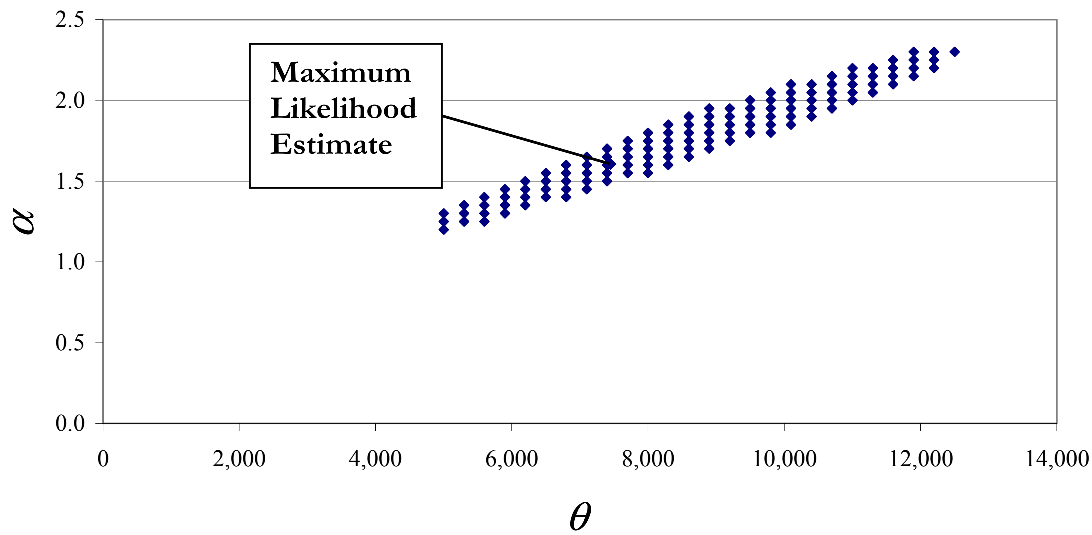
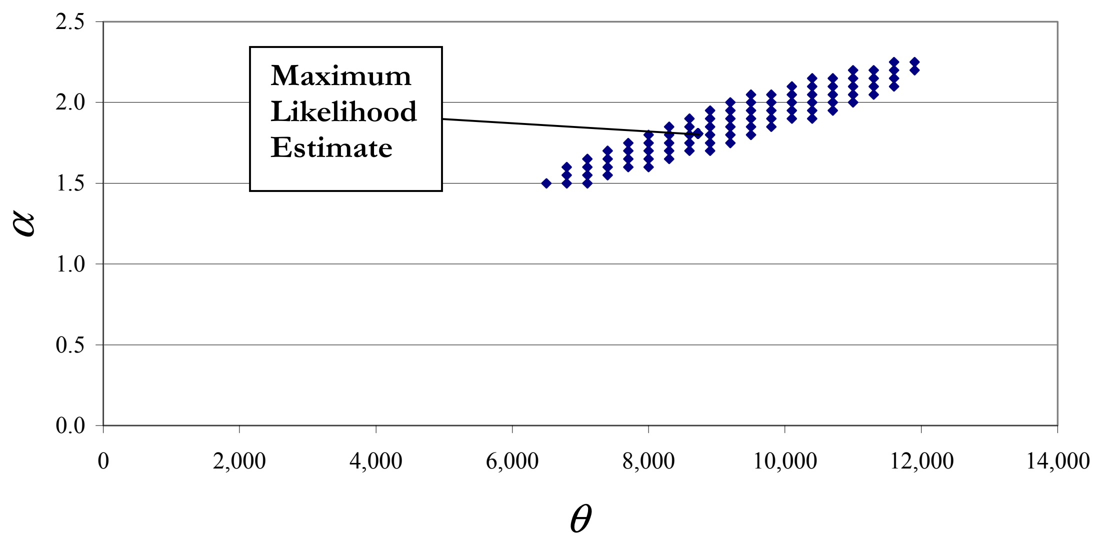



Figure 2
Confidence Region
Detailed Likelihood



This definition of a confidence region is somewhat unusual. The standard technique (described in KPW, Section 12.3) is to use the Fisher information matrix, substituting the maximum likelihood estimate of the parameters for the "true" parameters. However, the definition of confidence regions of the parameters used here has precedent. One of these is in the first edition of KPW[4].

[4] See Example 2.69 on page 131 of Klugman, Panjer and Willmot [1998]. I asked Professor Klugman why this example was not in the 2nd edition. He replied that it was an oversight, and made a note to put it back in the third edition.

## 3. THE COTOR CHALLENGE

Last year, the CAS Committee on the Theory of Risk (COTOR) issued a challenge. The committee published a list of 250 claims, and asked contestants to estimate the pure premium of a $5 million x $5 million layer. An additional requirement of the challenge was to put a 95% confidence interval around this estimate. A full description of the COTOR Challenge can be found on the CAS web site:

http://www.casact.org/cotor/round2.htm.

The "claims" were generated by a simulation from a transformed inverse gamma distribution, a fact that was not revealed until after all solutions were submitted.

The COTOR challenge has some of the elements that reinsurance actuaries face in their work. Most importantly:

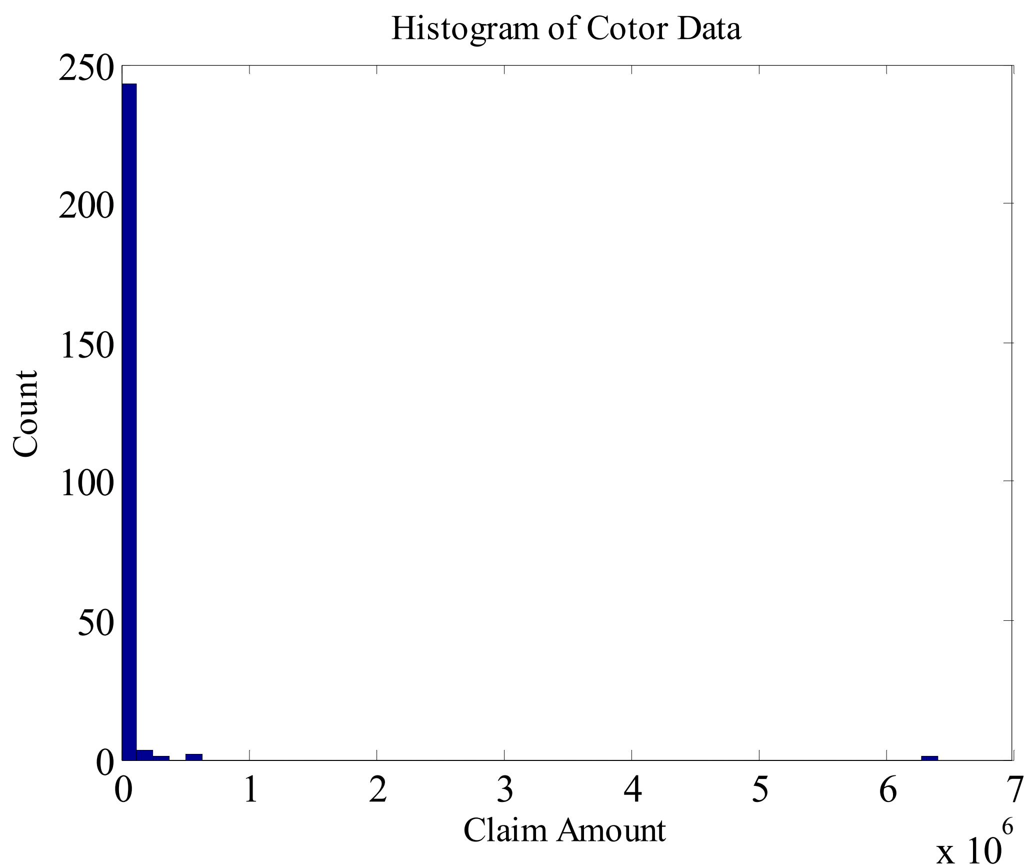
- The underlying loss distribution is unknown, and is very likely not one of the standard models that are in the typical actuarial toolbox.
- There are very few claims in the layer of interest. Actuaries typically try to project the frequency of claims in a high layer by looking at claims in a lower layer.

My solution, which is posted on the COTOR web site, provides an example that I believe to be of educational value as we move toward the ultimate goals stated in the introduction. I will describe it in some detail here.

The solution makes use on a software package called MATLAB. The software has tools for plotting histograms, calculating maximum likelihood estimates, and supporting statistics.

The first step one should almost always take when fitting a distribution to data is to plot a histogram.

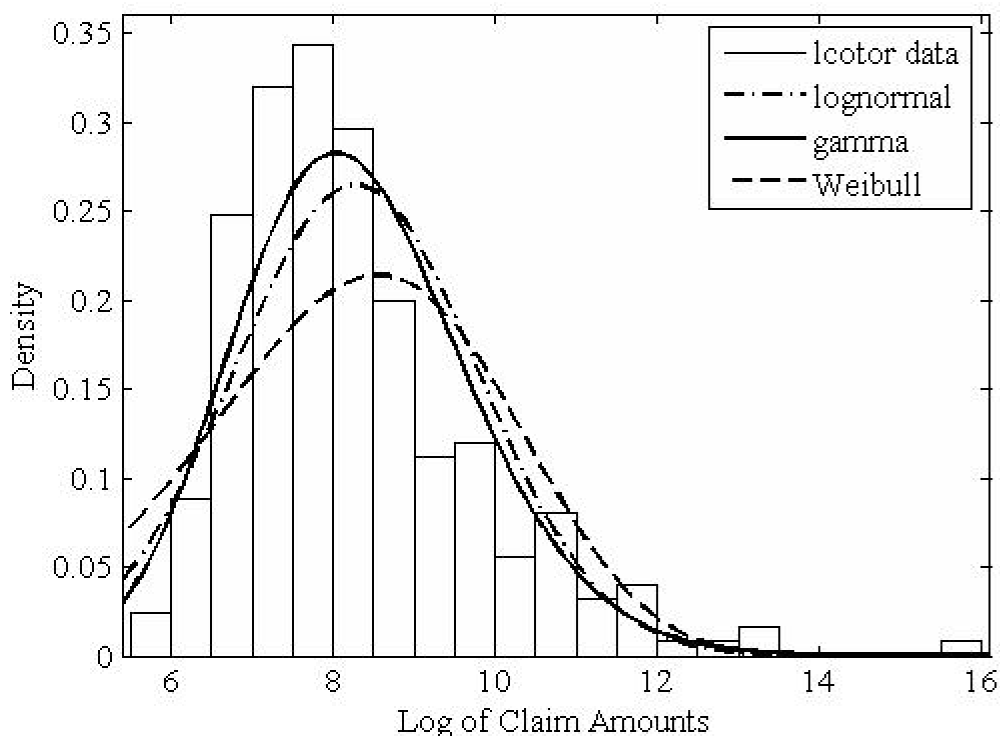
**Figure 3**




Note that there is only one claim in the $5 million x $5 million layer that contestants were asked to predict. The next highest claim was about $600,000. I did not even attempt to fit a distribution to this data.

