Actuarial Applications of the FFT to Computing Aggregate Loss Distributions

Stephen Mildenhall
September 2001

Overview

- What are aggregate loss distributions?
- Why are they important?
- How to compute ALDs using FFTs
- Application to modeling loss development

What are Aggregate Loss Distributions?

- Sum of a random number of IID random variables
- $S = X_1 + \ldots + X_N$
  - X_i are IID random variables, generally non-negative, continuous, may be counting
  - N is a frequency distribution, supported on non-negative or positive integers
- Trivial example: $X_i := 1$, $S=N$

Why are ALDs useful?

- Insurance
  - Determine aggregate losses from insured portfolio
  - Split total losses into number of claims or frequency, N, and size of each claim X_i
  - Frequency / severity split is ubiquitous
    - Number of patients and number of bed-days for each
    - Number of vehicles and number of occupants per vehicle
    - Number of accidents per insured and number of insureds having accidents
  - ALDs needed to price aggregate features
    - Health insurance, reinsurance, commercial policies
- Divisible distribution when N is Poisson

When are ALDs not needed?

- To compute impact of limit or deductible on a per occurrence basis
  - Auto insurance deductibles
- To compute mean of ALD:
  - Mean = (Avg Freq) x (Avg Severity)
- Needed for aggregate features:
  - Aggregate deductible
    - Health Insurance, reinsurance
    - Applies to total costs in a year, from one or more occurrences
  - Aggregate limit
    - Products Liability Insurance

How to Compute

- Simulation…
- Panjer Recursion
- Fourier Transform based methods:
  - $M_S(t) = M_N(\log(M_X(t))$
  - $M_X(t) = E[\exp(itX)]$ is characteristic function
  - Heckman-Meyers: continuous Fourier transform
  - Fast Fourier Transform methods

FFT

- FFT is Fast method of computing a discrete FT
- Discrete FT is a sample of continuous FT (in argument t)
- FFT is a transform $C^n \rightarrow C^n$ with an inverse
- FFT(v) = Wv where W is a matrix of roots of unity, v in C^n
- IFFT(v) = (1/n)W*v, W* = complex conjugate of W
  - W is symmetric
  - Orthogonal basis for C^n

