Working Paper n. 08 - 2016

AGGREGATE LOSS DISTRIBUTION AND DEPENDENCE: COMPOSITE MODELS, COPULA FUNCTIONS AND FAST FOURIER TRANSFORM FOR THE DANISH RE INSURANCE DATA

Rocco Roberto Cerchiara
Dipartimento di Economia, Statistica e Finanza
Università della Calabria
Ponte Pietro Bucci, Cubo 1/C
Tel.: +39 0984 492350
Fax: +39 0984 492421
e-mail: rocco.cerchiara@unical.it

Francesco Acri
Dipartimento di Scienze Statistiche
Università degli Studi di Roma “La Sapienza”
Viale Regina Elena, 295,
Palazzina F
e-mail: francesco.acri@uniroma1.it

Settembre 2016
Aggregate Loss Distribution and Dependence: Composite Models, Copula functions and Fast Fourier Transform for the Danish fire insurance data

Rocco Roberto Cerchiara¹ and Francesco Acri²

¹Assistant Professor at Department of Economics, Statistics and Finance, University of Calabria, Arcavacata di Rende (Cosenza), Italy
²Actuarial Science PhD student at Department of Statistical Science, Sapienza University of, Rome, Italy

¹rocco.cerchiara@unical.it
²francesco.acri@uniroma1.it

Abstract

Danish fire insurance data has been analyzed in several papers, using different models. In this paper we investigate the improving of the fitting for the Danish fire insurance data according to composite models, including dependence structure by copula functions and Fast Fourier Transform.

Keywords: Composite models, Copula, Fast Fourier Transform

JEL Classification: C10, C63, C65, G22

1 Introduction

The evaluation of the distribution of aggregate loss plays a fundamental role in the analysis of risk and solvency levels of an insurance company. In literature many different studies are based on definition of composite models which aim is to analyze this distribution and dependence between the main factors that characterize the risk profile of insurance companies, e.g. frequency-severity, attritional-large claims.

A composite model is a combination of two different models, one having a light tail below a threshold (attritional claims) and another with a heavy
tail suitable to model value that exceed this threshold (large claims). Composite distributions (also known as compound, spliced or piecewise distributions) have been introduced in many applications. Klugman et al. (2010) expressed the probability density function of a composite distributions as:

\[
f(x) = \begin{cases} 
  r_1 f^*_1(x) & k_0 < x < k_1 \\
  \vdots & \\
  r_n f^*_n(x) & k_{n-1} < x < k_n 
\end{cases}
\]

where \( f^*_j \) is truncated probability density function of marginal distribution \( f_j \), \( j=1,\ldots,n \); \( r_j \) are mixing weights; \( k_j \) define the range limit of the domain.

The Danish reinsurance data has been often analyzed according using a parametric approach and composite models. Cooray and Ananda (2005) and Scollnik (2007) show that the composite lognormal-Pareto model could fit better than standard univariate models. Following the previous two papers, Teodorescu and Vernic (2009 and 2013) fit the dataset firstly with a composite Exponential and Pareto distribution and then with a more general composite Pareto model, obtained by replacing the Lognormal distribution by an arbitrary continuous distribution, while Pigeon and Denne (2011) consider in the composite model the threshold value as the realization of a positive random variable. There have been other several approaches to model this dataset: Burr distribution for claim severity using XploRe computing environment (Burnecki and Weron, 2004), Bayesian estimation of finite time ruin probabilities (Ausin et al., 2009), hybrid Pareto models (Carreau and Bengio, 2009), beta kernel quantile estimation (Charpentier and Oulidi, 2010), bivariate compound Poisson process (Esmaeili and Kluppelberg, 2010). An example on non parametric modelling is shown in Guillotte et al. (2011) with a Bayesian inference on bivariate extremes. Drees and Muller (2008) show how to model dependence within joint tail regions. Nadarajah and Bakar (2014) improve the fittings for the Danish fire insurance data using various new composite models, including the composite Lognormal-Burr model.

Regarding the Danish fire insurance data in this paper we investigate the use of different composite models and Extreme Value Theory (EVT, see Embrechts et al., 1997 and McNeil et al., 2005), Copula function and Fast Fourier Transform-FFT (Klugman et al, 2010) in order to analyze the effect of the dependence between attritional and large claims as well.

