UNIVERSITY OF BUCHAREST
FACULTY OF MATHEMATICS AND COMPUTER SCIENCE
DOCTORAL SCHOOL OF MATHEMATICS

Stochastic models in actuarial science. Theoretical contributions and applications

Summary

Scientific supervisor:
PhD Prof. Preda Vasile

PhD Sudent:
Robe-Voinea
Elena-Graţiela

BUCUREŞTI
2016

Contents

Chapter 1

Introduction

The main objective of the research realized in the doctoral school program was the study of some extension of stochastic models concerning aggregate claims and their implementation through several practical applications.
We extend such models from the univariate case to the multivariate case, models in which we include some dependencies in order to help the actuaries to make a more realistic analysis of insurance scenarios.
In my opinion, the study of such models is a starting point in the theory of the aggregate claims distributions and opens a variety of ways of research that other young future PhD students could follow.
The operational objectives of my research were:

- The extension to a multivariate settings of the bivariate collective model introduced by Jin and Ren [14], to study aggregate claims when different types of claims simultaneously affect an insurance portfolio. The simpler multivatiate collective models were studied by Sundt [29].

- The description of several techniques used to evaluate the compound distribution such as: convolution, recursions, simulation, calculation with approximate distribution and inversion methods, which also includes the FFT, and moreover the study of how these methods interact with our new collective models in order to obtain the distribution of the aggregate amount of claims occurred in an in insurance portfolio within a given period of time. For details on these methods see Klugmann [17], while for the FFT method see Bühlmann [3], Grubel and Hermesmeier [8] and Embrechts et al. [7].

- A comparison between these techniques mentioned above in order to find the optimal way to be closer to the reality;

- We also focused on the so called alternative methods in order to simplify calculations and reduce the computing time, here we detailed a lot of aspects such as: a variety of errors that these methods imply but also a strongly comparison between simulation and recursive methods. For details on such errors see Klugman [17], Grubel and Hermesmeier [8], [9], and Sundt and Vernic [33].

- A study case using the Monte Carlo method in order to demonstrate the utility and the application of this method in the field of engineering, more precisely in the shipbuilding area. Fore more details regarding risk in shipbuilding see Dorp and Duffey [6] and Kolic and Calic [18], and fore mode generic details see also Storch and Lin, [27], and Winston [37] and Dlugokecki et al. [5].

The original results of this thesis were accepted and published in 4 international ISI journals such as:

1. **Elena-Gratiela Robe-Voinea**, Raluca Vernic, On a multivariate aggregate claims model with multivariate Poisson counting distribution, *Proceedings of the Romanian Academy, Series A*, ISSN: 1454-9069,cotat CNCSIS **A**, ID =789, and ISI indexed, **2015-IF=1.658**, accepted

2. **Elena-Gratiela Robe-Voinea**, Raluca Vernic, Another appproach to the evaluation of a certain multivariate compound distribution *Analele Universitatii "Ovidius" Constanta, seria Matematica*, ISSN 1224-1784, E-ISSN 1844-0835. (cotat CNCSIS **A**, ID =31, ISI indexed), **2015-IF= 0.383**, accepted.

3. **Elena-Gratiela Robe-Voinea**, Raluca Vernic, Fast Fourier Transform for multivariate aggregate claims, *Computational and Applied Mathematics, Springer International Publishing*, 2016, DOI 10.1007/s40314-016-0336-6, Print ISSN 0101-8205, Online ISSN 1807-0302. (ISI indexed), **2016-IF=0.802**.

4. **Elena-Gratiela Robe-Voinea**, Raluca Vernic, On the recursive evaluation of a certain multivariate compound distribution, *Acta Mathematicae Applicatae Sinica, English Series, Springer International Publishing*, 2016,. (ISI indexed), **2016-IF=0.250**, accepted.

Furthermore during the doctoral school there were attendace with presentantion to the following national and international conferences:

1. **Elena-Gratiela Robe-Voinea**, Raluca Vernic, Fast Fourier Transforms for Bivariate aggregate losses. A Matlab application, *The 16th Conference of Romanian Society of Statistics and Probabilities, SPSR* , April 26, 2013, The Bucharest University of Economic Studies, Bucharest, Romania.

