Using Maximum Likelihood Estimation to Parameterize Increased Limits Factors

To explain the proposed method, we show an example of using maximum likelihood estimation to parameterize increased limits factors by finding a loss distribution consistent with given factors. The general purpose of the distribution is for use in simulation of aggregate losses with reinsurance applied. To utilize MLE, a crucial step is to find the implied frequency of claims in each of the ILF bands. The ILF's here are for pure premiums, including risk loads.

Exhibit 1

(1)	(2)	(3)	(4)	(5)	(6)
Limit: x	ILF	Losses Excess Limit: G(x)	% Loss Between Upper & Lower Bounds	|Slope| |G'(x)|	Difference/ Frequency
$25,000	1.000	54.4%	45.6%	18.229	12.432
$30,000	1.064	51.5%	2.9%	5.797	0.322
$35,000	1.124	48.8%	2.7%	5.475	1.610
$40,000	1.166	46.9%	1.9%	3.865	0.322
$50,000	1.244	43.3%	3.5%	3.543	0.966
$75,000	1.385	36.9%	6.4%	2.576	0.837
$100,000	1.481	32.5%	4.3%	1.739	0.580
$125,000	1.544	29.6%	2.9%	1.159	0.129
$150,000	1.601	27.1%	2.6%	1.031	0.193
$175,000	1.647	25.0%	2.1%	0.837	0.193
$200,000	1.682	23.3%	1.6%	0.644	0.064
$250,000	1.746	20.5%	2.9%	0.580	0.193
$300,000	1.788	18.5%	1.9%	0.386	0.145
$500,000	1.894	13.7%	4.8%	0.242	0.142
$1,000,000	2.004	8.7%	5.0%	0.100	0.069
$2,000,000	2.071	5.6%	3.1%	0.031	0.021
$5,000,000	2.131	2.9%	2.7%	0.009	0.009
unlimited	2.194		2.9%		
			100.0%		

In Exhibit 1, $25,000 is the basic limits loss and the "unlimited" figure shown in the table is possible in cases such as Workers Compensation where benefits are uncapped.

Column 3, losses excess limit denoted by G(x), can simply be calculated as ILF at the limit divided by the largest ILF in the table subtracted from unity. For example, G($30,000)= 0.515 = 1 – 1.064/2.194. The ratio expresses the limited losses as a portion of the losses at the highest limit, and subtracting from one gives the portion of losses in-between.

The slope of G(x) in column 5, G′(x), is simply the change in excess losses over the change in the limit, or ΔG(x)/Δx. For the slope, we have taken the absolute value and scaled by 1M. Thus the slope at $50,000 is (0.469 – 0.433)/10,000 times a million, using unrounded numbers throughout. (The unrounded numbers are not shown in the exhibit as the pure premium ILF's themselves were from a calculation in this case.)

Column 6, the differences between the slopes, represents the frequency of claims in each interval. To see this, let's look at G(x):

$$G(x) = \int_x^\infty (y - x)\, f(y)dy \,/\, E[X], \quad \text{where } f(y) \text{ is the severity density.}$$

$$= \int_x^\infty y f(y)dy \,/\, E[X] - x \int_x^\infty f(y)dy \,/\, E[X]$$

$$= \int_x^\infty y f(y)dy \,/\, E[X] - x\, \{ 1 - F(x)\} / E[X]$$

After some calculus and some algebra, we find that the derivative of G comes out to:

$$G'(x) = \{F(x) - 1\}/E[X]$$

Therefore, we see that the difference between two points, x_1 & x_2, on the curve G′(x) is:

$G'(x_2) - G'(x_1) \propto F(x_2) - F(x_1) \propto$ The number of claims in the interval.

So you think we're ready to select a loss distribution and utilize the frequencies in Exhibit 1 to parameterize the ILFs? ***Not so fast!!!*** Here is the rub: As it turns out, the slope we calculate from the ILFs represents the secant through the limit endpoints and not a tangent along the curve G(x) itself. In addition, the midpoint of the interval doesn't always turn out to be the appropriate choice where the tangent is roughly equal to the secant. Through a process of back tracking and using theoretical distributions, we were able to develop a rule for determining the point in the interval where the secant is about equivalent to the tangent at the point. Instead of 50% through the interval, the rule finds the point where the slope of the tangent is equal to that of the secant as 18%×L/U + 33.5% where L is the limit at the lower endpoint and U is the limit at the upper endpoint. Since L < U, this point is always less than 33.5% + 18% = 51.5% through the interval.