The next step I took was to take the log of that data. The histogram looked promising so I tried to fit a couple of different distributions by maximum likelihood.

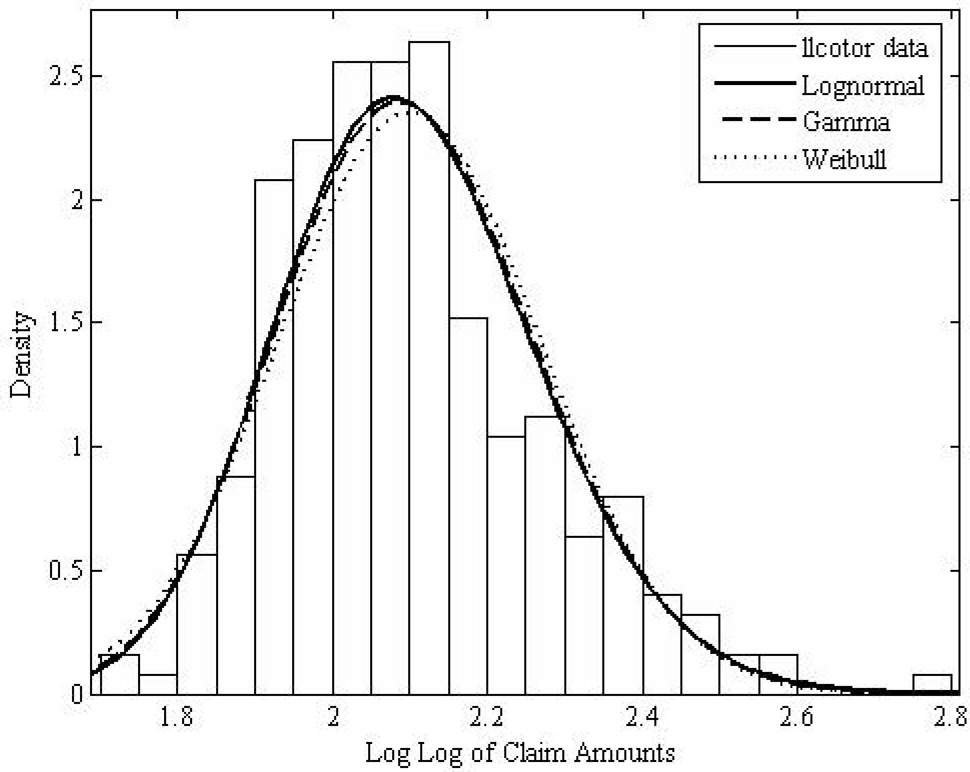
**Figure 4**




None of the selected distributions looked particularly good, with the Weibull providing the worst fit. The distribution was still skewed to the right.

In an attempt to reduce the skewness, I plotted a histogram of the double log of the data and fit some distributions to the double logged data by maximum likelihood.

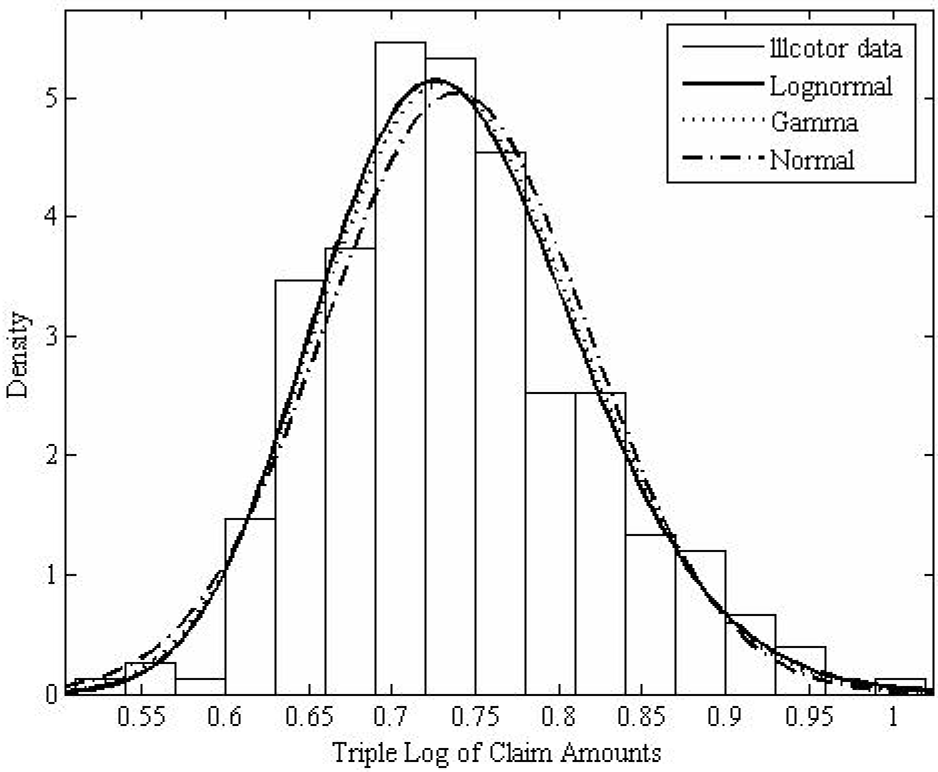
**Figure 5**




Here the fit of the three distributions is closer, but the double log of the data still looks more skewed than the distributions I tried.

Continuing the above logic, I took the triple log of the data and tried some more maximum likelihood fits.

**Figure 6**




At this stage, the maximum likelihood fits began to look reasonable. I examined these fits in more detail.

Here are some fitting statistics for the three distributions. Some observations:

- The lognormal distribution has the highest loglikelihood and the hence the best fit of the three.
- The loglikelihood of the gamma distribution is reasonably close to that of the lognormal distribution. The loglikelihood of the normal distribution is a bit lower, but not totally out of the running.
- Looking at Figure 7 below, we see that the maximum likelihood fit of all three distributions are nicely within the confidence bounds for lognormal fit.
- It is possible to go to a quadruple log transform since the triple logs of the claim amounts are still positive. But that is as far as we can go, since some triple logs are less than one. I stopped at the triple log transform since the lognormal is equivalent to a normal with the quadruple log transform.

**Table 2**

| | | |
|---|---|---|
| Distribution: | Lognormal | |
| Log likelihood: | 283.496 | |
| Mean: | 0.73835 | |
| Variance: | 0.00619 | |
| Parameter | Estimate | Std. Err. |
| $\mu$ | -0.30898 | 0.00672 |
| $\sigma$ | 0.10625 | 0.00477 |
| Estimated covariance of parameter estimates: | | |
| | $\mu$ | $\sigma$ |
| $\mu$ | 4.52E-05 | 1.31E-19 |
| $\sigma$ | 1.31E-19 | 2.27E-05 |

| | | |
|---|---|---|
| Distribution: | Gamma | |
| Log likelihood: | 282.621 | |
| Mean: | 0.73836 | |
| Variance: | 0.00615 | |
| Parameter | Estimate | Std. Err. |
| *a* | 88.6454 | 7.91382 |
| *b* | 0.00833 | 0.00075 |
| Estimated covariance of parameter estimates: | | |
| | *a* | *b* |
| *a* | 62.6286 | -0.00588 |
| *b* | -0.00588 | 5.56E-07 |

| | | |
|---|---|---|
| Distribution: | Normal | |
| Log likelihood: | 279.461 | |
| Mean: | 0.738355 | |
| Variance: | 0.006285 | |
| Parameter | Estimate | Std. Err. |
| $\mu$ | 0.738355 | 0.005014 |
| $\sigma$ | 0.079279 | 0.003556 |
| Estimated covariance of parameter estimates: | | |
| | $\mu$ | $\sigma$ |
| $\mu$ | 2.51E-05 | -1.14E-19 |
| $\sigma$ | -1.14E-19 | 1.26E-05 |

**Figure 7**

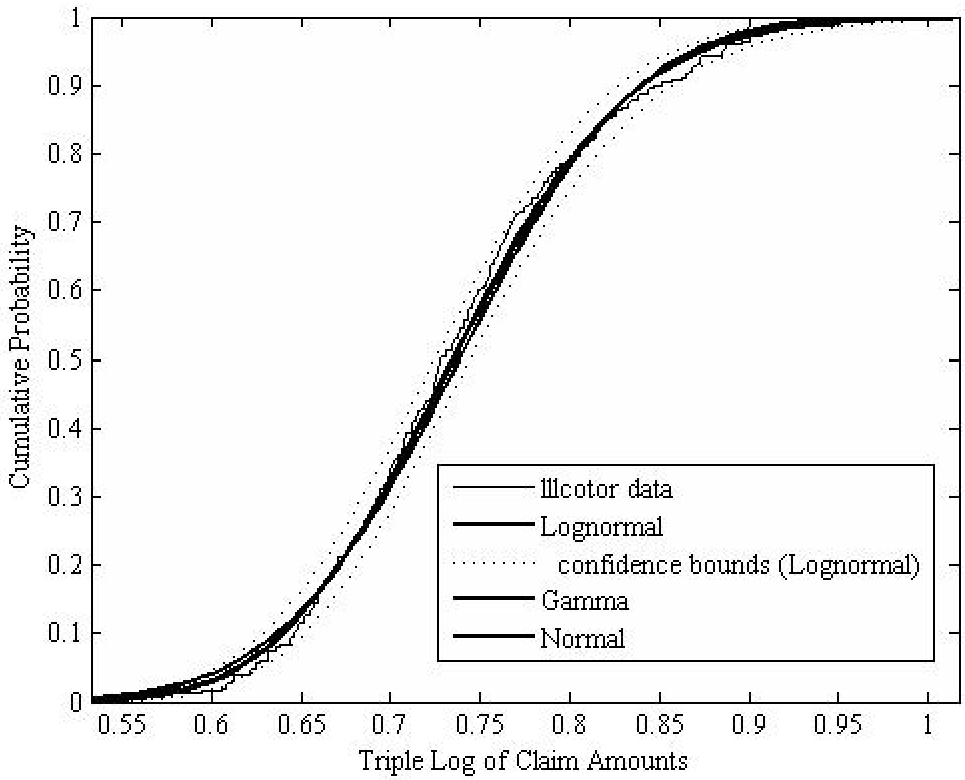



Figure 7 is a plot of the cumulative distribution functions of each of the fits and the data. The dotted lines give upper and lower confidence bounds for the best-fitting lognormal distribution. These confidence bounds contain the normal and gamma distributions and so we should consider all three of these distributions as potential models for the triple logs of the data.