2. **Elena-Gratiela Robe-Voinea**, Raluca Vernic, Fast Fourier transform for Multivariate aggregate losses, *The 17th Conference of Romanian Society of Statistics and Probabilities, SPSR* , April 25, 2014, The University of Bucharest, Romania.

3. **Elena-Gratiela Robe-Voinea**, Exponential tilting for computing multivariate compound distributions using the Fast Fourier Transform, *22nd Conference on Applied and Industrial Mathematics - CAIM 2014* , September 18-22, 2014, University of Bacău, Romania, Abstracts, pp. 38.

4. **Elena-Gratiela Robe-Voinea**, Raluca Vernic, What if different types of lcaims simultaneously affect an insurance portofolio?, *The 18th Conference of Romanian Society of Statistics and Probabilities,SPSR*, May 8, 2015, University of Bucharest, Romania.

5. **Elena-Gratiela Robe-Voinea**, Raluca Vernic, On the recursive evaluation of a certain multivariate compound distribution, *The 8th Congress of Romanian Mathematicians*, June 28 - July 1, 2015, "Alexandru Ioan Cuza" University of Iasi, Romania, Abstracts, pp. 120.

6. **Elena-Gratiela Robe-Voinea**, On the recursive evaluation of a certain multivariate compund distribution , *Scientific PhD Students Session of the PhD School of Mathematics*, June 22, 2015, University of Bucharest, Romania.

7. **Elena-Gratiela Robe-Voinea**, Raluca Vernic, A recursive procedure for a compound risk model with compound-type severity distributions, *The 12th Balkan Conference on Operational Research BALCOR*, September 9-13, 2015, "Romanian Naval Academy", Constanta, Romania.

8. **Elena-Gratiela Robe-Voinea**, Raluca Vernic, Risk analysis based on the Monte Carlo method for a ship design project, *The 19th Conference of Romanian Society of Statistics and Probabilities, SPSR*, May 27, 2016, Technical University of Civil engineering, Bucharest, Romania.

9. **Elena-Gratiela Robe-Voinea**, *Scientific PhD Students Session of the PhD School of Mathematics*, June 21, 2016, University of Bucharest, Romania.

10. **Elena-Gratiela Robe-Voinea**, Raluca Vernic, Multivariate aggregate claims evaluation using the Fast Fourier Transform, *13éme Colloque Franco-Roumain de Mathématiques Appliquées]*, August 25-29, 2016, "Alexandru Ioan Cuza" University of Iasi, Romania.

Chapter 2

Preliminary notions

2.1 Collective models

2.1.1 Univariate case

The collective model in the univariate case can be written as

$$S = \sum_{j=0}^{N} U_j, \tag{2.1}$$

where S represents the random variable (r.v.) aggregate claims, N the r.v. number of claims occurring during the given time period, $U_0 = 0$ and $(U_j)_{j\geq 1}$ are independent and identically distributed (i.i.d.) claim sizes positive r.v.s, which are also independent of N see [11].
The distribution of S is called compound, while the distribution of N is called counting distribution.

2.1.2 Bivariate case

The collective model in the bivariate case, can be written as

$$(S_1, S_2) = \left(\sum_{l=1}^{N_1} U_l, \sum_{j=1}^{N_2} V_j \right), \tag{2.2}$$

where S_1 represents the aggregate losses of type I, S_2 the aggregate losses of type II, U_l the cost of type I losses, i.i.d., V_j the cost of type II losses, i.i.d., N_1 the number of type I losses, N_2 the number of type II losses. Here N_1 and U_l are independent while N_2, V_j are also independent.
Recently, [14] introduced a bivariate collective model to study aggregate claims in the case when different types of claims simultaneously affect an insurance

portfolio (e.g., floods, storms or earthquakes).
Their model is

$$(S_1, S_2) = \left(\sum_{i=1}^{N_1} U_i + \sum_{k=1}^{N_3} L_k, \sum_{j=1}^{N_2} V_j + \sum_{k=1}^{N_3} Q_k \right), \tag{2.3}$$

where N_1 denotes the number of accidents that cause only type one claims, N_2 denotes the number of accidents that cause only type two claims and N_3 denotes the number of accidents that cause both types of claims. The claim number vector (N_1, N_2, N_3) has probability function $p(n_1, n_2, n_3) = P(N_1 = n_1, N_2 = n_2, N_3 = n_3)$; $(U_i)_{i \leq 1}$ and $(V_j)_{j \leq 1}$ are mutually independent and independent of the claim numbers (N_1, N_2, N_3) and claim sizes $(L_k, Q_k)_{k \leq 1}$, in the same manner the claim sizes vectors $(L_k, Q_k)_{k \leq 1}$ are mutually independent, identically distributed and independent of claim numbers (N_1, N_2, N_3) and claim sizes U_i and V_j.

