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

Rocco Roberto Cerchiara[1] and Francesco Acri[2]

[1]Assistant Professor at Department of Economics, Statistics and Finance, University of Calabria, Arcavacata di Rende (Cosenza), Italy

[2]Actuarial Science PhD student at Department of Statistical Science, Sapienza University of, Rome, Italy

[1]rocco.cerchiara@unical.it

[2]francesco.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 functions, Fast Fourier Transform

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:

$$f(x) = \begin{cases} r_1 f_1^*(x) & k_0 < x < k_1 \\ \vdots \\ r_n f_n^*(x) & k_{n-1} < x < k_n \end{cases} \tag{1}$$

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

The Danish fire insurance 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 fits 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 Denuit (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 fire insurance 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):

$$f(x) = \begin{cases} r f_1^*(x) & -\infty < x \leq u \\ (1-r) f_2^*(x) & u < x < \infty \end{cases} \tag{2}$$

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

$$\begin{cases} f_1^*(x) = \frac{f_1(x)}{F_1(u)} & -\infty < x \leq u \\ f_2^*(x) = \frac{f_2(x)}{1-F_2(u)} & u < x < \infty \end{cases} \tag{3}$$

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:

$$\begin{cases} \lim\limits_{x \to u} f(x) = f(u) \\ \lim\limits_{x \to u^-} f'(x) = \lim\limits_{x \to u^+} f'(x) \end{cases} \tag{4}$$

From the first we obtain

$$r = \frac{f_2(u) F_1(u)}{f_2(u) F_1(u) + f_1(u)(1 - F_2(u))} \tag{5}$$

while from the second

$$r = \frac{f_2'(u) F_1(u)}{f_2'(u) F_1(u) + f_1'(u)(1 - F_2(u))} \tag{6}$$

We can define distribution function F of (2)

$$F(x) = \begin{cases} r \frac{F_1(x)}{F_1(u)}, & -\infty < x \leq u \\ r + (1-r) \frac{F_2(x) - F_2(u)}{1 - F_2(u)}, & u < x < \infty \end{cases} \tag{7}$$

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

$$F^{-1}(x) = \begin{cases} F_1^{-1}\Big(\frac{p}{r} F_1(u)\Big), & p \leq r \\ F_2^{-1}\Big(\frac{p - r + (1-p) F_2(u)}{1-r}\Big), & p > r \end{cases} \tag{8}$$

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:

$$L(x_1, \dots, x_n; \boldsymbol{\theta}) = r^m (1-r)^{n-m} \prod_{i=1}^{m} f_1^*(x_i) \prod_{j=m+1}^{n} f_2^*(x_j) \tag{9}$$

where n is sample dimension, $\boldsymbol{\theta}$ 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 Pigeon and Denuit, 2011). In particular, given random sample $X = (X_1, \dots, X_n)$, 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, $i = 1, \dots, n$. For this reason, $u_1, \dots, u_n$ 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 completly 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:

$$f(x) = (1-r) \int_0^x f_2(x) g(u) du + r \int_x^\infty \frac{1}{\Phi(\xi\sigma)} f_1(x) g(u) du \tag{10}$$

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, Ψ is the Standard Normal distribution function, ξ is the shape parameter of GPD and σ is Lognormal scale parameter.

2.2 Kumaraswamy Distribution and some generalization

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

$$K(x; \alpha, \beta) = 1 - (1 - x^\alpha)^\beta, x \in (0, 1) \tag{11}$$

the distribution proposed in Kumaraswamy (1980), where parameter α and β establish it's trend. If G is the distribution function of a random variable, then we can define a new distribution by

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

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):

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

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

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

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

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

u	φ	a	b	Kurtosis	Skewness
0	5	1	1	5.4	1.1
0	1	0.5	0.5	7.1	1.6
-5	3	2	3	3.6	0.5
1	10	5	0.7	6.4	1.4
0	15	1	0.4	7.6	1.7