The paper is organized as follows. In Sections 2 and 3 we suppose there isn’t any dependence between attritional and large claims and we investigate the use of composite models and a compound model with random threshold in order to fit the Danish fire insurance data, comparing our numerical results with the fitting of Nadarajah and Bakar (2014) based on composite Lognormal-Burr model. In Sections 4 we try to appraise risk dependence through the concept of copula function and FFT. Section 5 concludes the work, where we present the estimation of VaR of aggregate loss distribution, comparing results under independence or dependence conditions.
2 Composite models

In the Danish reinsurance data we can find both frequent claims with low-medium severity and sporadic claims with high severity. If we want to define a joint distribution for these two types of claims we have to build a composite model.

Formally, the density distribution of a composite model can be written as a special case of (1):

$$f(x) = \begin{cases} \quad rf_1^*(x), & -\infty < x \leq u \\ (1-r)f_2^*(x), & u < x < \infty \end{cases}$$

(2)

where $r \in [0, 1]$, $f_1^*$ and $f_2^*$ are cut off density distributions of marginals $f_1$ and $f_2$ respectively. In details, if $F_i$ is distribution function of $f_i$, $i=1,2$, then we have

$$\begin{align*}
  f_1^*(x) &= f_1(x) F_1(u), \quad -\infty < x \leq u \\
  f_2^*(x) &= f_2(x), \quad u < x < \infty
\end{align*}$$

(3)

It’s simple note that (2) is a convex combination of $f_1^*$ and $f_2^*$ with weights $r$ and $1-r$. In addition, we want that (2) is a continuous, derivable and with continuous derivative density function and for this scope we fix the following limitation:

$$\begin{align*}
  \lim_{x \rightarrow u^-} f(x) &= f(u) \\
  \lim_{x \rightarrow u^+} f'(x) &= \lim_{x \rightarrow u^+} f'(x)
\end{align*}$$

(4)

From the first we obtain

$$r = \frac{f_2(u)F_1(u)}{f_2(u)F_1(u) + f_1(u)(1-F_2(u))}$$

(5)

while from the second

$$r = \frac{f_2(u)F_1(u)}{f_2(u)F_1(u) + f_1(u)(1-F_2(u))}$$

(6)

We can define distribution function $F$ of (2)

$$F(x) = \begin{cases} \quad rF_1(x), & -\infty < x \leq u \\
  r + (1-r)\frac{F_2(x)-F_2(u)}{1-F_2(u)}, & u < x < \infty \end{cases}$$

(7)

Suppose $F_1$ and $F_2$ admit inverse function; we can define quantile function via inversion method: let be $p$ a random number from a standard Uniform distribution, the quantile function results

$$\begin{align*}
  F^{-1}(x) &= \begin{cases} 
  F_1^{-1}\left(\frac{p}{r}F_1(u)\right), & p \leq r \\
  F_2^{-1}\left(\frac{p-r+(1-p)F_2(u)}{1-r}\right), & p > r
  \end{cases}
\end{align*}$$

(8)
To estimate the parameters of (7) we can proceed as follows: first of all we estimate marginal density function parameters separately (the underlying hypothesis is that there isn’t any relation between attritional and large claims); then these estimates will be the start values of density function in order to maximize the following likelihood:

\[ L(x_1, \ldots, x_n; \theta) = r^m (1 - r)^{n-m} \prod_{i=1}^{m} f^*_1(x_i) \prod_{j=m+1}^{n} f^*_2(x_j) \]  

where \( n \) is sample dimension, \( \theta \) is a vector including compound model parameters, while \( m \) is such that \( X_m \leq v \leq X_{m+1} \), otherwise it’s the level of order statistics immediately previous (or coincident) to \( v \).

The methodology described in Teodorescu & Vernic (2009 and 2013) has been used in order to estimate the threshold \( u \) which permit us to discriminate between attritional and large claims.

### 2.1 Compound model with random threshold

We can define a compound model using also a random threshold (see Pi- geon and Denuit, 2011). In particular, given random sample \( X = (X_1, \ldots, X_n) \), we can assume that every single component \( X_i \) provides a own threshold. So, for the generic observation \( x_i \) we’ll have the threshold \( u_i, i = 1, \ldots, n \). For this reason, \( u_1, \ldots, u_n \) are realizations of a random variable \( U \) with a distribution function \( G \). The random variable \( U \) is necessarily non-negative and with a heavy tailed distribution.