2.1.3 Multivariate case

The multivariate form for the collective model can be written as

$$(S_1, ..., S_m) = \left(\sum_{i=1}^{N_1} U_{1i}, ..., \sum_{i=1}^{N_m} U_{mi} \right) \tag{2.4}$$

where N_k denotes the number of claims of type k and $(U_{ki})_{i \leq 1}$ their corresponding costs, which are i.i.d. and independent of the number of claims. [14] obtained recursions for the bivariate form (2.3) under three different assumptions related to the dependency structure of the claim numbers; the resulting models were named A, B and C, corresponding to the similar ones from [11]. They also used FFT as an alternative method.

Remark 2.1 *We recall that a distribution belongs to the $R_1(a, b)$ class if its probability function (p.f.) satisfies the recursion*

$$\Pr(N = n) = \left(a + \frac{b}{n} \right) \Pr(N = n - 1), \forall n \geq 1,$$

for some constants $a, b \in \mathbb{R}$ (for details on the R_k classes see, e.g., [28] or [33]).

There are several techniques used to evaluate the compound distributions such as: convolutions, recursions, simulation, calculation with approximate distribution and inversion methods, which also includes the FFT. The recursive and the FFT method will be studied in detail in the following together with the manner in which they interact with our new collective models in order to obtain the distribution of the aggregate amount of claims occurred in an in insurance portfolio within a given period of time.

Chapter 3

Recursions for compound multidimensional models

3.1 Introduction

We consider the following multivariate aggregate claims model

$$(S_1, ..., S_m) = \left(\sum_{i=0}^{N_1} U_{1l} + \sum_{k=0}^{N_0} L_{1k}, ..., \sum_{l=0}^{N_m} U_{ml} + \sum_{k=0}^{N_0} L_{mk} \right), \tag{3.1}$$

where $m \geq 2$ is the number of different types of claims affecting the portfolio, S_k denotes the aggregate claims of type k, N_k the number of claims of only type k, N_0 the number of common claims (e.g., accidents that causes all m types of different claims). Each set of claim sizes $(U_{jl})_l \geq 1$ are positive, independent and identically distributed (i.i.d.) as the generic random variable (r.v.) $U_j, 1 \leq j \leq m$, independent of the claim numbers and of the other claim sizes, including $(L_{1k}, ..., L_{mk})$. The random vectors $(L_{1k}, ..., L_{mk})_{k\leq 1}$ are non-negative i.i.d. as the generic $(L_1, ..., L_m)$, and independent of the claim numbers. Clearly, $U_{j0} = L_{j0} = 0, \forall j$. In the following, by a bold faced letter we denote a vector, i.e., $\mathbf{X} = (X_1, ..., X_m)$ or $\mathbf{x} = (x_1, ..., x_m)$. We shall work with discrete r.v.s and if the claim sizes distributions are continuous, they should be discretized using, e.g., the rounding method, see [17]. If f is a probability function (p.f.), we denote by f^{*n} its n-fold convolution corresponding to the distribution of the sum of n i.i.d. r.v.s having p.f. f, note that$f^{*1} = f$ and, by convention, $f^{*0}(x) = \begin{cases} 1 & x = 0 \\ 0 & x \neq 0 \end{cases}$.
Let $f_{\mathbf{S}}$ denote the p.f. of $\mathbf{S}$, f_j the p.f. of U_j,$j = \overline{1, m}$, f_0 the p.f. of $\mathbf{L}$, and p the p.f. of $\mathbf{N} = (N_0, ..., N_m)$.
Then from [17], we easily obtain