FFT

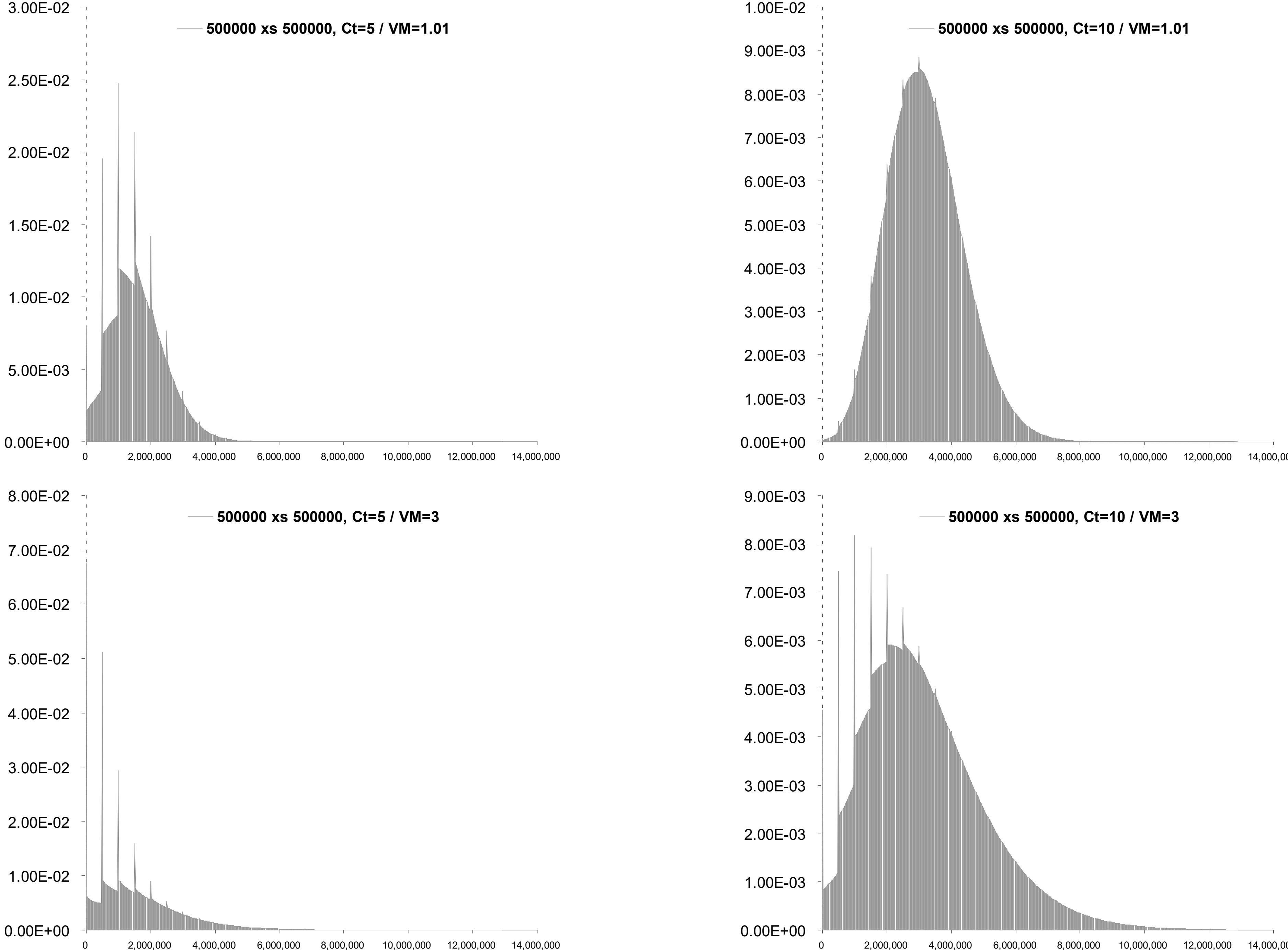
- Special features of W matrix allow FFTs to be computed very efficiently, especially when n is a power of 2 or product of small primes
  - O(n log(n)) time vs $O(n^2)$ for naïve approach
    - Practical vs impossible
- FFTW, http://www.fftw.org: C code which customizes itself to your machine
- Generative Programming: efficient C++ code for small FFT sizes
- Intel Signal Processing Library
- Matlab, SAS
- Don't use Excel built-in routines

Computing ALDs using FFTs

- Sample severity distribution X
  - Compute n x 1 vector of probabilities **v**
    - v_j = Pr((j-1/2)u < X <= (j+1/2)u), j = 0, … , n-1
    - u, the unit, determines the scale
    - **v** is an n x 1 vector of real numbers summing to 1.0
- Take FFT of **v**
  - FFT(**v**) is an n x 1 vector of complex numbers
- Compute **z** = M_N(log(FFT(**v**))
  - Component-by-component calculation
  - M_N(t) is often a function of exp(t), so no need to take logs
  - **z** is n x 1 vector of complex numbers

Computing ALDs using FFTs

- ALD is IFFT(**z**), another n x 1 real vector
  - **z** has symmetry properties required to make IFFT(z) real
- Makes working with aggregate distributions just as practical as working with Normal, lognormal, gamma etc.
- Examples
  - FFT method best suited to small claim count
  - Excess reinsurance perfect application
  - Real-world ALDs often not continuous


500000 xs 500000, Ct=5 / VM=1.01
500000 xs 500000, Ct=10 / VM=1.01
500000 xs 500000, Ct=5 / VM=3
500000 xs 500000, Ct=10 / VM=3
3.00E-02
2.50E-02
2.00E-02
1.50E-02
1.00E-02
5.00E-03
0.00E+00
1.00E-02
9.00E-03
8.00E-03
7.00E-03
6.00E-03
5.00E-03
4.00E-03
3.00E-03
2.00E-03
1.00E-03
8.00E-02
7.00E-02
6.00E-02
5.00E-02
4.00E-02
3.00E-02
2.00E-02
0
2,000,000
4,000,000
6,000,000
8,000,000
10,000,000
12,000,000
14,000,000