Table 1: Kurthosis and Skewness of Kum-Gum distribution

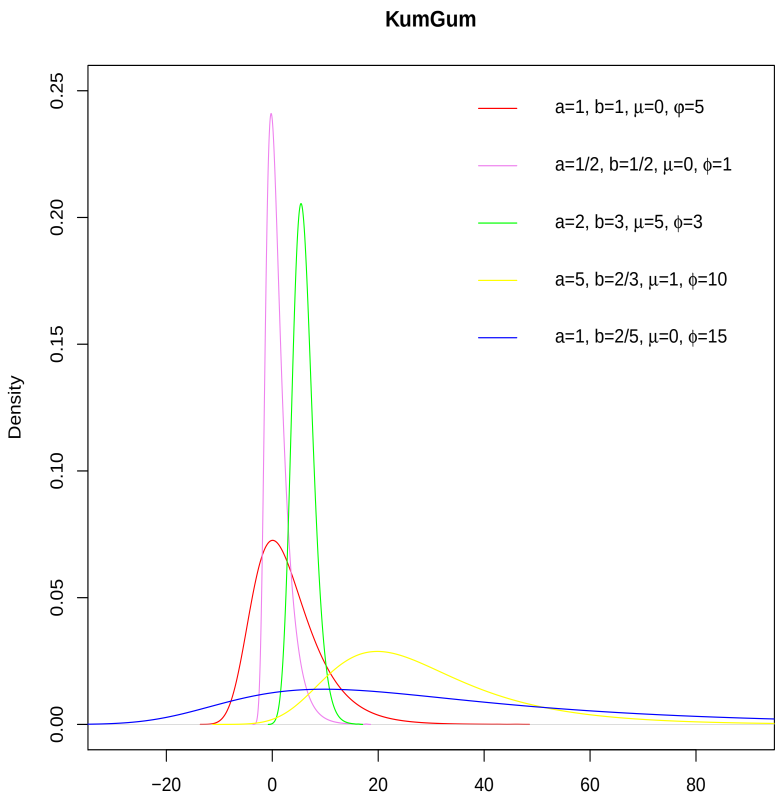

Figure 1: KumGum density functions.

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

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

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:

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

The corrisponding quantile function is

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

where $p \in (0,1)$. In the following figure we report Kum-Pareto density function varying the parameters:

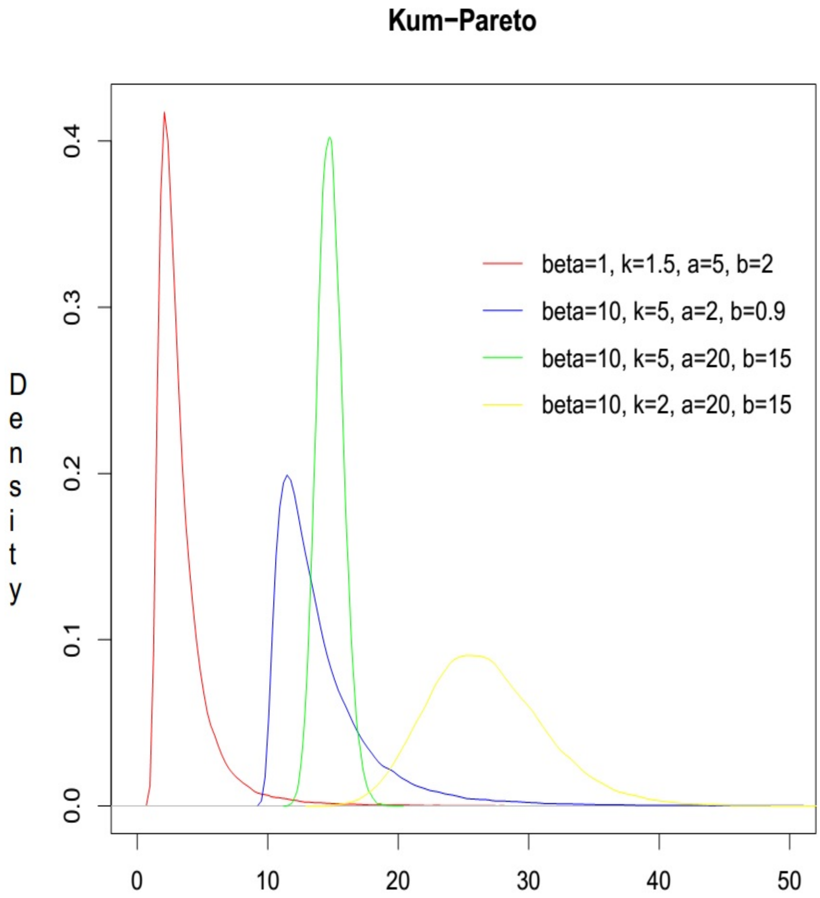

Figure 2: Kum-Pareto density functions.