$$f_{\mathbf{S}}(\mathbf{x}) = \sum_{n_0\geq 0,...,n_m\geq 0} p(n_0, ..., n_m) \sum_{k=0}^{x} \prod_{j=1}^{m} f_j^{*n_j}(k_j) f_0^{*n_0}(\mathbf{x} - \mathbf{k}), \mathbf{x} \geq \mathbf{0}, \tag{3.2}$$

where $\mathbf{0} = (0, ..., 0)$, while the inequality $\mathbf{x} \geq \mathbf{0}$ and the difference **x-k** are componentwise. Note that due to the convolutions in (3.2), $f_{\mathbf{S}}$ could be difficult to evaluate and time-consuming; therefore, alternative methods have been developed, from which recursions play an important role in actuarial mathematics. We recall the fact that the distribution of **N** is also called counting distribution, while the distribution of **S** is called compound. Let us denote the p.f. of a discrete random vector **X** by $f_{\mathbf{X}}$ and its probability generating function (p.g.f.) by $g_{\mathbf{X}}$; we recall that

$$g_{\mathbf{X}}(\mathbf{t}) \stackrel{def}{=} E\left[\prod_{j=1}^{m} t_j^{X_j}\right].$$

Moreover, as a property of the p.g.f., it holds that

$$g_{\mathbf{X}}(\mathbf{t}) = \sum_{x \geq 0} f_{\mathbf{X}}(\mathbf{x}) \prod_{i=1}^{m} t_i^{x_i} \tag{3.3}$$

and clearly, $g_{\mathbf{X}}(\mathbf{0}) = f_{\mathbf{X}}(\mathbf{0})$.

Proposition 3.1 *Under the assumptions of model (3.1), the p.g.f. of* $\boldsymbol{S}$ *is given by*

$$g_{\boldsymbol{S}}(\mathbf{t}) = g_{\boldsymbol{N}}(g_{\boldsymbol{L}}(\mathbf{t}), g_{U_1}(t_1), ..., g_{U_m}(t_m)). \tag{3.4}$$

3.2 Recursive evaluation: first case, corresponding to model B

This section is based on paper [21].
In this next section, we shall present a recursion to evaluate the distribution for model 3.1 when the number of claims **N** follows a multivariate Poisson distribution. In the bivariate setting, when $m = 2$, a recursion for this case has been recently presented in [14]; our recursion extends this one to a general m. In the simpler univariate case ($m = 1$), the history of recursions involving Poisson counting distributions starts in insurance with [20] and [35], and continues with more complex recursions for compound mixed Poisson distributions discussed in [36], [11] among others, or, in the multivariate case, in [32], etc. In particular, as we mentioned from the beginnig, if **N** follows the multivaritate Poisson distribution $MPo_{m+1}(\lambda, \lambda_0, ..., \lambda_m)$ with $\lambda > 0, \lambda_j > 0, \forall j$, then, from [16] we have that

$$g_{\mathbf{N}}(\mathbf{t}) = \exp\left\{\lambda\left(\prod_{j=0}^{m} t_j - 1\right) + \sum_{j=0}^{m} \lambda_j(t_j - 1)\right\}. \tag{3.5}$$

Proposition 3.2 *Under the assumptions of model (3.1), if* $\boldsymbol{N}$ *follows the multivariate Poisson distribution* $MPo_{m+1}(\lambda, \lambda_0, ..., \lambda_m)$*, then the p.f. of* $\boldsymbol{S}$ *satisfies the recursion*

$$\begin{aligned} f_S(\mathbf{x}) &= \frac{\lambda_k}{x_k} \sum_{z_k=0}^{x_k} z_k f_k(z_k) f_S(x_1, ..., x_{k-1}, x_k - z_k, x_{k+1}, ..., x_m) + \\ & + \frac{\lambda_0}{x_k} \sum_{z=0}^{\mathbf{x}} z_k f_0(\mathbf{z}) f_s(\mathbf{x}-\mathbf{z}) + \frac{\lambda}{x_k} \sum_{\mathbf{v}=\mathbf{0}}^{\mathbf{x}} \left(\sum_{\mathbf{u}=\mathbf{0}}^{\mathbf{v}} v_k f_i(u_i) f_0(\mathbf{v}-\mathbf{u}) \right), \end{aligned}$$