Loss Development

- What is loss development?
  - Insurance claims can take many years to be reported, and many more years to be settled
  - Losses from events occurring during a particular year are grouped and tracked as an “Accident Year”
  - Subsequent evaluations of an accident year increase, or develop
    - Displayed in loss development triangles
  - Insurance companies must estimate ultimate value of losses from each accident year as soon as possible
    - Process of setting reserves

Loss Development Triangle

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1984	1,000	3,796	9,604	19,536	29,651	54,121	87,576	114,209	123,156	132,280	115,923	133,159	132,372	126,845	135,858	125,719	120,468
1985	1,000	2,468	4,790	5,822	9,425	12,300	13,139	12,934	13,786	14,714	16,698	16,521	17,075	17,231	18,452	19,630	
1986	1,000	3,724	11,239	15,457	22,330	30,647	35,389	33,576	35,661	40,357	46,596	51,579	51,556	47,640	51,032		
1987	1,000	3,581	11,010	18,046	25,175	31,163	37,340	46,333	52,236	48,034	49,572	51,670	55,822	54,927			
1988	1,000	4,297	9,371	18,697	25,590	25,890	31,474	35,654	39,141	43,230	41,825	42,327	43,394				
1989	1,000	3,323	9,944	18,925	34,760	53,431	58,346	55,933	52,618	48,723	42,172	43,147					
1990	1,000	3,819	7,039	11,485	11,295	17,573	23,147	25,555	27,474	32,448	34,068						
1991	1,000	2,423	7,188	10,500	14,087	26,878	30,291	33,352	36,260	37,544							
1992	1,000	2,621	6,216	15,075	23,544	24,191	28,352	28,464	27,251								
1993	1,000	3,534	5,857	16,548	31,203	48,278	72,474	78,287									
1994	1,000	3,289	7,784	25,881	37,028	48,470	59,244										
1995	1,000	3,208	9,751	16,895	25,804	34,542											
1996	1,000	5,044	13,298	16,697	23,436												
1997	1,000	4,328	7,545	10,851													
1998	1,000	3,636	10,853														
1999	1,000	4,014															
2000	1,000																

Loss Development

- Loss development = reserving = complete the square
- Interested in distribution of R and splitting U=I+R where
  - U = ultimate losses
  - I = incurred-to-date losses (known at date)
  - R = reserves
- Can model each piece as an aggregate loss distribution
  - Both frequency and severity parts develop:
    - Incurred but not reported (IBNR): count development
    - Reported but not accurately reserved: severity development of known claims

Loss Development

- Development Factors: $U = f\,L$
  - f is a loss development factor
  - L is losses to date

Loss Development Triangle

Link Ratios

AY	1 : 2	2 : 3	3 : 4	4 : 5	5 : 6	6 : 7	7 : 8	8 : 9	9 : 10	10 : 11	11 : 12	12 : 13	13 : 14	14 : 15	15 : 16	16 : 17
1984	3.796	2.530	2.034	1.518	1.825	1.618	1.304	1.078	1.074	0.876	1.149	0.994	0.958	1.071	0.925	0.958
1985	2.468	1.941	1.215	1.619	1.305	1.068	0.984	1.066	1.067	1.135	0.989	1.034	1.009	1.071	1.064	
1986	3.724	3.018	1.375	1.445	1.372	1.155	0.949	1.062	1.132	1.155	1.107	1.000	0.924	1.071		
1987	3.581	3.075	1.639	1.395	1.238	1.198	1.241	1.127	0.920	1.032	1.042	1.080	0.984			
1988	4.297	2.181	1.995	1.369	1.012	1.216	1.133	1.098	1.104	0.968	1.012	1.025				
1989	3.323	2.992	1.903	1.837	1.537	1.092	0.959	0.941	0.926	0.866	1.023					
1990	3.819	1.843	1.632	0.983	1.556	1.317	1.104	1.075	1.181	1.050						
1991	2.423	2.966	1.461	1.342	1.908	1.127	1.101	1.087	1.035							
1992	2.621	2.372	2.425	1.562	1.028	1.172	1.004	0.957								
1993	3.534	1.657	2.826	1.886	1.547	1.501	1.080									
1994	3.289	2.367	3.325	1.431	1.309	1.222										
1995	3.208	3.040	1.733	1.527	1.339											
1996	5.044	2.637	1.256	1.404												
1997	4.328	1.743	1.438													
1998	3.636	2.984														
1999	4.014															