For example, between $25,000 and $30,000 the scaled slope when x = $27,425 ={25000 + (.335+.18×25000/30000) × (30000-25000)} is thought to be comparable to the scaled slope of the secant, shown in Exhibit 1, of 5.797. For most of these intervals the tangent does reside roughly halfway through the band, however, it does fall to as low as 41% through the band in the highest layer – See Exhibit 2.

NOW, we are ready to parameterize the ILFs using Maximum Likelihood Estimation. Utilizing the rule above, we can determine the "Adjusted limits" for which our differences/frequencies are appropriate.

Exhibit 2

Original Limit	Adjusted Limit	Fraction of Band width	Difference/ Frequency
$25,000	$27,426	0.485	
$30,000	$32,447	0.489	0.322
$35,000	$37,464	0.493	1.610
$40,000	$44,791	0.479	0.322
$50,000	$61,376	0.455	0.966
$75,000	$86,751	0.470	0.837
$100,000	$111,976	0.479	0.580
$125,000	$137,126	0.485	0.129
$150,000	$162,233	0.489	0.193
$175,000	$187,314	0.493	0.193
$200,000	$223,951	0.479	0.064
$250,000	$274,251	0.485	0.193
$300,000	$388,601	0.443	0.145
$500,000	$712,501	0.425	0.142
$1,000,000	$1,425,001	0.425	0.069
$2,000,000	$3,221,001	0.407	0.021
$5,000,000			0.009

Note the reversals in the last column. These are not intuitively likely, suggesting some inconsistencies in the ILF's. In any case, the loss distribution function we have decided to parameterize here is the transformed beta. This distribution was found to fit our data very well. The frequencies in the exhibit are shown on the line with the upper limit.

Grouped MLE with the data truncated at $27,426 was utilized. In addition to these constraints, an additional constraint was imposed to force the % of losses xs of $5m to be

2.9% as indicated by the data. This constraint was added to use the last data point that says 2.9% of the dollar losses are above $5M. This is difficult to convert to an assumption about claim counts, so it was set as a separate constraint. This forces the excess loss function G to be very close to this point, where a typical MLE would give weight to the claim count above the last point but would not require an exact match. In addition, because one goal is to scale this in the future to the average claim cost, the mean was forced to be the sample mean by making the scale parameter a function of the other parameters and the mean. Thus the MLE was for the shape parameters, with the scale parameter determined to match the mean, and the last data point representing the sum of the censored claims, not their count.

The function to minimize is thus the negative log likelihood of the grouped data plus the added constraint, i.e.:

$$-\sum n_i \times \mathrm{Ln}[\{F(y_i) - F(y_{i-1})\}/\{1-F(27,426)\}] + |G(5,000,000) - .029|$$

Where n_i are the number of implied claims in each interval. G(5,000,000) is also a function of the parameters being estimated.

As an alternative approach, parameters were also estimated by minimizing the sum of absolute percentage errors of the excess factors (EFs). This minimizes:

$$\sum |\text{fitted EF} - \text{Actual EF}|/[\text{Actual EF}].$$

Here the mean was also forced to be the sample mean by making the scale parameter a function of the other parameters and the mean. A comparison of fitted EFs to actual EFs is shown for both approaches in Exhibit 3. The percentage error fit is closer here, as that is what was being optimized, but the MLE appears very reasonable on this basis. The percentage error forces the last point to be very close to the actual as does the constrained MLE.

Exhibit 4 is a comparison of parameters and the implied layer claim counts from fits based on MLE and minimum percentage error. Exhibit 5 is a comparison of the two fits on a log scale. The fits imply a different distribution for small claims, but this is not likely to be an issue in reinsurance. Exhibit 6 displays each of the PDFs.