$x_k \geq 1, 1 \leq k \leq m$,*with starting value*

$$f_{\mathbf{S}}(\mathbf{0}) = \exp\left\{ \lambda \left(f_0(\mathbf{0}) \prod_{j=1}^{m} f_j(0) - 1 \right) + \sum_{j=1}^{m} \lambda_j (f_j(0) - 1) + \lambda_0 (f_0(\mathbf{0}) - 1) \right\}. \tag{3.6}$$

3.3 Recursive evaluation: second case corresponding to model A

This section is based on [25].
In this case, we start by extending the recursion of model A from [14] in a multivariate setting under the assumption that the distribution of the total number of claims N belongs to the $R_1(a, b)$ class, while, conditionally on it, the distribution of $\mathbf{N}$ is multinomial.
Therefore, we derive the recursive formula corresponding to the multivariate model A, and provide a numerical example in the trivariate case $m = 3$ making some remarks in what regards the computations order.
According to the model A in Jin and Ren [14], we consider the following supplementary assumptions on model (3.1):

A1 The first one is related to the total number of claims $N = N_0 + N_1 + ... + N_m$, whose probability function (p.f.) is assumed to satisfy the Panjer-type recursion

$$\Pr(N = n) = \left(a + \frac{b}{n} \right) \Pr(N = n - 1), \forall n \geq 1,$$

for some constants $a, b \in \mathbb{R}$ (for details on Panjer's class, see [19] or [33]);

A2 The second one concerns the conditional distribution of $\mathbf{N}$ given $N = n$, which is assumed to be multinomial $Mnom(n; p_1, ..., p_m)$ with the parameters $n \in \mathbb{N}$ and $p_1, ..., p_m \in (0, 1)$ such that $p_0 := 1 - \sum_{i=1}^{m} p_i \in (0, 1)$. We recall that (see, e.g., [16]), with $n = \sum_{i=0}^{m} n_i$,

$$\Pr(N_0 = n_0, N_1 = n_1, ..., N_m = n_m | N = n) = \frac{n!}{\prod_{i=0}^{m} n_i!} \prod_{i=0}^{m} p_i^{n_i}.$$

Moreover, denoting by $\mathbb{E}$ the expected value operator, based on the p.g.f. formula of this multinomial distribution (see, e.g., [16]), the p.g.f. of $\mathbf{N}$ becomes

$$g_{\mathbf{N}}(\mathbf{n}) = \mathbb{E}\left[\mathbb{E}\left[\prod_{j=0}^{m} n_j^{N_j} \,\middle|\, N\right]\right] = \mathbb{E}\left[\left(\sum_{j=0}^{m} p_j n_j\right)^{N}\right] = g_N\left(\sum_{j=0}^{m} p_j n_j\right).$$

From Proposition 2.1 in [21], we have that for the general model (3.1), the p.g.f. of $\mathbf{S}$ is given by

$$g_{\mathbf{S}}(\mathbf{t}) = g_{\mathbf{N}}(g_{\mathbf{L}}(\mathbf{t}), g_{U_1}(t_1), ..., g_{U_m}(t_m)),$$

and inserting the above formula of $g_{\mathbf{N}}$ yields that for model A,

$$g_{\mathbf{S}}(\mathbf{t}) = g_N\left(\sum_{j=1}^{m} p_j g_{Uj}(t_j) + p_0 g_{\mathbf{L}}(\mathbf{t})\right). \tag{3.7}$$

The following recursive formulas were proved in the thesis in two alternative ways: first based on the properties of the p.g.f. and, alternatively, based on some recursions presented in [33]. Next, we redenote $f_{\mathbf{L}}$ by f_0.

Proposition 3.3 *Under the assumptions (A1-A2) of model (3.1), the following starting value and recursive formulas hold:*

$$f_{\mathbf{s}}(\mathbf{0}) = g_{\mathbf{S}}(\mathbf{0}) = g_N\left(\sum_{j=1}^{m} p_j f_j(0) + p_0 f_{\mathbf{L}}(\mathbf{0})\right);$$

$$\begin{aligned} f_{\mathbf{S}}(\mathbf{x}) &= K\Bigg\{ a\sum_{\substack{j=1\\ j\neq k}}^{m} p_j \sum_{y_j=1}^{x_j} f_j(y_j) f_{\mathbf{S}}(x_1, ..., x_{j-1}, x_j - y_j, x_{j+1}, ..., x_m) \\ &\quad + p_k \sum_{y_k=1}^{x_k}\left(a + b\frac{y_k}{x_k}\right) f_k(y_k) f_{\mathbf{S}}(x_1, ..., x_{k-1}, x_k - y_k, x_{k+1}, ..., x_m) \\ &\quad + p_0 \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}}\left(a + b\frac{y_k}{x_k}\right) f_{\mathbf{L}}(\mathbf{y}) f_{\mathbf{S}}(\mathbf{x}-\mathbf{y})\Bigg\}, \end{aligned} \tag{3.8}$$