Up to now, this analysis has been fairly classical. A typical classical analysis of the data would take the cumulative distribution function, $F(x)$, of the best-fitting model, (in this case the quadruple lognormal with the parameters in Table 2) and integrate the following formula to estimate the pure premium of the \$5 million x \$5 million layer.

$$\text{Layer Pure Premium}^5 = \int_{5,000,000}^{10,000,000} \left(1-F(x)\right)dx\,. \tag{1}$$

Such an estimate of the layer pure premium reflects the uncertainty of the loss given the model and the parameters of the model. It does not reflect the uncertainty of the model and the uncertainty in the parameters given the model.

If we are to reflect these additional uncertainties, we need to get the probability of the potential models and parameters. The only information we have to get these probabilities is the data. Now we have the ability to calculate the likelihood function (i.e., the probability of the data) for any given model and parameter set. To carry out this program, I made use of Bayes' Theorem to calculate the probability of each model and parameter set given the data.

Here is an outline of the methodology underlying my solution.

- I began by hypothesizing a series of "models" for the data. I interpret the term "model" broadly to include choices of the parametric form of the '*model*' (in the narrow sense; e.g., lognormal or gamma) as well as a choice of parameters for each '*model*.' I am intentionally blurring the distinction between parameter uncertainty and '*model*' uncertainty.
- For each model, I calculated the likelihood (or probability) of the data given each model. Using Bayes' theorem, the posterior probability of each model, given the data, is calculated by the following formula.

$$\text{Posterior}\left(\text{model}|\text{data}\right) \propto \text{Likelihood}\left(\text{data}|\text{model}\right)\cdot\text{Prior}\left(\text{model}\right).$$

- For each model, I calculated the cost of the \$5 million x \$5 million layer using Equation 1 above. I then calculated various statistics of the posterior distribution of the layer costs using posterior probabilities. For example:

[5] One should distinguish between the layer pure premium and the layer average severity. The layer average severity is the average severity given that the claim has pierced the layer. The layer pure premium is equal to the layer average severity times the probability of piercing the layer.

- The posterior expected cost of the layer was the posterior probability-weighted average of the layer cost for each model. Calculations such as this led to the mean and standard deviation in my solution.
- The posterior percentile of a selected layer cost is the sum of the posterior probabilities of all the models for which the layer cost is less than the selected layer cost. Calculations such as this led to the median and confidence interval for my solution.

Now let's look at the details.

The above analysis identified three potential '*models*' for the data with the triple log transform — the lognormal, the gamma and the normal. The fitting statistics give an indication of the range of possible parameters for each '*model*.'

Here are the steps in my calculations.

1. For each '*model*,' I calculated the confidence interval at the 0.1% level for both parameters.
2. I divided the confidence interval into 50 intervals and created a 51 by 51 grid of possible parameters for each '*model*.' The three '*models*' along with the 2,601 parameters yielded 7,803 "models" from which to do the Bayesian calculations.
3. I calculated the loglikelihood of each model. I assumed that the prior probabilities for each model were equal. I then exponentiated the posterior loglikelihoods and normalized them so that the posterior probabilities sum to one.
4. I then calculated the pure premiums for each model using Equation 1 and MATLAB's numerical quadrature function.
5. Finally I transferred the MATLAB arrays into Excel, sorted the models in increasing order of the pure premiums, and calculated the statistics reported in the results below.

The MATLAB code for executing the first four steps along with the spreadsheet for Step 5 can be downloaded from the COTOR web site.

I should point out that the '*model*' uncertainty did not have a significant effect on the final answer. The lognormal models got 95.33% of the posterior probability, the gamma models got 2.98% of the posterior probability and the normal models got the remaining 1.69%.

Here are the results.

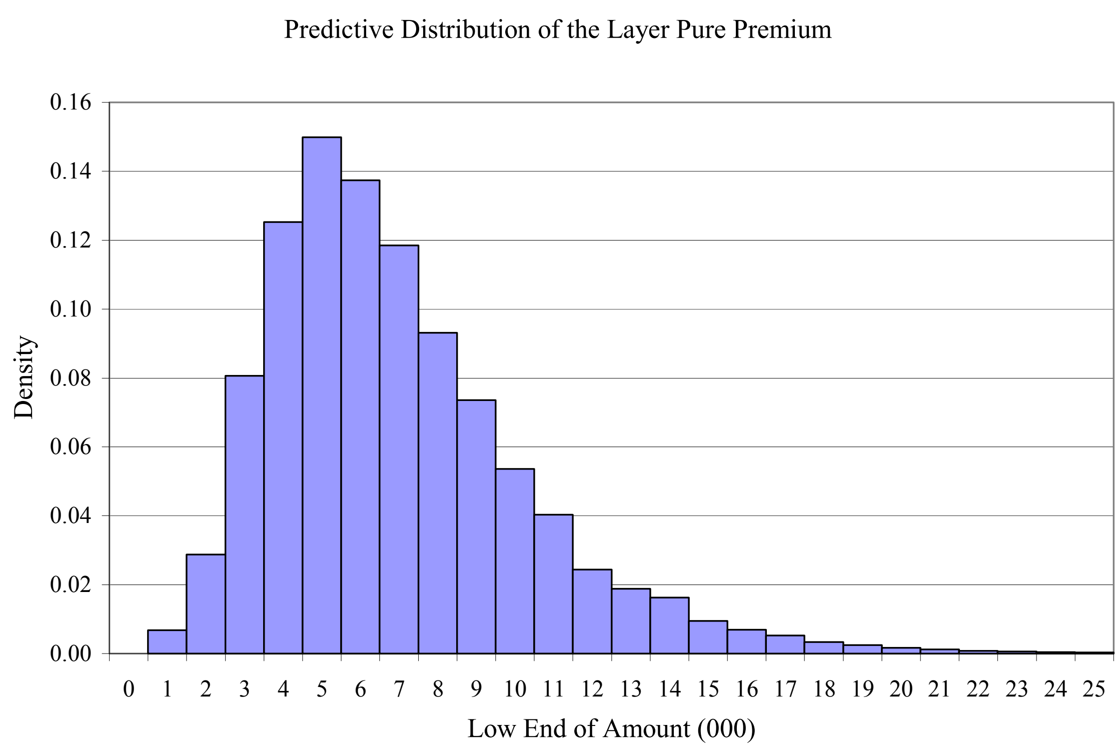
**Table 3**

**Predictive Statistics for the Layer Pure Premium**

| | |
|---|---|
| Mean | 6,430 |
| Standard Deviation | 3,370 |
| Median | 5,780 |
| Range | |
| Low at 2.5% | 1,760 |
| High at 97.5% | 14,710 |

Here is a histogram for the predictive distribution.

**Figure 8**



Using Bayes' Theorem in the solution above is similar to the likelihood ratio approach described in Section 2 in that both approaches use the likelihood function to identify potential models. The likelihood ratio test only provides you with a "yes/no" decision. And this "yes/no" decision is only correct if the underlying '*model*' is correct. But if you are comfortable with assigning prior probabilities to models, Bayes' Theorem allows you to use the likelihood function of the data associated with each model to calculate posterior probabilities for each model. And with probabilities assigned to each model, you can calculate any desired statistic of a function (e.g., layer pure premium) of the potential models. And the Bayesian approach can deal with '*model*' uncertainty.

Using Bayes' theorem to fit claim severity distributions is not new to the CAS literature. Meyers [1994], Klugman [1994], and Kreps [1997] have papers on this subject.

## 4. AN EXAMPLE BASED ON INSURANCE DATA

I believe the Bayesian methodology underlying the COTOR challenge is potentially useful for predicting pure premiums for high layers of insurance, but the methodology is far from complete. In this section, I will give an example that uses this Bayesian methodology that also addresses two of the more serious shortcomings.

1. The models that make up the prior information need careful consideration. In my solution to the COTOR Challenge, I developed the prior information using preliminary fits to the data and the standard errors of the parameter estimates. A true Bayesian would think hard and develop models that they believe are plausible in the absence of any data.
2. Liability claims can take a long time to settle. We are often given the task of predicting the ultimate claim severity distribution given an incomplete sample of claims. We do not know the ultimate values for many of the claims.

The fact that reinsurers go through great effort to examine the excess claims experience of their prospective contracts indicates that they believe that there are significant differences in the excess loss potential between insurers. Otherwise all reinsurance contracts would be priced using claim severity distributions based on industry aggregate experience such as those available from my employer, Insurance Services Office, Inc. (ISO).

To test this belief, I asked our (ISO's) increased limits ratemaking division to extract the empirical claim severity distributions for a liability coverage by individual insurer[6]. We the then fit mixed exponential distributions separately to 20 large insurers[7]. Each model had 10 parameters. Thus I think it is more appropriate to think of the "fitting" as "smoothing," and I do not expect each insurer's result to be necessarily predictive of future results.

The mixed exponential models were fit separately by settlement lag.