3 Numerical example considering composite models

In this section we present some numerical results on the fitting of the Danish fire insurance 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), obtnaing $u = 1,022,125$€. Regarding the main statistics of Danish fire insurance data see Embrechts et al. (1997). We start with a compound model Lognormal-KumPareto, choosing $f_1 \sim Lognormal$ and $f_2 \sim Kum - Pareto$. From the following table we can compare some theoretical and empirical quantiles:

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

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 rigth tails of the distribution functions (empirical and theoretical):

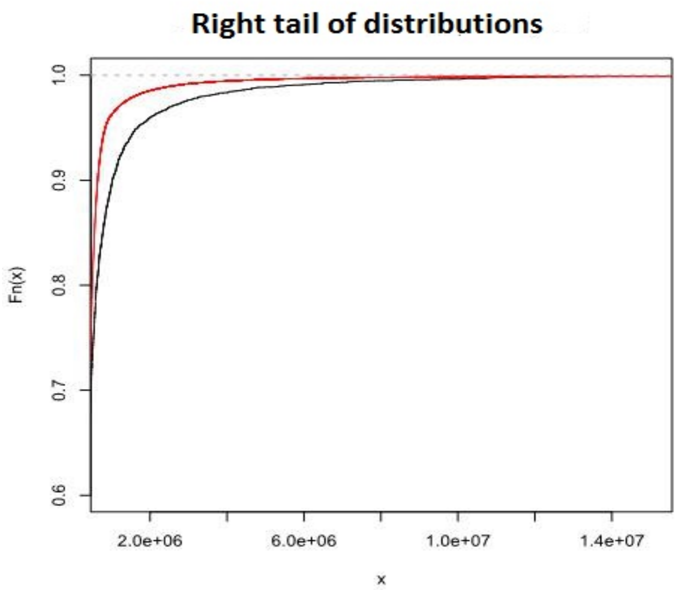

Figure 3: Rigth tails of Lognormal-KumPareto (red line) and empirical distribution (dark line) functions.

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 $f_1 \sim Lognormal$ and $f_2 \sim Burr$ as suggested in Nadarajah and Bakar (2014). The parameters are estimated using the CompLognonormal R package as shown in Nadarajah and Bakar (2014). From the following table we can compare some theoretical quantiles with empirical ones:

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

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

The model seems to be more feasible to catch the rigth tail of emprical distribution respect to the previous Lognormal-KumPareto, as we can see from the figure below:

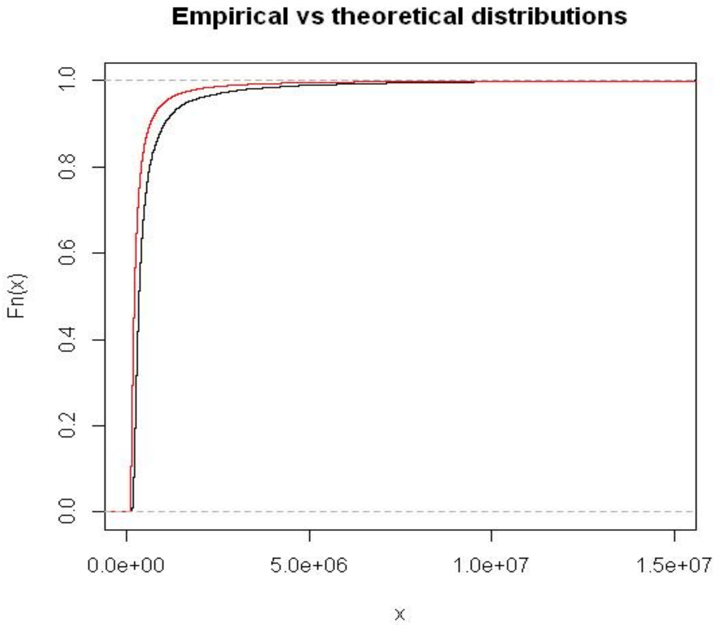

Figure 4: Lognormal-Burr and empirical distribution functions (red and dark lines).

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 $f_1 \sim Lognormal$ and $f_2 \sim GPD$ and then we generate pseudo-random numbers from quantile function (8). In the following we report the estimates of parameters and the QQ-plot:

	low extreme	best estimate	high extreme
μ_1	12.82	12.84	12.86
σ	0.59	0.61	0.62
σ_μ	1,113,916	1,115,267	1,116,617
ξ	0.33	0.45	0.56

Table 4: Estimated parameters of Lognormal-GPD

μ_1 and σ are the Lognormal parameters, while σ_μ and ξ are GPD parameters.