Averages

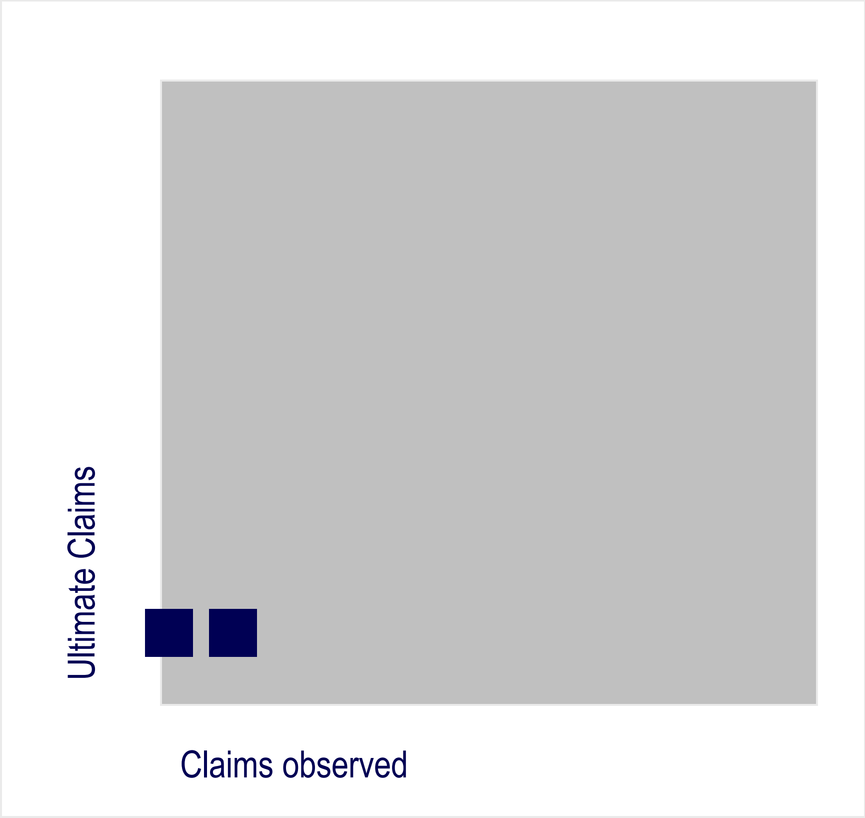
Strt All	3.569	2.490	1.875	1.486	1.415	1.244	1.086	1.055	1.055	1.012	1.054	1.027	0.969	1.071	0.995	0.958
Last 5	4.046	2.554	2.115	1.562	1.426	1.268	1.050	1.032	1.033	1.014	1.035	1.027	0.969	1.071	0.995	0.958
Last 3	3.993	2.455	1.476	1.454	1.398	1.298	1.062	1.040	1.047	0.961	1.026	1.035	0.972	1.071	0.995	0.958
Wtd All	3.569	2.477	1.827	1.495	1.406	1.278	1.112	1.056	1.045	0.964	1.082	1.017	0.960	1.071	0.942	0.958
Wtd L5	4.046	2.524	1.964	1.548	1.385	1.291	1.042	1.021	1.011	1.007	1.043	1.017	0.960	1.071	0.942	0.958
Wtd L3	3.993	2.437	1.453	1.451	1.396	1.324	1.069	1.041	1.020	0.949	1.027	1.036	0.963	1.071	0.942	0.958
Selecte	3.569	2.490	1.875	1.486	1.415	1.244	1.086	1.055	1.055	1.012	1.054	1.027	0.969	1.071	0.995	0.958
FTU	56.98	15.97	6.413	3.419	2.301	1.627	1.307	1.204	1.142	1.082	1.070	1.015	0.989	1.021	0.953	0.958