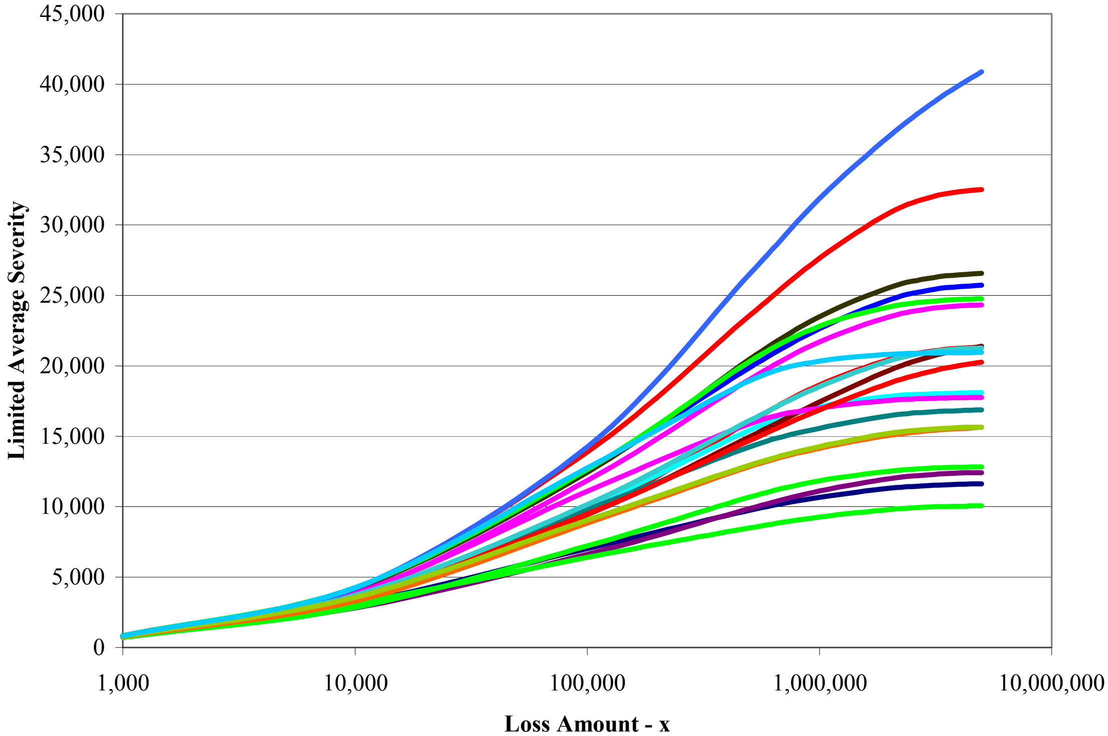
The limited average severity, $E[X^x]$, is the average severity of claims subject to a limit of $x$. Mathematically:

$$E\left[X \wedge x\right] = \int_{0}^{x} \left(1 - F(u)\right) du \,.$$

Figure 9 gives the ultimate limited average severity curves, based on the fitted mixed exponential distributions for each of the 20 insurers.

[6] ISO's standard increased limits ratemaking procedure also includes data from excess and umbrella claims that are reported separately to ISO. These claims were not included in this study.
[7] See Keatinge [1999] for information and details of fitting the mixed exponential distribution.

**Figure 9 – Initial Insurer Models**




If you are looking at Figure 9 in color, it should be apparent that the relationship between limited average severities at low loss amounts and high loss amounts is by no means perfect, but there does seem to be a general trend. The lack of correlation can be due to a lack of a fundamental relationship between losses at low levels and high levels, or it could be due to a lack of credibility of the data (as realized through the smoothing procedure.)

If there is a fundamental relationship between low-level losses and high-level losses, it makes the job of estimating high layer losses more reliable since low-level claims are more numerous. Ultimately this is a judgment call, and it is one that reinsurance actuaries routinely make.

The examples below will consist of estimates of the pure premium for the \$500,000 x \$500,000 layer, and the \$1 million x \$1 million layer. Figures 10 and 11 below respectively show how probabilities of exceeding \$5,000 and \$100,000 track with the pure premium for the \$500,000 x \$500,000 layer. The correspondence appears to be stronger in the latter case.

At this point in the analysis, I decided to use only claims that are in excess of $100,000 to estimate the cost of these layers.

Following the methodology of the previous section, the next step is to hypothesize a series of models for the data. Each model should represent the probability distribution of claims over $100,000. In developing this series of models, a good place to start is with models that were fit to individual insurer data. After all, the object is to project future losses to individual insurers.

I first attempted to use the fits directly. But in spite of the general pattern of higher layer losses increasing with lower layer losses observed in Figure 11, the Bayesian methodology would assign posterior probabilities to models where this was not the case. Given the general trend observed in Figure 11 and my prior actuarial experience (otherwise known as preconceptions) I decided to smooth out the initial set of company models. The process was informal. Loosely speaking, I dropped company models that did not behave "correctly" and replaced them with mixtures of company models that did behave "correctly." I was not able to reduce the noise entirely. Before putting this plan into practice, the choice of priors needs to be addressed more fully. I welcome debate on my notion of "correct" models. One of the advantages of the Bayesian methodology is that if forces one to make the assumptions explicit for all to see and open them to debate.

Figures 12 and 13 give the limited average severity curves and the layer pure premiums for the final set of models. These should be compared with Figures 9 and 11, respectively.