A compound model with random threshold shows a completely new and original aspect: we cannot be able to choose only a value for \( u \) but its whole distribution and the parameters of the latter are implicit in the definition of the compound model. In particular, we define the density function of Lognormal-Generalized Pareto Distribution model (GPD, see Embrechts et al, 1997) with random threshold in the following way:

\[ f(x) = (1 - r) \int_0^x f_2(x)g(u)du + r \int_x^\infty \frac{1}{\Phi(\xi\sigma)} f_1(x)g(u)du \]  

where \( r \in [0, 1] \), \( u \) is the random threshold whit density function \( g \), \( f_1 \) and \( f_2 \) are Lognormal and GPD density functions, respectively, \( \Phi \) is the Standard Normal distribution function, \( \xi \) is the shape parameter of GPD and \( \sigma \) is Lognormal scale parameter.

### 2.2 Kumaraswamy Distribution and some generalization

In this section we describe the Kumaraswamy Distribution (see Ku- maraswamy, 1980) and a generalization of Gumbel distribution (see Cordeiro et al., 2012). In particular, let

\[ K(x; \alpha, \beta) = 1 - (1 - x^\alpha)^\beta, x \in (0, 1) \]
the distribution proposed in Kumaraswamy (1980), where parameter $\alpha$ and $\beta$ establish it’s trend. If $G$ is the distribution function of a random variable, then we can define a new distribution by

$$F(x; a, b) = 1 - (1 - G(x)^a)^b$$ \hspace{1cm} (12)

where $a > 0$ and $b > 0$ are shape parameters that influence kurtosis and skewness. The Kumaraswamy-Gumbel (KumGum) distribution is defined throughout (12) with the following distribution function (see Cordeiro et al., 2012):

$$F_{KG}(x; a, b) = 1 - (1 - \Lambda(x)^a)^b$$ \hspace{1cm} (13)

where $\Lambda(x)$ is Gumbel distribution function. The quantile function of KumGum is obtained by inverting (13) and expliciting Gumbel parameters ($u$ and $\phi$):

$$x_p = F^{-1}(p) = u - \varphi \log \left[ -\log (1 - (1 - p)^{1/b})^{1/a} \right]$$ \hspace{1cm} (14)

with $p \in (0, 1)$.

The following table and Figure 1 show Kurtosis and Skewness of KumGum density function by varying the four parameters:

<table>
<thead>
<tr>
<th>$u$</th>
<th>$\varphi$</th>
<th>$a$</th>
<th>$b$</th>
<th>Kurtosis</th>
<th>Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>5.4</td>
<td>1.1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0.5</td>
<td>0.5</td>
<td>7.1</td>
<td>1.6</td>
</tr>
<tr>
<td>-5</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3.6</td>
<td>0.5</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>5</td>
<td>0.7</td>
<td>6.4</td>
<td>1.4</td>
</tr>
<tr>
<td>0</td>
<td>15</td>
<td>1</td>
<td>0.4</td>
<td>7.6</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Table 1: Kurtosis and Skewness of Kum-Gum distribution
Figure 1: KumGum density functions.

Another generalization of Kum distribution is the Kumaraswamy-Pareto distribution (Kum-Pareto); in particular, we can evaluate equation (12) in the Pareto distribution function $P$ which is

$$P(x; \beta, \kappa) = 1 - \left( \frac{\beta}{x} \right)^\kappa, x \geq \beta$$  \hspace{1cm} (15)

where $\beta > 0$ is a scale parameter and $\kappa \geq 0$ is a shape one. So from (11), (12) and (15) we obtain the Kum-Pareto distribution function:

$$F_{KP}(x; \beta, \kappa, a, b) = 1 - \left\{ 1 - \left[ 1 - \left( \frac{\beta}{x} \right)^\kappa \right]^a \right\}^b$$  \hspace{1cm} (16)