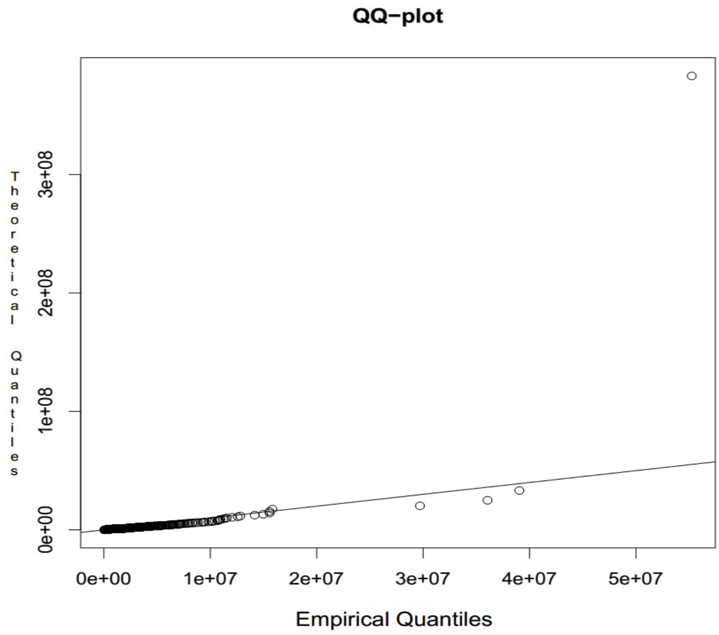

Figure 5: Observed-theoretical quantile plot for the Lognormal-GPD model.

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:

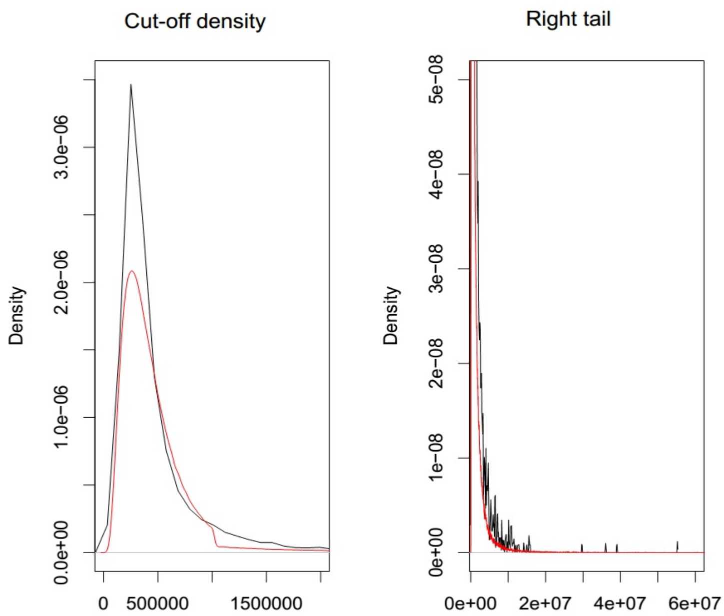

Figure 6: Left, comparison between cut-off density functions. Right, empirical and theoretical (red) right tail.

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:

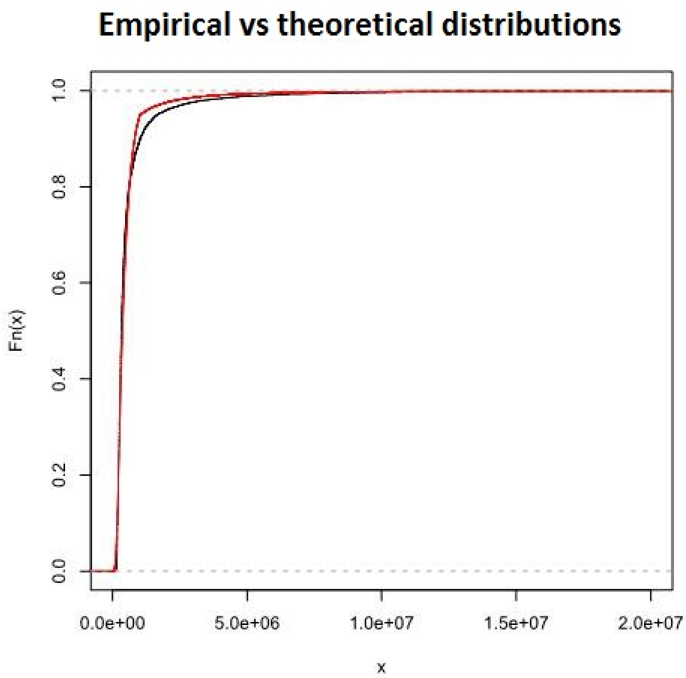

Figure 7: Lognormal-GPD (red) and empirical (dark) distribution function.

The corresponding Kolmogorov-Smirnov test has return 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) :

	low extreme	best estimate	high extreme
μ_1	12.78	12.79	12.81
σ	0.52	0.54	0.55
u (threshold)	629,416	630,768	632,121
σ_μ	1,113,915	1,115,266	1,116,616
ξ	0.22	0.29	0.37

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:

Levels	50%	75%	90%	95%	99%	99.5%
Empirical percentile	327,016	532,757	1,022,213	1,675,219	5,484,150	8,216,877
Theoretical percentile	360,574	517,996	1,103,309	2,077,792	5,266,116	7,149,253

Table 6: Comparison between empirical and Lognormal-GPD-Gamma quantiles

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

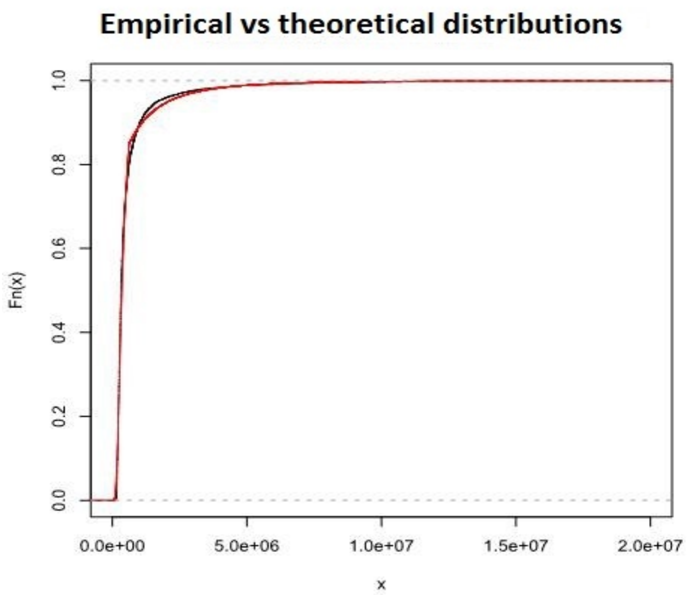

Figure 8: Lognormal-GPD-Gamma (red) versus empirical (dark) distribution functions.

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:

Index	KumPareto	Burr	GPD	GPD-Gamma
AIC	193,374	191,459	191,172	190,834
BIC	193,409	191,494	191,207	190,882

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 $\mathbf{R}$:

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

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:

Method	$\hat{\theta}$	Standard error
Maximum pseudo-likelihood	1.11	0.008
Canonical maximum pseudo-likelihood	1.11	0.008
Simulated maximum likelihood	1.11	-
Minimum distance	1.09	-
Moments based on Kendall's tau	1.13	-

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

We remind that Gumbel's parameter θ assumes values in $[1, \inf)$ 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 Indipendent Copula. We can say that because we have verified, using bootstrap procedures, θ 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:

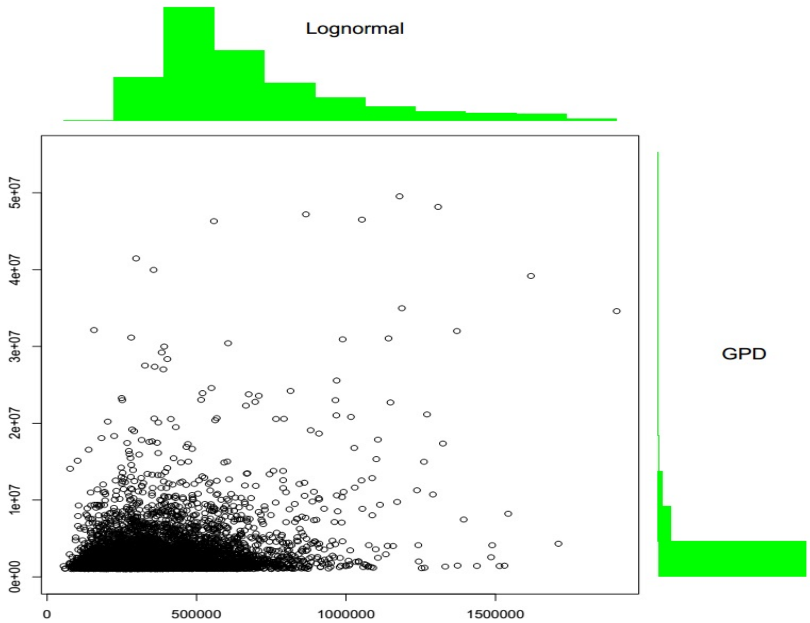

Figure 9: Lognormal (top) and GPD (right) marginal histograms and Gumbel Copula simulated values plot.

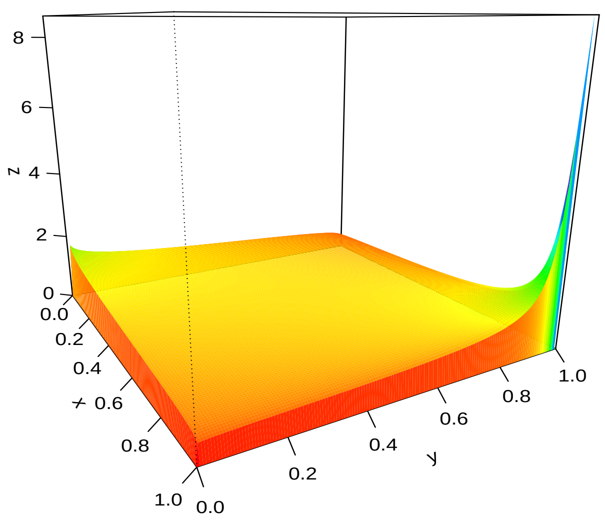

Figure 10: Density function of estimated Gumbel Copula. Attritional claims losses on X-axis, large claims losses on Y-axis.

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 difference 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

$$PGF_{\tilde{k}}(t) = \sum_{j=0}^{\infty} t^j Pr(\tilde{k} = j) = E[t^k] \tag{18}$$

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

$$(x_0, x_1, \ldots, x_{n-1}) = IFFT(PGF(FFT(\mathbf{z}))) \tag{19}$$

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

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

$$\phi(z) = \int_{-\infty}^{+\infty} f(x) \exp(izx) dx \tag{20}$$

The characteristic function creates an association between a probability density 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 its 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

$$\mathbf{M_z} = \begin{pmatrix} 0.5 & 0 & 0 \\ 0.2 & 0.25 & 0 \\ 0 & 0.05 & 0 \end{pmatrix}$$

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

$$\mathbf{M_x} = IFFT(PGF(FFT(\mathbf{M_z}))) \tag{21}$$

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:

Min	**Max**	**Q1**	**Mean**
154	447	238	299
Median	**Q3**	**Variance**	**Skewness**
310	381	8,482	-0.12

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

$$PGF(t) = \Big(\frac{1-p}{1-pt}\Big)^m \tag{22}$$

We have estimated its parameters ($m = 5$ and $p = 0.82$) and obtained the matrix $PGF_{\tilde{k}}(FFT(\mathbf{M_z}))$. 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.