**Figure 10 – Initial Insurer Models**

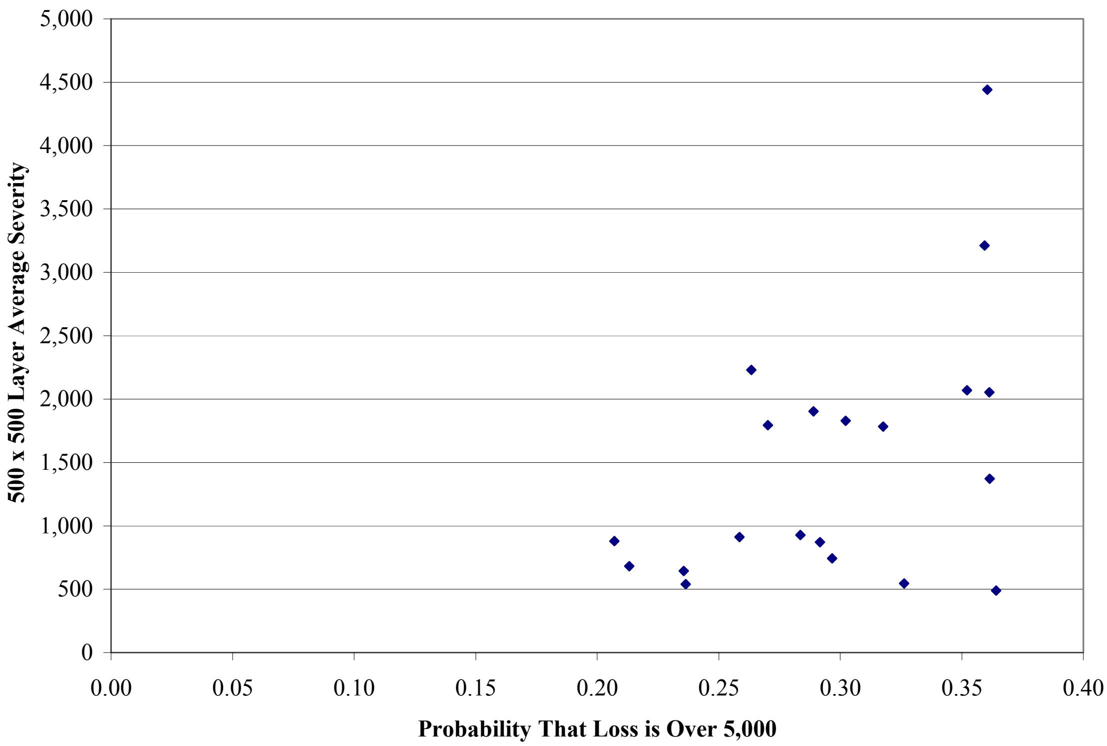


**Figure 11 – Initial Insurer Models**

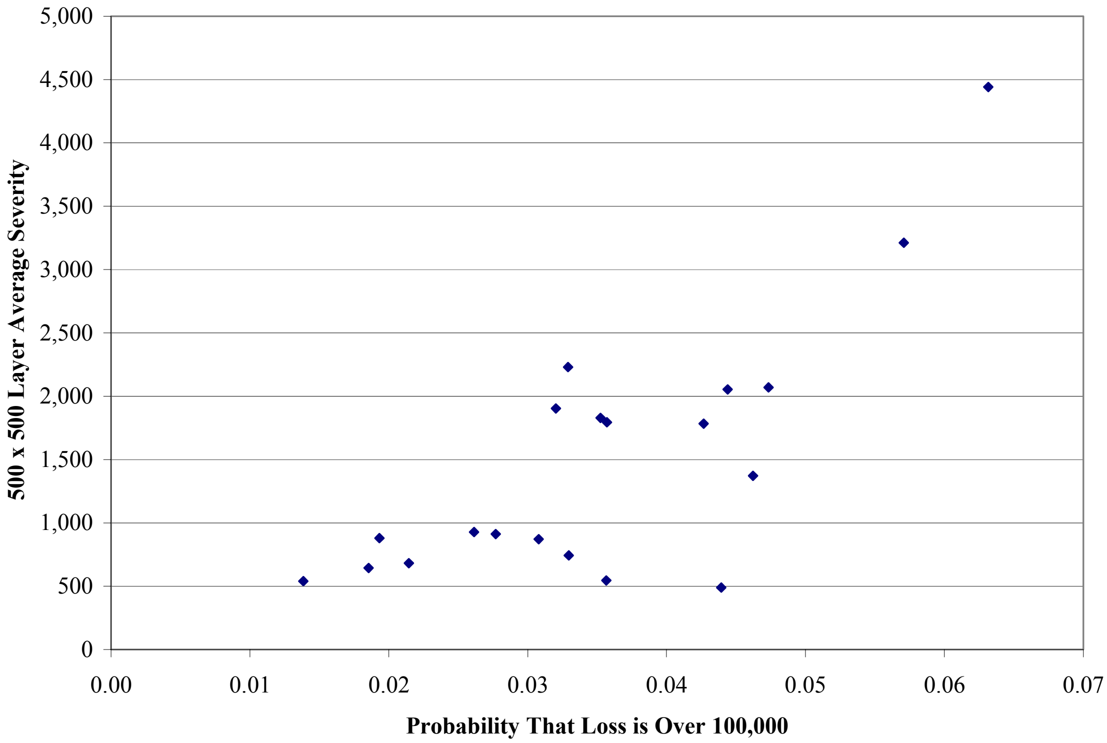


**Figure 12 – Selected Insurer Models**

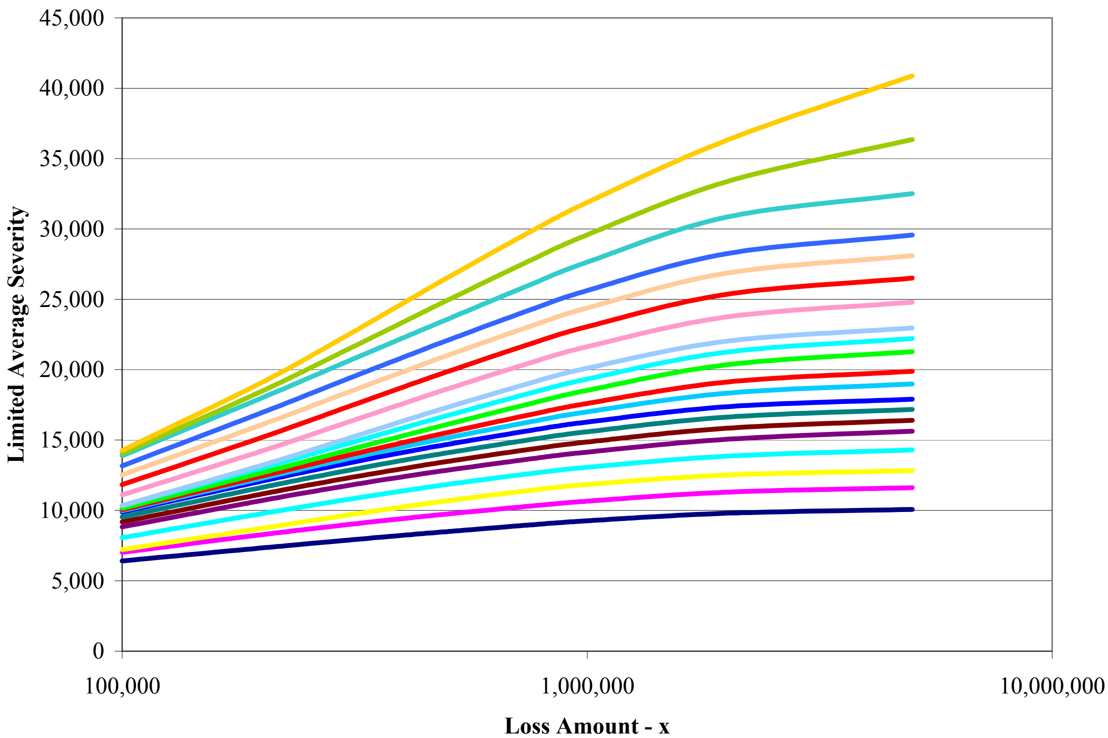
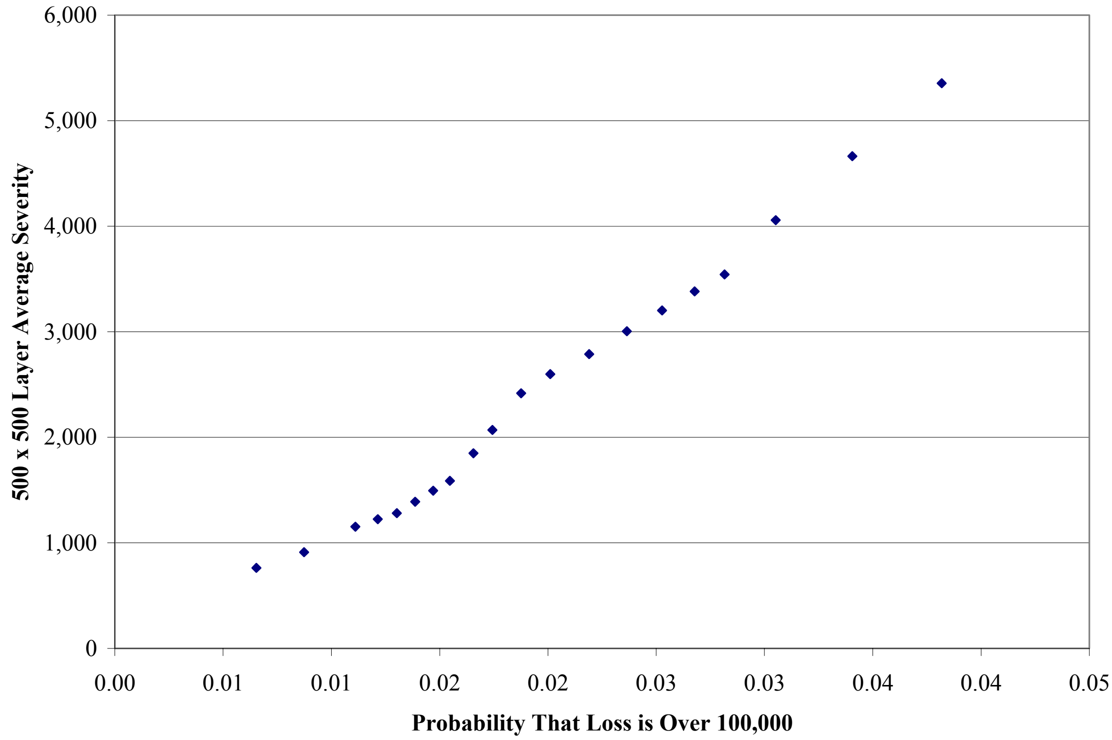


**Figure 13 – Selected Insurer Models**




The insurer models that underlie Figures 12 and 13 have been developed to ultimate. The insurer data that is presented for evaluation for an excess of loss treaty usually come from years that are too recent to contain all claims at their ultimate value. To make use of the data that reinsurers typically get, we need to have distributions of the data that are available for each insurer model.

Let's look at some examples. These examples will consist of three years of settled claim data. This data will be used to calculate the likelihood of each of 20 models. The prior probability for each model will be equal to 1/20. Then using the Bayesian methodology described in the previous section, I will calculate posterior layer pure premiums for the \$500,000 x \$500,000 layer, and for the \$1 million x \$1 million layer.

The 20 models used in this section's example will consist of the following distributions:

- The claim severity distribution for all claims settled within 1 year – Table 4.
- The claim severity distribution for all claims settled within 2 years – Table 5.

- The claim severity distribution for all claims settled within 3 years – Table 6.
- The ultimate claim severity distribution for all claims – Table 7.
- The ultimate limited average severity curve – Table 8.

As mentioned above, the models are a bit noisy, but I think they are good enough to illustrate the principles involved in this Bayesian methodology.

**Table 4**
**Cumulative Probability for Lag 1**

| Claim Amount | Prior Model Number 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 100,000 | 0.998424 | 0.997549 | 0.999228 | 0.999234 | 0.999241 | 0.999097 | 0.998949 |
| 200,000 | 0.999522 | 0.999158 | 0.999779 | 0.999784 | 0.999789 | 0.999729 | 0.999668 |
| 300,000 | 0.999741 | 0.999563 | 0.999921 | 0.999923 | 0.999925 | 0.999888 | 0.999849 |
| 400,000 | 0.999837 | 0.999737 | 0.999971 | 0.999972 | 0.999973 | 0.999947 | 0.999921 |
| 500,000 | 0.999891 | 0.999827 | 0.999989 | 0.999990 | 0.999990 | 0.999972 | 0.999953 |
| 750,000 | 0.999954 | 0.999924 | 0.999999 | 0.999999 | 0.999999 | 0.999991 | 0.999982 |
| 1,000,000 | 0.999979 | 0.999962 | 1.000000 | 1.000000 | 1.000000 | 0.999996 | 0.999992 |
| 1,500,000 | 0.999996 | 0.999988 | 1.000000 | 1.000000 | 1.000000 | 0.999999 | 0.999998 |
| 2,000,000 | 0.999999 | 0.999995 | 1.000000 | 1.000000 | 1.000000 | 1.000000 | 0.999999 |

| Claim Amount | Prior Model Number 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|
| 100,000 | 0.998806 | 0.998725 | 0.998659 | 0.998562 | 0.998122 | 0.997094 | 0.996603 |
| 200,000 | 0.999609 | 0.999610 | 0.999611 | 0.999612 | 0.999422 | 0.998979 | 0.998818 |
| 300,000 | 0.999812 | 0.999817 | 0.999822 | 0.999828 | 0.999727 | 0.999492 | 0.999413 |
| 400,000 | 0.999895 | 0.999899 | 0.999903 | 0.999908 | 0.999848 | 0.999710 | 0.999663 |
| 500,000 | 0.999935 | 0.999938 | 0.999940 | 0.999943 | 0.999905 | 0.999817 | 0.999786 |
| 750,000 | 0.999974 | 0.999974 | 0.999974 | 0.999974 | 0.999960 | 0.999928 | 0.999911 |
| 1,000,000 | 0.999987 | 0.999987 | 0.999986 | 0.999985 | 0.999980 | 0.999968 | 0.999957 |
| 1,500,000 | 0.999996 | 0.999995 | 0.999994 | 0.999993 | 0.999993 | 0.999994 | 0.999987 |
| 2,000,000 | 0.999999 | 0.999998 | 0.999997 | 0.999996 | 0.999997 | 0.999999 | 0.999995 |

| Claim Amount | Prior Model Number 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|
| 100,000 | 0.996112 | 0.995621 | 0.995130 | 0.994197 | 0.995956 | 0.997715 |
| 200,000 | 0.998658 | 0.998498 | 0.998337 | 0.997573 | 0.998259 | 0.998944 |
| 300,000 | 0.999335 | 0.999256 | 0.999177 | 0.998684 | 0.999032 | 0.999381 |
| 400,000 | 0.999616 | 0.999570 | 0.999523 | 0.999183 | 0.999392 | 0.999601 |
| 500,000 | 0.999754 | 0.999722 | 0.999690 | 0.999443 | 0.999585 | 0.999728 |
| 750,000 | 0.999893 | 0.999875 | 0.999858 | 0.999730 | 0.999807 | 0.999884 |
| 1,000,000 | 0.999945 | 0.999933 | 0.999921 | 0.999848 | 0.999898 | 0.999948 |
| 1,500,000 | 0.999981 | 0.999975 | 0.999969 | 0.999939 | 0.999964 | 0.999989 |
| 2,000,000 | 0.999992 | 0.999988 | 0.999984 | 0.999969 | 0.999983 | 0.999998 |

