MLE for Claims with Several Retentions

Gary G Venter, Guy Carpenter Instrat

In the *Loss Models* readings, CAS students learn how to fit severity distributions by MLE, including the case of fitting a ground-up distribution where only losses above a deductible are available. In that case the MLE looks for the ground-up distribution parameters that provide the best fit to the known excess losses. This procedure falls apart, however, when different deductibles are used and there are different degrees of exposure to each. This note derives the likelihood function for that situation.

Acknowledgment

This problem was posed by Claude Yoder, who recognized the shortcomings of the usual method. The likelihood function was worked out one day after lunch by Rodney Kreps, Paul Silberbush, John Pagliaccio and the author. The usual caveat applies.

MLE for Claims with Several Retentions

A not atypical fitting problem for reinsurance losses is trying to find the severity distribution that is generating claims, where the data is provided for groups of excess policies, each with its own retention and limit. The case considered here is where information is also presented about how much exposure is included in each group. Losses are of course truncated from below at their retentions and censored from above at their limits.

What is the likelihood function for severity in this situation? It turns out that that question is intrinsically linked with the likelihood function for frequency, as the exposure information comes in through frequency. First some notation. To simplify the typing and also to increase the visibility of the sub-variables, subscripts will not be used.

So suppose you have k groups of claims, and the jth group has retention Rj, upper limit or plafond (i.e., retention plus policy limit) Uj, and Ej exposures. The data for the group consists of Mj claims at the policy limit, (some of which are probably censored by the limit, so would have been larger without the limit) and Nj claims less than the limit. All the ground-up claim sizes are assumed to come from the same distribution, with severity distribution function F and density f. The exposures for all the groups are assumed to have the same ground-up Poisson loss frequency h per unit of exposure, so the expected excess frequency for the jth group is: $hEj(1-F(Rj)) = hEjS(Rj)$, which will be denoted by hj, and is still Poisson. Estimating h is part of the problem to be addressed.

With this setup, what is the likelihood function for the set of losses observed in the jth group? This is the product of the frequency and severity probabilities of observing that many claims of those sizes. Let **a** denote the severity parameters, considered to be a vector, and Xji the ground-up amount of the ith loss in the jth group. Then the likelihood function at h and **a** for the jth group is:

$Lj(h,\mathbf{a}) = hj^{Nj+Mj}\exp(-hj)[\prod_{i=1}^{Nj} f(Xji|\mathbf{a})]S(Uj|\mathbf{a})^{Mj}S(Rj)^{-Mj-Nj}$.

The log-likelihood for all the groups combined is the log of the product of these:

$LL(h,\mathbf{a}) = \Sigma_{j=1}^{k}\{\ln(hj)(Nj+Mj) - hj + \Sigma_{i=1}^{Nj} \ln[f(Xji|\mathbf{a})] + Mj\ln[S(Uj|\mathbf{a})] - (Mj+Nj)\ln[S(Rj)]\}$

Since hj is a function of F, this cannot be separated into frequency and severity sections. But note that ln(hj) = ln h + ln Ej + ln S(Rj). Putting this into the first term cancels out the last term. The ln Ej term does not contain any parameters, so can be eliminated. Thus LL can be written:

(1) $LL(h,\mathbf{a}) = \Sigma_{j=1}^{k}\{\ln(h)(Nj+Mj) - hj + \Sigma_{i=1}^{Nj} \ln[f(Xji|\mathbf{a})] + Mj\ln[S(Uj|\mathbf{a})]\}$

Formula (1) is the answer, in the sense that this is the function that has to be maximized to estimate h and **a**. Some insight can be gained by considering its partials, however. First, wrt h:

$\partial LL/\partial h = \Sigma_{j=1}^{k}\{(Nj+Mj)/h - EjS(Rj|\mathbf{a})\}$, which setting to zero gives:

(2): $h = \Sigma_{j=1}^{k} (Nj+Mj) / \Sigma_{j=1}^{k} EjS(Rj|\mathbf{a})$

This gives the ground up frequency in terms of the severity parameters.

The partial of the LL wrt **a** can be seen to have two components – the partial of the first two terms is a frequency component, and the partial of the second two is the usual severity component that does not consider exposures. If these separately become zero at the maximum, then the exposure information is not affecting the parameters. But it can be shown that for this to happen, a different value of h would result. Thus the exposure information does make a difference.

A large number of exposures with high retentions would be expected to produce several large ground-up claims, but no small ones. But if retentions are small, the same sample would suggest that the severity distribution tends to produce larger claims. Thus including the exposure information should make a difference of this kind. Some practical testing of MLE with this LL could discern if this is the case.

Example

To illustrate the application of this approach, a simplified example is provided. Assume that there are two groups of policies, with retentions of 1M and 10M. Each has experienced a single claim, with ground-up amounts of 2M and 20M. It is assumed that ground-up claim frequency is Poisson and severity ballasted Pareto: $F(x) = 1 - (1+x/\theta)^{-\alpha}$ with $\alpha = 1.5$. The problem is to estimate the θ parameter by MLE.

First suppose that each group has 50 exposures. The Poisson parameter h can be expressed in terms of the losses and the Pareto parameters using formula (2). Thus the LL function has only the single parameter θ. It can be estimated by MLE to be $\theta = 4.32$M, giving h = 4.46%. Thus 4+ claims are expected ground up, two of which are observed above the retentions. The expected excess claim counts are 1.63 and 0.37 claims for the two groups, respectively, and the median ground-up claim size is $(2^{2/3}-1)\theta = 0.5874\theta = 2.54$M. In effect, both claims are modeled as coming from equally likely independent draws from the same distribution. The expected count difference is due only to the fall-off of the severity distribution for larger claims.

Now take the case where the 1M retention group has 5 exposures and the 10M retention has 95. In this case the MLE for θ is 0.516M with h= 98.7% and the median claim is much smaller at 0.303M. More claims are imputed below the retentions. The expected excess claim counts are now 0.98 and 1.02 for the two groups. The difference in exposure makes a considerable difference in both the parameters and the expected claim counts. All the additional exposures in the upper group make the single claim observed a more remote severity than it was with equal exposures in both groups, but with much more exposure that group is more likely to have a claim

Even if there are no claims in a group, hj will not be zero, and so it will influence the estimation. To illustrate this, take the 2nd example, with 5 exposures excess of 1M and 95 excess of 10M. Now assume there are an additional 100 exposures with retention 50M, and they were claim-free. The θ parameter is now estimated to be 0.437M, which gives a median claim of 0.257M and h = 115.4%. Expected counts by group are now 0.97, 0.94, and 0.09. The lack of larger claims when there is exposure that could have generated them has a downward push on the estimated severity distribution.

The estimation results are summarized in the table below.

Exposures	θ	**Median**	**h**	**h1**	**h2**	**h3**
50/50/0	4.32M	2.54M	4.46%	1.63	0.37	0
5/95/0	0.516M	0.303M	98.7%	0.98	1.02	0
5/95/100	0.437M	0.257M	115.4%	0.97	0.94	0.09