Exhibit 3

Limit: x	Excess Losses: G(x)	% error	MLE
25,000	54.4%	54.9%	59.1%
30,000	51.5%	51.8%	55.9%
35,000	48.8%	49.3%	53.1%
40,000	46.9%	47.0%	50.6%
50,000	43.3%	43.3%	46.5%
75,000	36.9%	36.8%	39.3%
100,000	32.5%	32.5%	34.4%
125,000	29.6%	29.3%	30.9%
150,000	27.1%	26.8%	28.1%
175,000	25.0%	24.8%	25.9%
200,000	23.3%	23.1%	24.1%
250,000	20.5%	20.5%	21.2%
300,000	18.5%	18.6%	19.1%
500,000	13.7%	13.8%	14.0%
1,000,000	8.7%	8.9%	8.9%
2,000,000	5.6%	5.6%	5.5%
5,000,000	2.9%	2.9%	2.9%

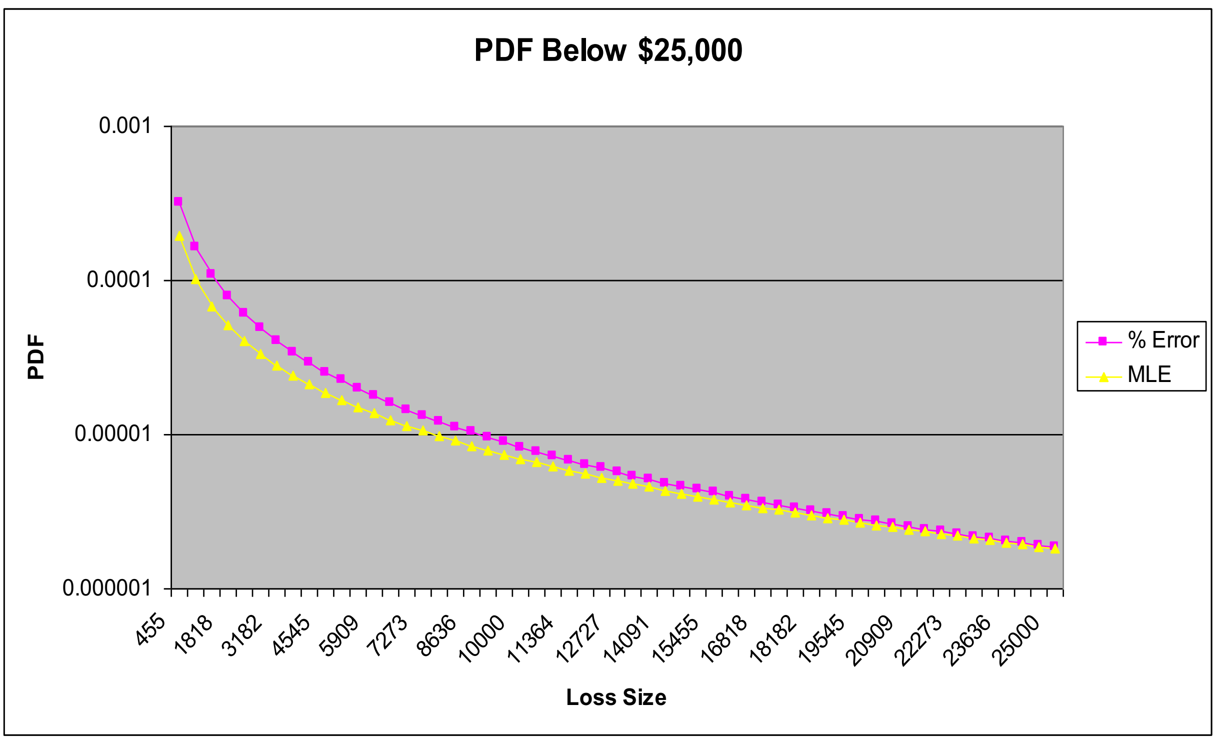
Exhibit 4

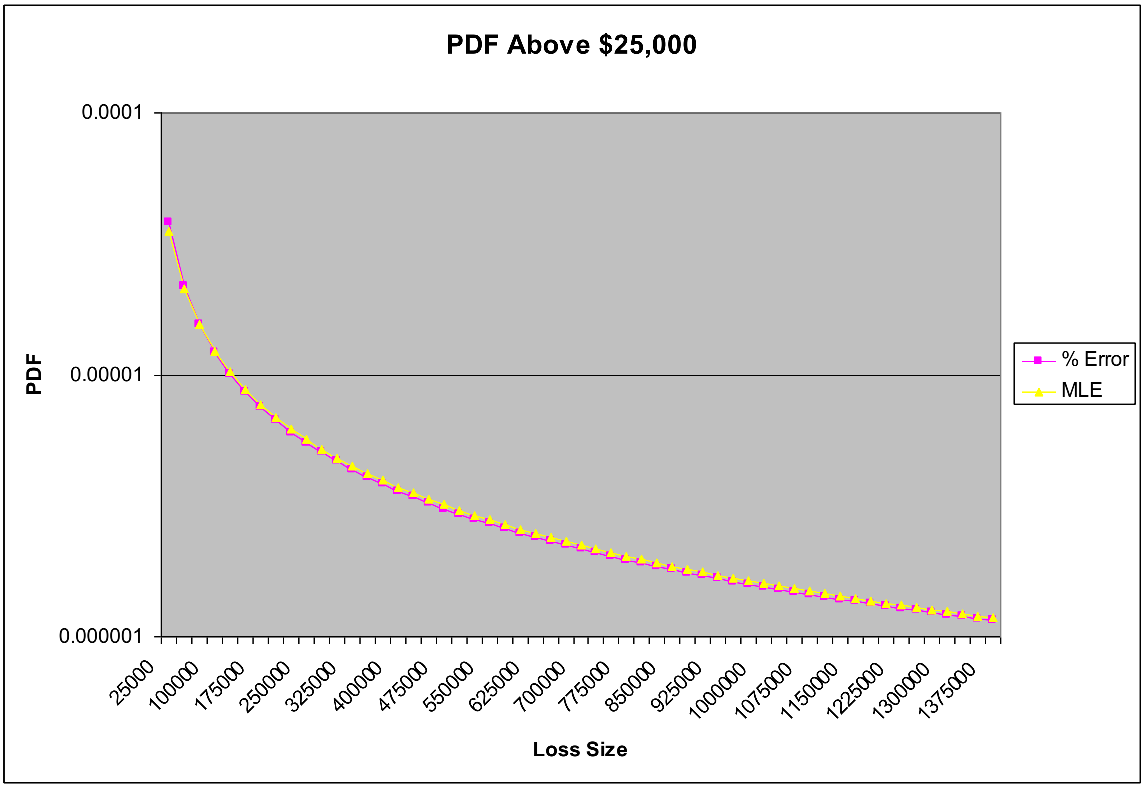
	min % error	MLE
Beta	0.484	0.156
Alpha	1.839	1.745
Tau	0.43	0.69
Theta	18963	46078
Mean	7,259	7,259

				Claims above Upper Limit	
Lower Lim	Upper Lim	Error Method	MLE	Error Method	MLE
0	10,000	892,751	894,900	107,249	105,100
10,001	25,000	58,691	53,296	48,558	51,804
25,001	50,000	24,711	25,395	23,847	26,409
50,001	100,000	13,172	14,477	10,675	11,932
100,001	250,000	7,434	8,395	3,241	3,538
250,001	500,000	2,033	2,262	1,208	1,276
500,001	1,000,000	784	844	423	432
1,000,001	2,000,000	282	292	141	140
2,000,001	3,000,000	68.3	68.6	72.7	71.4
3,000,001	4,000,000	27.6	27.4	45.1	44.0
4,000,001	5,000,000	14.1	13.8	31.0	30.2
5,000,001	10,000,000	21.5	20.9	9.5	9.3
10,000,001	15,000,000	4.8	4.7	4.7	4.6
15,000,001	20,000,000	1.8	1.8	2.9	2.8
20,000,001	30,000,000	1.5	1.4	1.4	1.4
30,000,000 Over		1.4	1.4	-	-
		1,000,000	1,000,000		

Exhibit 5

Exhibit 6





Conclusions

Fitting by minimizing absolute relative errors in the excess factors is ad hoc, but it gets closest to matching the actual factors. Sometimes this method can go wild and find some remote part of the distribution that best matches the excess factors, but which doesn't make sense in applications. Constraining the fit to match the mean, when possible, might keep the parameters in a reasonable range. The closeness of the MLE distribution to the minimum relative error in this example supports both. The MLE method is less arbitrary in its choice of function to minimize, so has some appeal in theory. It might best be used as a reasonableness check or as an alternative when the percent error method gives odd results. The fitting method can also point out anomalies in the excess factors themselves, which may call for further examination of their calculation.

MLE implies that the NCCI procedure produces a distribution with several modes. That's not impossible but doesn't seem likely to be right. It might come from mixing of different injury types, each with their own modes. Fitting a distribution to the excess factors by using minimum % error seems to give more rational factors than the original (in that only one mode is implied).