**Table 5**
**Cumulative Probability for Lags 1-2**

| Claim Amount | Prior Model Number 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 100,000 | 0.996249 | 0.994479 | 0.996598 | 0.995650 | 0.994980 | 0.994376 | 0.993730 |
| 200,000 | 0.998776 | 0.997967 | 0.998770 | 0.998387 | 0.998117 | 0.997858 | 0.997582 |
| 300,000 | 0.999329 | 0.998904 | 0.999393 | 0.999171 | 0.999012 | 0.998860 | 0.998698 |
| 400,000 | 0.999579 | 0.999315 | 0.999648 | 0.999499 | 0.999391 | 0.999291 | 0.999184 |
| 500,000 | 0.999717 | 0.999529 | 0.999768 | 0.999659 | 0.999579 | 0.999508 | 0.999432 |
| 750,000 | 0.999881 | 0.999770 | 0.999886 | 0.999824 | 0.999779 | 0.999744 | 0.999706 |
| 1,000,000 | 0.999947 | 0.999869 | 0.999931 | 0.999891 | 0.999862 | 0.999842 | 0.999821 |
| 1,500,000 | 0.999989 | 0.999947 | 0.999968 | 0.999948 | 0.999933 | 0.999926 | 0.999918 |
| 2,000,000 | 0.999998 | 0.999973 | 0.999983 | 0.999971 | 0.999962 | 0.999959 | 0.999955 |

| Claim Amount | Prior Model Number 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|
| 100,000 | 0.993080 | 0.993542 | 0.993901 | 0.994418 | 0.993722 | 0.992133 | 0.990703 |
| 200,000 | 0.997303 | 0.997541 | 0.997726 | 0.997995 | 0.997530 | 0.996472 | 0.995866 |
| 300,000 | 0.998534 | 0.998696 | 0.998821 | 0.999003 | 0.998640 | 0.997820 | 0.997472 |
| 400,000 | 0.999075 | 0.999190 | 0.999279 | 0.999407 | 0.999126 | 0.998490 | 0.998252 |
| 500,000 | 0.999355 | 0.999437 | 0.999501 | 0.999593 | 0.999375 | 0.998882 | 0.998699 |
| 750,000 | 0.999667 | 0.999704 | 0.999733 | 0.999775 | 0.999656 | 0.999388 | 0.999268 |
| 1,000,000 | 0.999800 | 0.999817 | 0.999830 | 0.999849 | 0.999780 | 0.999625 | 0.999538 |
| 1,500,000 | 0.999910 | 0.999913 | 0.999916 | 0.999920 | 0.999891 | 0.999827 | 0.999778 |
| 2,000,000 | 0.999952 | 0.999952 | 0.999953 | 0.999954 | 0.999939 | 0.999907 | 0.999877 |

| Claim Amount | Prior Model Number 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|
| 100,000 | 0.989248 | 0.987769 | 0.986267 | 0.983454 | 0.987292 | 0.991076 |
| 200,000 | 0.995249 | 0.994620 | 0.993980 | 0.992118 | 0.993669 | 0.995200 |
| 300,000 | 0.997118 | 0.996756 | 0.996388 | 0.995124 | 0.995968 | 0.996801 |
| 400,000 | 0.998009 | 0.997762 | 0.997510 | 0.996571 | 0.997128 | 0.997676 |
| 500,000 | 0.998513 | 0.998323 | 0.998129 | 0.997388 | 0.997808 | 0.998222 |
| 750,000 | 0.999146 | 0.999022 | 0.998895 | 0.998429 | 0.998698 | 0.998964 |
| 1,000,000 | 0.999449 | 0.999359 | 0.999266 | 0.998945 | 0.999134 | 0.999320 |
| 1,500,000 | 0.999727 | 0.999675 | 0.999622 | 0.999451 | 0.999539 | 0.999626 |
| 2,000,000 | 0.999847 | 0.999816 | 0.999784 | 0.999685 | 0.999717 | 0.999749 |

**Table 6**
**Cumulative Probability for Lags 1-3**

| Claim Amount | Prior Model Number 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 100,000 | 0.993117 | 0.991183 | 0.991960 | 0.990009 | 0.988313 | 0.987309 | 0.986241 |
| 200,000 | 0.997248 | 0.996461 | 0.996618 | 0.995966 | 0.995416 | 0.994950 | 0.994452 |
| 300,000 | 0.998325 | 0.997996 | 0.998083 | 0.997766 | 0.997507 | 0.997228 | 0.996931 |
| 400,000 | 0.998843 | 0.998687 | 0.998752 | 0.998567 | 0.998419 | 0.998233 | 0.998034 |
| 500,000 | 0.999143 | 0.999056 | 0.999110 | 0.998984 | 0.998886 | 0.998753 | 0.998611 |
| 750,000 | 0.999529 | 0.999487 | 0.999527 | 0.999458 | 0.999404 | 0.999337 | 0.999265 |
| 1,000,000 | 0.999710 | 0.999680 | 0.999711 | 0.999664 | 0.999627 | 0.999589 | 0.999549 |
| 1,500,000 | 0.999865 | 0.999848 | 0.999867 | 0.999841 | 0.999819 | 0.999805 | 0.999790 |
| 2,000,000 | 0.999927 | 0.999917 | 0.999928 | 0.999911 | 0.999898 | 0.999891 | 0.999885 |

| Claim Amount | Prior Model Number 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|
| 100,000 | 0.985174 | 0.986141 | 0.986900 | 0.988014 | 0.987115 | 0.985052 | 0.982547 |
| 200,000 | 0.993953 | 0.994405 | 0.994760 | 0.995282 | 0.994564 | 0.992851 | 0.991672 |
| 300,000 | 0.996631 | 0.996928 | 0.997161 | 0.997502 | 0.996895 | 0.995426 | 0.994721 |
| 400,000 | 0.997834 | 0.998038 | 0.998197 | 0.998430 | 0.997942 | 0.996754 | 0.996259 |
| 500,000 | 0.998468 | 0.998608 | 0.998718 | 0.998879 | 0.998494 | 0.997557 | 0.997171 |
| 750,000 | 0.999193 | 0.999247 | 0.999289 | 0.999351 | 0.999139 | 0.998625 | 0.998370 |
| 1,000,000 | 0.999509 | 0.999526 | 0.999540 | 0.999559 | 0.999437 | 0.999140 | 0.998956 |
| 1,500,000 | 0.999774 | 0.999771 | 0.999769 | 0.999765 | 0.999715 | 0.999594 | 0.999488 |
| 2,000,000 | 0.999878 | 0.999873 | 0.999870 | 0.999864 | 0.999839 | 0.999777 | 0.999715 |

| Claim Amount | Prior Model Number 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|
| 100,000 | 0.980077 | 0.977642 | 0.975241 | 0.970705 | 0.974003 | 0.977352 |
| 200,000 | 0.990513 | 0.989375 | 0.988255 | 0.985651 | 0.986512 | 0.987325 |
| 300,000 | 0.994028 | 0.993349 | 0.992682 | 0.990980 | 0.991195 | 0.991346 |
| 400,000 | 0.995774 | 0.995298 | 0.994831 | 0.993577 | 0.993620 | 0.993606 |
| 500,000 | 0.996793 | 0.996421 | 0.996058 | 0.995061 | 0.995077 | 0.995047 |
| 750,000 | 0.998121 | 0.997876 | 0.997635 | 0.996980 | 0.997026 | 0.997049 |
| 1,000,000 | 0.998774 | 0.998596 | 0.998421 | 0.997952 | 0.997998 | 0.998030 |
| 1,500,000 | 0.999385 | 0.999283 | 0.999182 | 0.998920 | 0.998912 | 0.998892 |
| 2,000,000 | 0.999653 | 0.999592 | 0.999533 | 0.999378 | 0.999321 | 0.999251 |

**Table 7**
**Ultimate Cumulative Probability**

| Claim Amount | Prior Model Number 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 100,000 | 0.986144 | 0.981451 | 0.978563 | 0.975297 | 0.972292 | 0.970836 | 0.969375 |
| 200,000 | 0.993462 | 0.991264 | 0.988893 | 0.987858 | 0.986981 | 0.986135 | 0.985300 |
| 300,000 | 0.995749 | 0.994592 | 0.992756 | 0.992349 | 0.992056 | 0.991480 | 0.990917 |
| 400,000 | 0.996907 | 0.996197 | 0.994830 | 0.994621 | 0.994495 | 0.994066 | 0.993649 |
| 500,000 | 0.997603 | 0.997108 | 0.996110 | 0.995958 | 0.995865 | 0.995527 | 0.995200 |
| 750,000 | 0.998547 | 0.998270 | 0.997838 | 0.997697 | 0.997579 | 0.997370 | 0.997172 |
| 1,000,000 | 0.999030 | 0.998845 | 0.998670 | 0.998530 | 0.998397 | 0.998261 | 0.998132 |
| 1,500,000 | 0.999499 | 0.999403 | 0.999387 | 0.999273 | 0.999155 | 0.999095 | 0.999039 |
| 2,000,000 | 0.999714 | 0.999659 | 0.999668 | 0.999582 | 0.999490 | 0.999467 | 0.999445 |

| Claim Amount | Prior Model Number 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|
| 100,000 | 0.967995 | 0.966848 | 0.965978 | 0.964755 | 0.962971 | 0.961104 | 0.957388 |
| 200,000 | 0.984525 | 0.983441 | 0.982566 | 0.981238 | 0.979893 | 0.978100 | 0.976354 |
| 300,000 | 0.990400 | 0.989463 | 0.988683 | 0.987462 | 0.986407 | 0.984925 | 0.983884 |
| 400,000 | 0.993269 | 0.992457 | 0.991774 | 0.990694 | 0.989877 | 0.988749 | 0.988008 |
| 500,000 | 0.994904 | 0.994197 | 0.993601 | 0.992655 | 0.992024 | 0.991202 | 0.990608 |
| 750,000 | 0.996993 | 0.996480 | 0.996049 | 0.995368 | 0.995031 | 0.994695 | 0.994273 |
| 1,000,000 | 0.998018 | 0.997636 | 0.997316 | 0.996813 | 0.996621 | 0.996506 | 0.996183 |
| 1,500,000 | 0.998990 | 0.998768 | 0.998583 | 0.998294 | 0.998218 | 0.998230 | 0.998036 |
| 2,000,000 | 0.999427 | 0.999294 | 0.999184 | 0.999012 | 0.998974 | 0.998998 | 0.998881 |

| Claim Amount | Prior Model Number 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|
| 100,000 | 0.953900 | 0.950641 | 0.947611 | 0.942936 | 0.940188 | 0.936834 |
| 200,000 | 0.974729 | 0.973226 | 0.971844 | 0.969483 | 0.965947 | 0.961818 |
| 300,000 | 0.982924 | 0.982042 | 0.981240 | 0.979667 | 0.976492 | 0.972815 |
| 400,000 | 0.987327 | 0.986706 | 0.986145 | 0.984848 | 0.982275 | 0.979308 |
| 500,000 | 0.990063 | 0.989568 | 0.989122 | 0.987937 | 0.985918 | 0.983600 |
| 750,000 | 0.993887 | 0.993537 | 0.993223 | 0.992192 | 0.991065 | 0.989781 |
| 1,000,000 | 0.995888 | 0.995620 | 0.995380 | 0.994510 | 0.993768 | 0.992925 |
| 1,500,000 | 0.997859 | 0.997698 | 0.997554 | 0.996986 | 0.996440 | 0.995814 |
| 2,000,000 | 0.998774 | 0.998677 | 0.998589 | 0.998237 | 0.997707 | 0.997093 |