The corresponding quantile function is

$$F^{-1}(p) = \beta \left\{ \left\{ 1 - \left[ 1 - \left( 1 - p \right)^{1/b} \right]^{1/a} \right\}^{1/\kappa} \right\}^{-1}$$  \hspace{1cm} (17)

where $p \in (0, 1)$. In the following figure we report Kum-Pareto density function varying the parameters:
In this section we present some numerical results on the fitting of the Danish reinsurance data by composite models with a constant and with random threshold between attritional and large claims. As already mentioned, for the composite models with a constant threshold, we used the methodology described in Teodorescu & Vernic (2009 and 2013), obtaining $u = 1,022,125\€$. Regarding the main statistics of Danish reinsurance data see Embrechts et al. (1997). We start with a compound model Lognormal-KumPareto, choosing $f_1 \sim \text{Lognormal}$ and $f_2 \sim \text{KumPareto}$. From the following table we can compare some theoretical and empirical quantiles:

<table>
<thead>
<tr>
<th>Level</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
<th>99.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empirical quantile</td>
<td>327,016</td>
<td>532,757</td>
<td>1,022,213</td>
<td>1,675,219</td>
<td>5,484,150</td>
<td>8,216,877</td>
</tr>
<tr>
<td>Theoretical quantile</td>
<td>333,477</td>
<td>462,852</td>
<td>642,196</td>
<td>840,161</td>
<td>2,616,338</td>
<td>4,453,476</td>
</tr>
</tbody>
</table>

Table 2: Comparison between empirical and Lognormal-KumPareto quantiles
Only the fiftieth percentile of theoretical distribution function is very close to the same empirical quantile: from this percentile onwards the differences increase. In the following figure we show only right tails of the distribution functions (empirical and theoretical):

![Right tail of distributions](image)

Figure 3: Right tails of Lognormal-KumPareto (red line) and empirical distribution (dark line) functions.

The red line stands ever over dark line. Kumaraswamy generalized families of distributions are very versatile to analyze different types of data this means, but in this case the Lognormal-KumPareto model underestimates the right tail.

So we consider the compound model \( f_1 \sim \text{Lognormal} \) and \( f_2 \sim \text{Burr} \) as suggested in Nadarajah and Bakar (2014). The parameters are estimated using the CompLognormal R package as shown in Nadarajah and Bakar (2014). From the following table we can compare some theoretical quantiles with empirical ones:

<table>
<thead>
<tr>
<th>Level</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
<th>99.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empirical quantile</td>
<td>327,016</td>
<td>532,757</td>
<td>1,022,213</td>
<td>1,675,219</td>
<td>5,484,150</td>
<td>8,216,877</td>
</tr>
<tr>
<td>Theoretical quantile</td>
<td>199,681</td>
<td>332,341</td>
<td>634,531</td>
<td>1,029,262</td>
<td>3,189,937</td>
<td>5,181,894</td>
</tr>
</tbody>
</table>

Table 3: Comparison between empirical quantiles and Lognormal-Burr ones

The model seems to be more feasible to catch the right tail of empirical distribution respect to the previous Lognormal-KumPareto, as we can see from the figure below:
As Lognormal-KumPareto model, the Lognormal-Burr distribution stands ever over the empirical distribution but not always at the same distances.

We go forward modelling a Lognormal-Generalized Pareto Distribution (GPD), that is we choose $f_1 \sim \text{Lognormal}$ and $f_2 \sim \text{GPD}$ and then we generate pseudo-random numbers from quantile function (8). In the following we report the estimates of parameters and the QQ-plot:

<table>
<thead>
<tr>
<th></th>
<th>low extreme</th>
<th>best estimate</th>
<th>high extreme</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1$</td>
<td>12.82</td>
<td>12.84</td>
<td>12.86</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.59</td>
<td>0.61</td>
<td>0.62</td>
</tr>
<tr>
<td>$\sigma_\mu$</td>
<td>1,113,916</td>
<td>1,115,267</td>
<td>1,116,617</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.33</td>
<td>0.45</td>
<td>0.56</td>
</tr>
</tbody>
</table>

Table 4: Estimated parameters of Lognormal-GPD

$\mu_1$ and $\sigma$ are the Lognormal parameters, while $\sigma_\mu$ and $\xi$ are GPD parameters.
We observe that this compound model has a good adaptation to empirical distribution; in fact, except many, theoretical quantiles are close to corresponding empirical quantiles. In the following figure we compare theoretical cut-off density function with corresponding empirical one and theoretical right tail with empirical one:

Figure 5: Observed-theoretical quantile plot for the Lognormal-GPD model.