Loss Development

- What is distribution of f?
  - Bootstrap from age-to-age factors
  - Quick, easy, few assumptions
  - Simulation from triangles with known distributions indicates method works reasonably well
    - Variance under estimated

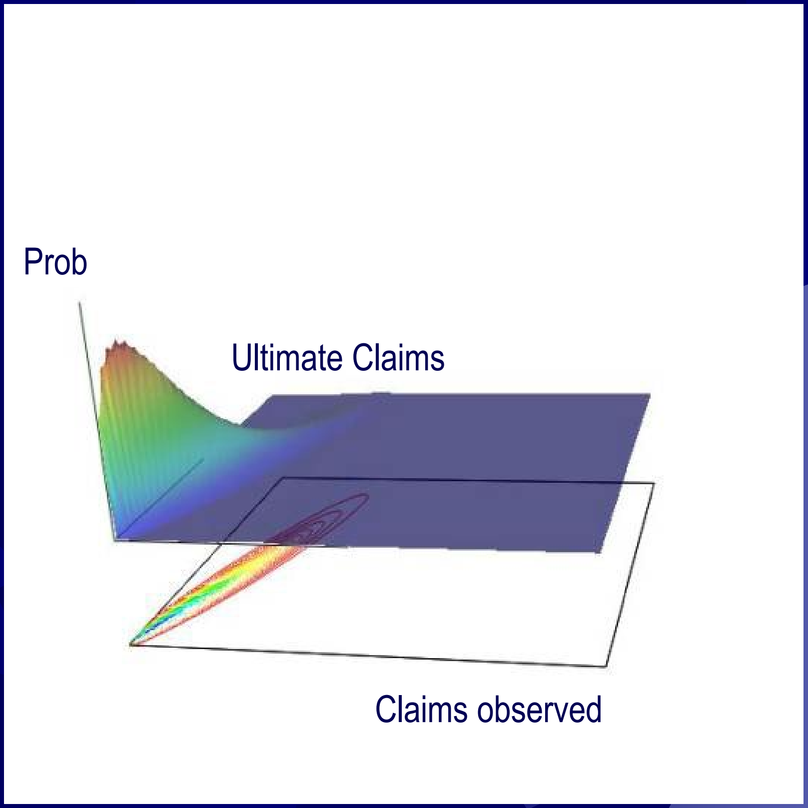
Simple Model for Claim Counts

- Assume severity ≡ 1
- At evaluation point, estimate
  - Pr(Claim observed)
  - Binomial distribution
- Use 2D FFT to aggregate over expected ultimate claims
- Bivariate distribution of (Obs Claims, Ult Clms)
- Transform for distribution of f
- Posterior ultimate distributions
- Extend for losses in obvious way



Simple Model

- Method produces f-distributions with CV and skewness too low
- "Over-dispersion" parameters possible but lacking motivation
- Mixture possible



Application of Bivariate Dist

- Given distribution of (L, U), observed loss L and prior ultimate loss U
- Observe actual losses L=l
- Conditional distribution gives revised distribution of U | L=l
  - Confidence intervals for reserves
  - Iterate over years
- Reserving considering relative variability in U and L or f
  - Would represent major advance in technique

References

- My website for aggregate distribution tools and further examples of techniques: http://www.mynl.com/MALT/home.html
- Casualty actuarial website: http://www.casact.org
- FFT tools: http://www.fftw.org
- Generative Programming, in Techniques for Scientific C++ at http://extreme.indiana.edu/~tveldhui/papers/techniques/
- Intel Signal Processing library (can be called from Excel VBA) at http://developer.intel.com/software/products/perflib/spl/index.htm
- Graphics library: http://www.kitware.com