**Table 8**
**Ultimate Limited Average Severity**

| Claim Amount | Prior Model Number 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 100,000 | 6,412 | 7,021 | 7,217 | 8,067 | 8,835 | 9,181 | 9,540 |
| 200,000 | 7,340 | 8,276 | 8,735 | 9,774 | 10,712 | 11,168 | 11,637 |
| 300,000 | 7,864 | 8,961 | 9,629 | 10,735 | 11,725 | 12,251 | 12,788 |
| 400,000 | 8,226 | 9,413 | 10,241 | 11,375 | 12,385 | 12,960 | 13,546 |
| 500,000 | 8,498 | 9,744 | 10,689 | 11,841 | 12,861 | 13,474 | 14,097 |
| 750,000 | 8,964 | 10,302 | 11,417 | 12,604 | 13,650 | 14,329 | 15,016 |
| 1,000,000 | 9,261 | 10,655 | 11,842 | 13,065 | 14,142 | 14,864 | 15,591 |
| 1,500,000 | 9,612 | 11,073 | 12,297 | 13,583 | 14,725 | 15,493 | 16,264 |
| 2,000,000 | 9,802 | 11,300 | 12,524 | 13,860 | 15,054 | 15,842 | 16,631 |

| Claim Amount | Prior Model Number 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|
| 100,000 | 9,891 | 9,977 | 10,045 | 10,148 | 10,337 | 10,344 | 11,107 |
| 200,000 | 12,091 | 12,290 | 12,448 | 12,682 | 13,025 | 13,217 | 14,235 |
| 300,000 | 13,306 | 13,606 | 13,847 | 14,209 | 14,671 | 15,028 | 16,180 |
| 400,000 | 14,107 | 14,495 | 14,809 | 15,286 | 15,841 | 16,328 | 17,567 |
| 500,000 | 14,692 | 15,155 | 15,533 | 16,111 | 16,739 | 17,323 | 18,628 |
| 750,000 | 15,668 | 16,283 | 16,787 | 17,567 | 18,311 | 19,032 | 20,461 |
| 1,000,000 | 16,279 | 17,004 | 17,602 | 18,528 | 19,336 | 20,110 | 21,632 |
| 1,500,000 | 16,990 | 17,863 | 18,583 | 19,702 | 20,571 | 21,361 | 23,009 |
| 2,000,000 | 17,374 | 18,333 | 19,125 | 20,356 | 21,253 | 22,032 | 23,756 |

| Claim Amount | Prior Model Number 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|
| 100,000 | 11,832 | 12,518 | 13,166 | 13,884 | 14,036 | 14,257 |
| 200,000 | 15,197 | 16,105 | 16,958 | 18,006 | 18,489 | 19,102 |
| 300,000 | 17,267 | 18,289 | 19,246 | 20,485 | 21,307 | 22,315 |
| 400,000 | 18,735 | 19,830 | 20,854 | 22,234 | 23,344 | 24,684 |
| 500,000 | 19,856 | 21,006 | 22,080 | 23,583 | 24,922 | 26,526 |
| 750,000 | 21,803 | 23,057 | 24,223 | 26,002 | 27,721 | 29,758 |
| 1,000,000 | 23,058 | 24,388 | 25,623 | 27,641 | 29,585 | 31,880 |
| 1,500,000 | 24,550 | 25,985 | 27,314 | 29,688 | 31,939 | 34,581 |
| 2,000,000 | 25,367 | 26,865 | 28,251 | 30,851 | 33,369 | 36,320 |

We are now ready to work through our examples in detail. Exhibits 1-3 below give the results of the Bayesian methodology for a small, a medium, and a large insurer. The exhibits take claim severity data from each insurer and provide estimates of the layer pure premium for a $500,000 x $500,000 layer and for a $1 million x $1 million layer. For the record, I note that the "data" for each insurer was produced from a simulation taken from a single claim severity distribution. The "true" expected pure premiums for the layers are $1,382 and $1,015, respectively.

Here is a step by step description of the calculations in those exhibits.

**Lags** – As mentioned above, the claim severity distributions underlying the models were fit by settlement lag. Claims from the most recent accident year consist of claims that were settled in Lag 1. Claims from the second most recent year consist of claims that were settled in Lags 1 and 2. The designation Lag 1 corresponds to accident year $AY$=1, Lags 1-2 corresponds to $AY$=2 and Lags 1-3 corresponds to $AY$=3.

**Interval Lower Bound and Claim Count** – We summarized the claim amounts in intervals, with the lower bound of the interval being specified to the left of the claim count. Let $n_{i,AY}$ be the observed claim count in the $i^{th}$ interval for accident year $AY$. For example, in Exhibit 1, there were 15 claims in the interval (100,000, 200,000] and there were no claims more than 2,000,000 in Lag 1. The underlying exposure was the same for each accident year. Note that there are more high severity claims in the earlier accident years.

**Prior Model #** – These are the models described in Tables 4-8 above. Each table gives a different part of each model as described above.

**Posterior Probability** – This is calculated for each prior model. Let:

- $n_{i,AY}$ be the number of claims in the $i^{th}$ interval of the $AY^{th}$ accident year.
- $x_{i,\mathrm{AY}}$ be the lower bound of the ith interval, $x_{i+1,\mathrm{AY}}$ be the upper bound of the ith interval. Note $x_{10,AY}$ = □.
- Let $P_{i,\mathrm{AY},m}$ be the probability that a claim is observed in $i^{th}$ cell given that it is in the $AY^{th}$ accident year for model $m$. Let $x_i$ be the lower bound of the $i^{th}$ interval. Let $F_{AY,m}(x_i)$ be the probability that a claim is $\leq x_i$ in accident year $AY$ for model $m$. These probabilities are given in Tables 4-6 above. Then:

$$P_{i,AY,m} = \frac{F_{AY,m}(x_{i+1}) - F_{AY,m}(x_i)}{1 - F_{AY,m}(x_1)}.$$

- The likelihood of the data $\{n_{i,AY}\}$ for model $m$, $l_m$, is given by:

$$l_m = \prod_{i=1}^{9} \prod_{AY=1}^{3} \left(P_{i,AY,m}\right)^{n_{i,AY}}.$$

- Let Prior($m$) be the prior probability associated with model $m$. (In this example, Prior($m$) = 1/20 for all $m$.) Then according to Bayes' theorem:

$$\text{Posterior}(m) \propto l_m \cdot \text{Prior}(m).$$

  As was done in the COTOR Challenge example, you first calculate the product $l_m$·Prior($m$) and then normalize.

**Layer Pure Premium** – The layer pure premium for each model is calculated from the limited average severity curves in Table 8. For example, the layer pure premium for the \$500 thousand x \$500 thousand layer is calculated as the difference between the limited average severity for \$1 million and the limited average severity at \$500 thousand[8].

**Posterior Mean and Standard Deviation –** These quantities are calculated by the following formulas.

$$\text{Posterior Mean} = \sum_{m=1}^{20} \text{Layer Pure Premium}(m) \cdot \text{Posterior}(m).$$

$$\text{Posterior Standard Deviation} = \sqrt{\sum_{m=1}^{20} \text{Layer Pure Premium}(m)^2 \cdot \text{Posterior}(m) - \text{Posterior Mean}^2}.$$

Note that as we increase the exposure, and hence the number of observations, the posterior probability tends to be concentrated on fewer models. As the posterior standard deviations indicate, increasing exposure leads to less uncertainty in the final estimate.

[8] It is often the case that the reinsurer will have an independent estimate of the probability that a claim is more than \$100,000. To make use of this information, the reinsurer should multiply the layer pure premium times the ratio of this probability to each model's probability that a claim is more than \$100,000. I did not do this in these examples.

## 5. CONCLUDING REMARKS

In this paper, I gave some examples showing how to use the likelihood function and Bayes' theorem to estimate the costs of high layers of reinsurance. Many of the assumptions need to be debated, and regardless of how the debate is resolved, much work is needed to complete the job. I hope this paper provides strong evidence that such an approach can succeed and provide a sound methodology for reinsures to use in pricing coverage of high layers of reinsurance.