Figure 6: Left, comparison between cut-off density functions. Right, empirical and theoretical (red) right tail.
The model exhibits a non-negligible right tail (kurtosis index is 115,656.2) which can be evaluated comparing observed distribution function with the theoretical one:

![Empirical vs theoretical distributions](image)

Figure 7: Lognormal-GPD (red) and empirical (dark) distribution function.

The corresponding Kolmogorov-Smirnov test has returned a p-value equal to 0.8590423, using 50,000 bootstrap samples.

Finally, we report the best estimate and 99% confidence intervals of the compound model Lognormal-GPD with a Gamma random threshold (see Pingeon and Denuit, 2011):

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Low Extreme</th>
<th>Best Estimate</th>
<th>High Extreme</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1$</td>
<td>12.78</td>
<td>12.79</td>
<td>12.81</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.52</td>
<td>0.54</td>
<td>0.55</td>
</tr>
<tr>
<td>$u$ (threshold)</td>
<td>629,416</td>
<td>630,768</td>
<td>632,121</td>
</tr>
<tr>
<td>$\sigma_{\mu}$</td>
<td>1,113,915</td>
<td>1,115,266</td>
<td>1,116,616</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.22</td>
<td>0.29</td>
<td>0.37</td>
</tr>
</tbody>
</table>

Table 5: Estimated parameters and 99% confidence intervals of Lognormal-GPD-Gamma distribution

The threshold $u$ is a parameter which value depends on Gamma parameters. In the following we report the theoretical and empirical quantiles:
Table 6: Comparison between empirical and Lognormal-GPD-Gamma quantiles

We can see from the following figure that Lognormal-GPD-Gamma model can be considered a good fitting model:

![Empirical vs theoretical distributions](image)

Figure 8: Lognormal-GPD-Gamma (red) versus empirical (dark) distribution functions.

The Kolmogorov-Smirnov adaptive test returns a p-value equal to 0.1971361. So we cannot reject the null hypothesis under which the investigate model is a feasible model for our data.

Finally Lognormal-KumPareto, Lognormal-Burr, Lognormal-GPD with fixed threshold and Lognormal-GPD with a Gamma random threshold, can be compared using the AIC and BIC values:

<table>
<thead>
<tr>
<th>Index</th>
<th>KumPareto</th>
<th>Burr</th>
<th>GPD</th>
<th>GPD-Gamma</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>193,374</td>
<td>191,459</td>
<td>191,172</td>
<td>190,834</td>
</tr>
<tr>
<td>BIC</td>
<td>193,409</td>
<td>191,494</td>
<td>191,207</td>
<td>190,882</td>
</tr>
</tbody>
</table>

Table 7: AIC and BIC indices for a comparison between different models

The previous analysis suggests that the Lognormal-GPD-Gamma gives the better fit.
4 Introducing dependence structure: Copula Functions and Fast Fourier Transform

In the last section we restricted our analysis to the case of independence between attritional and large claims. We now try to extend this paper to a dependence structure. Firstly we’ll define a compound model using a copula function to evaluate the possible dependence. As marginal distributions we’ll make reference to a Lognormal distribution for attritional claims and a GPD for large ones. The empirical correlation matrix $R$:

$$R = \begin{pmatrix} 1 & 0.01259155 \\ 0.01259155 & 1 \end{pmatrix}$$

and Kendall’s Tau and Spearman’s Rho measures of association:

$$\begin{pmatrix} 1 & 0.002526672 \\ 0.002526672 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0.003730766 \\ 0.003730766 & 1 \end{pmatrix}$$

suggest a weak but positive correlation between normal and large claims.

For this reason, the individuation of an appropriate copula function will not be easy, but we present an illustrative example based on a Gumbel Copula.