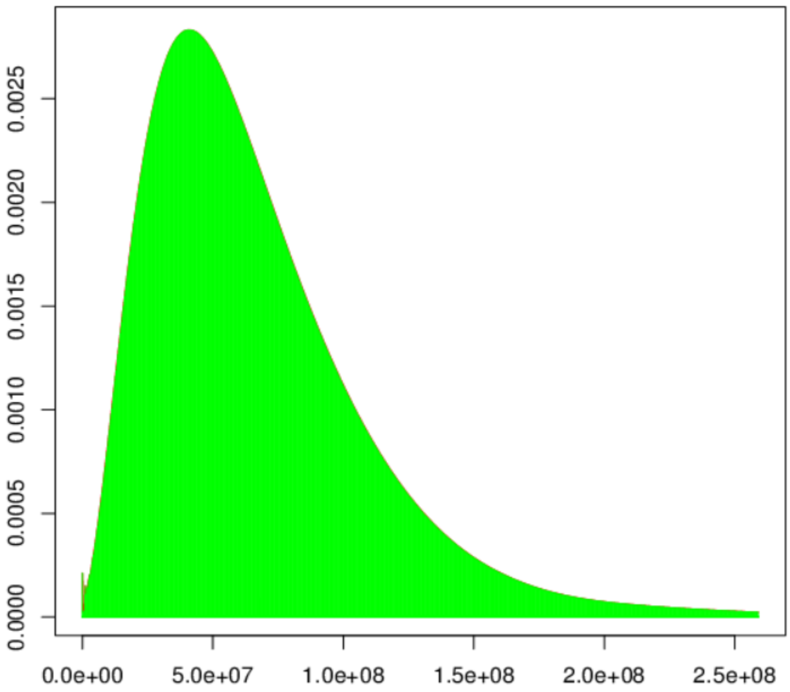

Figure 11: Aggregate loss density function obtained by FFT procedure.

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; repeating the procedure many times (1,000,000 in our paper), 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:

Model	Claim frequency	VaR
Lognormal-Burr	300	€ 205,727,356
Lognormal-GPD-Gamma	"	€ 209,057,172
Gumbel Copula	"	€ 649,006,035
FFT	Negative Binomial	€ 703,601,564

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

[1] Ausin, M. C., Wiper, M. P. & Lillo, R. E. (2009). Bayesian estimation of finite time ruin probabilities. Applied Stochastic Models in Business and Industry 25, 787-805.

[2] Burnecki, K. & Weron, R. (2004). Modeling the risk process in the XploRe computing environment. *Lecture Notes in Computer Science.* **3039**, 868-875.

[3] Carreau, J. & Bengio, Y. (2009). A hybrid pareto model for asymmetric fat-tailed distribution. Extremes 12. 53-76.

[4] Charpentier, A. & Oulidi, A. (2010). Beta kernel quantile estimators of heavy-tailed loss distributions, *Statistics and computing.* **20**, 35-55.

[5] Cooray, K. & Ananda, M. M. A. (2005). Modeling actuarial data with a composite lognormal-Pareto model. *Scandinavian Actuarial Journal*, **5**, 321-334.

[6] Cordeiro, G. M. & Nadarajah, S. & Ortega, E.M.M. (2012). The Kumaraswamy Gumbel distribution. *Statistical Methods & Applications*, **21**, 139-168.

[7] Drees, H. & Müller, P. (2008). Fitting and validation of a bivariate model for large claims, *Insurance: Mathematics and Economics*, **42**, 638-650.

[8] Embrechts, P. & Klüppelberg, C. & Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*, Springer.

[9] Esmaeili, H. & Klüppelberg, C. (2010). Parameter estimation of a bivariate compound Poisson process. *Insurance: Mathematics and Economics*, **47(2)**, 224-233.

[10] Guillotte, S. & Perron, F. & Segers, J. (2011). Non-parametric Bayesian inference on bivariate extremes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **73**, 377-406.

[11] Klugman, S. & Panjer, H.H. & Wilmot, G.E. (2010). *Loss Models: From Data to Decisions*, John Wiley & Sons.

[12] Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes. Journal of Hydrology, 46 (1-2): 79â€“88

[13] McNeil, A. J. & Frey, R. & Embrechts, P. (2005). *Quantitative Risk Management. Concepts, Techniques and Tools*, Princeton University Press.

[14] Nadarajah, S. & Bakar, A.A.S. (2014). *New composite models for the Danish-fire data*, *Scandinavian Actuarial Journal*, **2**, 180-187.

[15] Pigeon, M. & Denuit, M. (2011). Composite Lognormal-Pareto model with random threshold. *Scandinavian Actuarial Journal*, **3**, 177-192.

[16] Scollnik, D. P. M. (2007). On composite lognormal-Pareto models. *Scandinavian Actuarial Journal*, **1**, 20-33.

[17] Teodorescu, S. & Vernic, R. (2009). Some composite Exponential-Pareto models for actuarial prediction. *Rom. J. Econ. Forecast.*, **12**, 82-100.

[18] Teodorescu, S. & Vernic, R. (2013). On Composite Pareto Models. *MATH. REPORTS*, **15(65)**, 1, 11-29.