**Exhibit 1 – Small Insurer**

| Lags | Interval Lower Bound | Claim Count |
|---|---|---|
| 1 | 100,000 | 15 |
| 1 | 200,000 | 2 |
| 1 | 300,000 | 1 |
| 1 | 400,000 | 2 |
| 1 | 500,000 | 0 |
| 1 | 750,000 | 0 |
| 1 | 1,000,000 | 0 |
| 1 | 1,500,000 | 0 |
| 1 | 2,000,000 | 0 |
| 1-2 | 100,000 | 40 |
| 1-2 | 200,000 | 10 |
| 1-2 | 300,000 | 1 |
| 1-2 | 400,000 | 0 |
| 1-2 | 500,000 | 2 |
| 1-2 | 750,000 | 0 |
| 1-2 | 1,000,000 | 2 |
| 1-2 | 1,500,000 | 0 |
| 1-2 | 2,000,000 | 0 |
| 1-3 | 100,000 | 76 |
| 1-3 | 200,000 | 26 |
| 1-3 | 300,000 | 11 |
| 1-3 | 400,000 | 3 |
| 1-3 | 500,000 | 8 |
| 1-3 | 750,000 | 0 |
| 1-3 | 1,000,000 | 0 |
| 1-3 | 1,500,000 | 0 |
| 1-3 | 2,000,000 | 0 |

| | | Layer Pure Premium | |
|---|---|---|---|
| Prior Model # | Posterior Probability | $500K x $500K | $1M x $1M |
| 1 | 0.016406 | 763 | 541 |
| 2 | 0.041658 | 911 | 645 |
| 3 | 0.089063 | 1,153 | 682 |
| 4 | 0.130281 | 1,224 | 796 |
| 5 | 0.157593 | 1,281 | 912 |
| 6 | 0.110614 | 1,390 | 978 |
| 7 | 0.075702 | 1,494 | 1,040 |
| 8 | 0.053226 | 1,587 | 1,095 |
| 9 | 0.080525 | 1,849 | 1,328 |
| 10 | 0.104056 | 2,069 | 1,523 |
| 11 | 0.129925 | 2,417 | 1,828 |
| 12 | 0.010896 | 2,598 | 1,916 |
| 13 | 0.000007 | 2,788 | 1,922 |
| 14 | 0.000009 | 3,004 | 2,124 |
| 15 | 0.000011 | 3,202 | 2,309 |
| 16 | 0.000013 | 3,382 | 2,477 |
| 17 | 0.000014 | 3,543 | 2,628 |
| 18 | 0.000000 | 4,058 | 3,211 |
| 19 | 0.000000 | 4,663 | 3,784 |
| 20 | 0.000000 | 5,354 | 4,440 |
| **Posterior Mean** | | **1,572** | **1,113** |
| **Posterior Std. Dev.** | | **463** | **385** |

**Exhibit 2 – Medium Insurer**

| Lags | Interval Lower Bound | Claim Count |
|---|---|---|
| 1 | 100,000 | 31 |
| 1 | 200,000 | 12 |
| 1 | 300,000 | 4 |
| 1 | 400,000 | 2 |
| 1 | 500,000 | 1 |
| 1 | 750,000 | 0 |
| 1 | 1,000,000 | 0 |
| 1 | 1,500,000 | 0 |
| 1 | 2,000,000 | 0 |
| 1-2 | 100,000 | 107 |
| 1-2 | 200,000 | 33 |
| 1-2 | 300,000 | 14 |
| 1-2 | 400,000 | 3 |
| 1-2 | 500,000 | 7 |
| 1-2 | 750,000 | 2 |
| 1-2 | 1,000,000 | 0 |
| 1-2 | 1,500,000 | 0 |
| 1-2 | 2,000,000 | 0 |
| 1-3 | 100,000 | 191 |
| 1-3 | 200,000 | 47 |
| 1-3 | 300,000 | 31 |
| 1-3 | 400,000 | 22 |
| 1-3 | 500,000 | 6 |
| 1-3 | 750,000 | 5 |
| 1-3 | 1,000,000 | 1 |
| 1-3 | 1,500,000 | 2 |
| 1-3 | 2,000,000 | 1 |

| Prior Model # | Posterior Probability | Layer Pure Premium $500K x $500K | Layer Pure Premium $1M x $1M |
|---|---|---|---|
| 1 | 0.000973 | 763 | 541 |
| 2 | 0.021135 | 911 | 645 |
| 3 | 0.221357 | 1,153 | 682 |
| 4 | 0.235280 | 1,224 | 796 |
| 5 | 0.209597 | 1,281 | 912 |
| 6 | 0.123874 | 1,390 | 978 |
| 7 | 0.059523 | 1,494 | 1,040 |
| 8 | 0.028986 | 1,587 | 1,095 |
| 9 | 0.037532 | 1,849 | 1,328 |
| 10 | 0.037637 | 2,069 | 1,523 |
| 11 | 0.023539 | 2,417 | 1,828 |
| 12 | 0.000567 | 2,598 | 1,916 |
| 13 | 0.000000 | 2,788 | 1,922 |
| 14 | 0.000000 | 3,004 | 2,124 |
| 15 | 0.000000 | 3,202 | 2,309 |
| 16 | 0.000000 | 3,382 | 2,477 |
| 17 | 0.000000 | 3,543 | 2,628 |
| 18 | 0.000000 | 4,058 | 3,211 |
| 19 | 0.000000 | 4,663 | 3,784 |
| 20 | 0.000000 | 5,354 | 4,440 |
| **Posterior Mean** | | **1,344** | **909** |
| **Posterior Std. Dev.** | | **278** | **245** |

**Exhibit 3 - Large Insurer**

| Lags | Interval Lower Bound | Claim Count | Prior Model # | Posterior Probability | Layer Pure Premium $500K x $500K | Layer Pure Premium $1M x $1M |
|---|---|---|---|---|---|---|
| 1 | 100,000 | 77 | 1 | 0.000000 | 763 | 541 |
| 1 | 200,000 | 20 | 2 | 0.000193 | 911 | 645 |
| 1 | 300,000 | 7 | 3 | 0.000481 | 1,153 | 682 |
| 1 | 400,000 | 1 | 4 | 0.050204 | 1,224 | 796 |
| 1 | 500,000 | 1 | 5 | 0.689060 | 1,281 | 912 |
| 1 | 750,000 | 1 | 6 | 0.179377 | 1,390 | 978 |
| 1 | 1,000,000 | 0 | 7 | 0.015896 | 1,494 | 1,040 |
| 1 | 1,500,000 | 0 | 8 | 0.001625 | 1,587 | 1,095 |
| 1 | 2,000,000 | 0 | 9 | 0.009032 | 1,849 | 1,328 |
| | | | 10 | 0.022443 | 2,069 | 1,523 |
| | | | 11 | 0.031690 | 2,417 | 1,828 |
| 1-2 | 100,000 | 193 | 12 | 0.000000 | 2,598 | 1,916 |
| 1-2 | 200,000 | 60 | 13 | 0.000000 | 2,788 | 1,922 |
| 1-2 | 300,000 | 22 | 14 | 0.000000 | 3,004 | 2,124 |
| 1-2 | 400,000 | 14 | 15 | 0.000000 | 3,202 | 2,309 |
| 1-2 | 500,000 | 10 | 16 | 0.000000 | 3,382 | 2,477 |
| 1-2 | 750,000 | 7 | 17 | 0.000000 | 3,543 | 2,628 |
| 1-2 | 1,000,000 | 2 | 18 | 0.000000 | 4,058 | 3,211 |
| 1-2 | 1,500,000 | 1 | 19 | 0.000000 | 4,663 | 3,784 |
| 1-2 | 2,000,000 | 1 | 20 | 0.000000 | 5,354 | 4,440 |
| | | | | | | |
| | | | **Posterior Mean** | | **1,360** | **966** |
| 1-3 | 100,000 | 431 | **Posterior Std. Dev.** | | **234** | **188** |
| 1-3 | 200,000 | 117 | | | | |
| 1-3 | 300,000 | 40 | | | | |
| 1-3 | 400,000 | 25 | | | | |
| 1-3 | 500,000 | 24 | | | | |
| 1-3 | 750,000 | 4 | | | | |
| 1-3 | 1,000,000 | 5 | | | | |
| 1-3 | 1,500,000 | 0 | | | | |
| 1-3 | 2,000,000 | 1 | | | | |

## 6. REFERENCES


1. Keatinge, Clive L., "Modeling Losses with the Mixed Exponential Distribution", *Proceedings of the Casualty Actuarial Society*, 1999. http://www.casact.org/pubs/proceed/proceed99/99578.pdf

2. Klugman, S.A., Panjer, H.H. and Willmot, G.E., *Loss Models: From Data to Decisions*, John Wiley and Sons, 1998.

3. Klugman, S.A., Panjer, H.H. and Willmot, G.E., *Loss Models: From Data to Decisions*, Second Edition, John Wiley and Sons, 2004.

4. Klugman, Stuart A., "Discussion of Meyers – Quantifying the Uncertainty in Claim Severity Estimates when Using the Single Parameter Pareto", *Proceedings of the Casualty Actuarial Society*, 1994. http://www.casact.org/pubs/proceed/proceed94/94114.pdf

5. Kreps, Rodney E. "Parameter Uncertainty in (Log)normal Distributions", *Proceedings of the Casualty Actuarial Society*, 1997. http://www.casact.org/pubs/proceed/proceed97/97553.pdf

6. Meyers, Glenn G., "Quantifying the Uncertainty in Claim Severity Estimates when Using the Single Parameter Pareto", *Proceedings of the Casualty Actuarial Society*, 1994. http://www.casact.org/pubs/proceed/proceed94/94091.pdf

## Biography of the Author

**Glenn Meyers** is the Chief Actuary for ISO Innovative Analytics. He holds a bachelor's degree in mathematics and physics from Alma College in Alma, Mich., a master's degree in mathematics from Oakland University, and a Ph.D. in mathematics from the State University of New York at Albany. Glenn is a Fellow of the Casualty Actuarial Society and a member of the American Academy of Actuaries. Before joining ISO in 1988, Glenn worked at CNA Insurance Companies and the University of Iowa.

Glenn has assumed positions of increasing responsibility at ISO, including leadership of increased limits and catastrophe ratemaking. He created ISO's Multi-Distributional Increased Limits Developer (MILD), ISO's increased limits software, and Property Size-of-Loss Database (PSOLD), ISO's model for commercial property size of loss distributions.

Glenn was the principal author of several books in the *ISO Issues Series*. His current projects include developing ISO's Dynamic Financial Analysis and insurance scoring products solutions.

Glenn's work has been published in *Proceedings of the Casualty Actuarial Society (CAS)*. He is a three-time winner of the Woodward-Fondiller Prize, a two-time winner of the Dorweiller Prize and a winner of the Dynamic Financial Analysis Prize. He is a frequent speaker at CAS meetings and seminars.

His service to the CAS includes long-term membership of the Examination Committee and the Committee on the Theory of Risk. He serves on the International Actuarial Association Solvency Committee and co-chairs the CAS Working Party on Correlation. He also serves on the CAS Dynamic Risk Modeling Committee and the Modeling Workshop Task Force.