The parameters of the Gumbel Copula can be estimated through different methods:

<table>
<thead>
<tr>
<th>Method</th>
<th>$\hat{\theta}$</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum pseudo-likelihood</td>
<td>1.11</td>
<td>0.008</td>
</tr>
<tr>
<td>Canonical maximum pseudo-likelihood</td>
<td>1.11</td>
<td>0.008</td>
</tr>
<tr>
<td>Simulated maximum likelihood</td>
<td>1.11</td>
<td>-</td>
</tr>
<tr>
<td>Minimum distance</td>
<td>1.09</td>
<td>-</td>
</tr>
<tr>
<td>Moments based on Kendall’s tau</td>
<td>1.13</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 8: Different methods for estimating the dependence parameter of a Gumbel Copula

We remind that Gumbel’s parameter $\theta$ assumes values in $[1, \infty)$ and for $\theta \to 1$ we have independence between marginal distributions. We observe that estimates are significantly different from 1 and so our Gumbel Copula doesn’t correspond to Independent Copula. We can say that because we have verified, using bootstrap procedures, $\theta$ parameter has a Normal distribution. In fact, Shapiro-Wilk test has given a $p$-value equals to 0.08551 and so, fixed a significance level of 5%, it’s not possible reject null hypothesis. In addition, the 99% confidence interval obtained with Maximum pseudo-likelihood method results $[1.090662; 1.131003]$ which doesn’t include the value 1. We report two useful graphics, obtained by simulation of estimated Gumbel:
Figure 9: Lognormal (top) and GPD (right) marginal histograms and Gumbel Copula simulated values plot.

Figure 10: Density function of estimated Gumbel Copula. Attritional claims losses on X-axis, large claims losses on Y-axis.

The density function assumes greater values in correspondence of great values both for Lognormal and GPD marginal; in other words, using that Gumbel Copula, the probability that attritional claims produce losses near...
to the threshold \( u \) and that large claims produce extreme losses is greater
than probability of any other joined event.
In our numerical examples, we’ll refer to Gumbel Copula function despite
having estimated and analyzed other copulas for which no significant dif-
ference for the aims of this paper.

4.1 An alternative to Copula Function: the Fast Fourier Transform

Considering the fact that it is not easy to define an appropriate copula for
this dataset, now we’ll try to model the aggregate loss distribution directly
with the Fast Fourier Transform (FFT) using empirical data. That approach
allows us to avoid the dependence assumption between attritional and large
claims (necessary instead with the copula approach).
To build aggregate loss distribution by FFT it’s necessary, first of all, make
the severity distribution \( Z \) discrete (see Klugman et al., 2010) and obtain
the vector \( z = (z_0, \ldots, z_{n-1}) \) which element \( z_i \) is the probability that single
claim produce a loss equals to \( ic \), where \( c \) is a fixed constant such that, given
\( n \) the length of the vector \( z \), the loss \( cn \) has a negligible probability. We
consider also frequency claim distribution \( \tilde{k} \) through Probability-Generating
function (PGF) defined as

\[
PGF_k(t) = \sum_{j=0}^{\infty} t^j Pr(\tilde{k} = j) = E[t^k]
\]

In particular, let \( FFT(z) \) and \( IFFT(z) \) be the FFT and its inverse respec-
tively, we obtain the discretized probability distribution for the aggregate
loss \( X \) as

\[
(x_0, x_1, \ldots, x_{n-1}) = IFFT(PGF(FFT(z)))
\]

Both \( FFT(z) \) and \( IFFT(z) \) are \( n \)-dimensional vectors which generic ele-
ments are, respectively, \( \tilde{z}_k = \sum_{j=0}^{n-1} z_j \exp(\frac{2\pi i}{n}jk) \) and
\( z_k = \frac{1}{n} \sum_{j=0}^{n-1} \tilde{z}_j \exp(-\frac{2\pi i}{n}jk), i = \sqrt{-1} \).

From a theoretical point of view, this is a discretized version of Fourier
Transform (DFT):

\[
\phi(z) = \int_{-\infty}^{+\infty} f(x) \exp(izx)dx
\]

The characteristic function creates an association between a probability den-
sity function and continue complex one, while the DFT makes an association
between an \( n \)-dimensional vector and an \( n \)-dimensional complex vector. The
former one-to-one association can be done through the algorithm FFT.

For two-dimensional case it’s necessary a matrix \( M_Z \) as input; that matrix
contains joined probabilities of attritional and large claims and is such that
its possible obtain corresponding marginal distributions adding long rows and columns respectively. For example, let

\[
M_z = \begin{pmatrix}
0.5 & 0 & 0 \\
0.2 & 0.25 & 0 \\
0 & 0.05 & 0
\end{pmatrix}
\]

be that matrix. The vector \((0.5, 0.45, 0.05)\), obtained adding long three rows, contains attritional claims marginal distribution, while the vector \((0.7, 0.3, 0)\), obtained adding long three columns, contains large claims marginal distribution. The single element of the matrix, instead, is the joined probability. The aggregate loss distribution will be a matrix \(M_X\) given by

\[
M_x = IFFT(PGF(FFT(M_z)))
\]  

(21)

We decided to discretize observed distribution function without a reference to a specific theoretical distribution, using the discretize R function available in the actuar package (see Klugman et al., 2010). This discretization allows us to build the matrix \(M_Z\) to which apply the two-dimensional FFT version. In this way, we have a new matrix \(FFT(M_Z)\) that acts as input of the random \(\tilde{k}\) probability generating function.

We need to define the distribution function of \(k\). The losses have been split by year, so we can report some descriptive statistics for frequency claims:

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Max</th>
<th>Q1</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>154</td>
<td>447</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>238</td>
<td>299</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 9: Statistics of frequency claims empirical distribution

We note 50% of frequencies are included between 237 and 380 claims and there is a light negative asymmetry. In addition, the variance is greater than mean value, so its possible suppose a Negative Binomial distribution for frequency claims; the corresponding probability generating function is defined by

\[
PGF(t) = \left( \frac{1-p}{1-pt} \right)^m
\]  

(22)

We have estimated its parameters \((m = 5\) and \(p = 0.82)\) and obtained the matrix \(PGF_k(FFT(M_z))\). As last stage we have applied the IFFT whose output is the matrix \(M_X\). Adding long counter-diagonals of \(M_X\) we can individuate discretized probability distribution of aggregate loss claims, having maintained the distinction between normal and large claims and, above all, preserving the dependence structure.
5 Final Results and Discussion

Now we are interested to estimate the $VaR_p$ using the previous models. According to the collective approach of risk theory, aggregate loss is the sum of a random number of random variables and so it requests convolution or simulative methods. We remember that among considered methodologies only FFT returns directly aggregate loss.

For illustrative purposes, considering the statistics of frequency in the Danish fire insurance data, we can assume the claim frequency constant and equals to $k = 300$.

A single simulation of aggregate loss can be achieved adding the losses of $k$ single claims and repeating the procedure 1,000,000 times, we obtain the aggregate loss distribution.

In the following table, we report the VaRs obtained using compound models Lognormal-Burr, Lognormal-GPD-Gamma, Gumbel Copula and FFT:

<table>
<thead>
<tr>
<th>Model</th>
<th>Claim frequency</th>
<th>VaR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lognormal-Burr</td>
<td>300</td>
<td>€ 205,727,356</td>
</tr>
<tr>
<td>Lognormal-GPD-Gamma</td>
<td>&quot;</td>
<td>€ 209,057,172</td>
</tr>
<tr>
<td>Gumbel Copula</td>
<td>&quot;</td>
<td>€ 649,006,035</td>
</tr>
<tr>
<td>FFT</td>
<td>Negative Binomial</td>
<td>€ 703,601,564</td>
</tr>
</tbody>
</table>

Table 10: Estimate of VaR at %99 level with different models
If we consider the independence assumption, aggregate loss distribution will return a VaR significantly smaller than those calculated relating dependence hypothesis.

All the previous approaches has advantages and disadvantages. With the first two composite models we can fit robustly each of the two underlying distribution of attritional and large claims, without a clear identification of the dependency structure. With the copula we can model dependency, but it is not easy to determine what is the right copula to use and that is the typical issue that the companies have under capital modelling purposes using copula approach. FFT allows to not simulate claim process and to not estimate a threshold, working directly on empirical data, but includes some implicit bias due to the discretization methods. Anyway, we realize that is fundamental take into account dependence between claims, regarding its shape and intensity, because VaR increase drastically respect to the independence case.

References