$x_k \geq 1, x_j \geq 0, \forall j \neq k;$

$$\begin{aligned} f_{\mathbf{S}}(\mathbf{x}) &= K\Bigg\{ \sum_{j=1}^{m} p_j \sum_{y_j=1}^{x_j}\left(a + b\frac{y_j}{x_+}\right) f_j(y_j) f_{\mathbf{S}}(x_1, ..., x_{j-1}, x_j - y_j, x_{j+1}, ..., x_m) \\ &\quad + p_0 \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}}\left(a + b\frac{y_+}{x_+}\right) f_{\mathbf{L}}(\mathbf{y}) f_{\mathbf{S}}(\mathbf{x}-\mathbf{y})\Bigg\}, \ \mathbf{x} > \mathbf{0}, \end{aligned} \tag{3.9}$$

where $K = \left[1 - a\left(\sum_{j=1}^{m} p_j f_j(0) + p_0 f_{\mathbf{L}}(\mathbf{0})\right)\right]^{-1}$ *and* $x_+ = \sum_{i=1}^{m} x_i$.

Chapter 4

Alternative methods

In the following, we shall need the characteristic function of model (3.1) presented in next proposition.

Proposition 4.1 *Under the assumptions of model (3.1), it holds that the characteristic function of* $\mathbf{S}$ *is given by*

$$\varphi_{\mathbf{S}}(\mathbf{t}) = g_{\mathbf{N}}(\varphi_{\mathbf{L}}(\mathbf{t}), \varphi_{U_1}(t_1), ..., \varphi_{U_m}(t_m)). \tag{4.1}$$

4.1 Multidimensional discrete Fourier transforms and FFT algorithm

Let $f(\mathbf{x})$ be an m-variate function defined on the integer values $x_j = 0, 1, ..., r_j - 1, 1 \leq j \leq m$. Then its discrete Fourier transform (DFT) $\tilde{f}$ can defined by (definition used in Matlab)

$$\tilde{f}(\mathbf{c}) = \sum_{x_1=0}^{r_1-1} ... \sum_{x_m=0}^{r_m-1} f(\mathbf{x}) \exp\left\{-2\pi i \sum_{j=1}^{m} \frac{x_j c_j}{r_j}\right\}, \ c_j = 0, ..., r_j - 1, \ 1 \leq j \leq m.$$

Its inverse mapping is given by

$$f(\mathbf{x}) = \frac{1}{\prod_{j=1}^{m} r_j} \sum_{c_1=0}^{r_1-1} ... \sum_{c_m=0}^{r_m-1} \tilde{f}(\mathbf{c}) \exp\left\{2\pi i \sum_{j=1}^{m} \frac{x_j c_j}{r_j}\right\}, \ x_j = 0, ..., r_j - 1, \ 1 \leq j \leq m.$$

For model (3.1), the following algorithm based on the FFT and its inverse (IFFT) can be used to obtain an approximate distribution of $\mathbf{S}$.
Algorithm 1
Step 1. Set the truncation points for the r.v.s claim sizes U_j at $r_j, 1 \leq j \leq m$, and for $\mathbf{L}$ at $(r_1, ..., r_m)$. The truncated claim size distributions result as $\mathbf{f}_j =$

$\{f_j(0), f_j(1), ..., f_j(r_j-1)\}$ for $U_j, 1 \leq j \leq m$, and $\mathbf{f}_0 = [f_0(j_1, ..., j_m)]_{j_1,...,j_m}$ for $\mathbf{L}$, where $0 \leq j_l \leq r_l - 1, 1 \leq l \leq m$. If necessary, the resulting vectors $\mathbf{f}_j$ or the table $\mathbf{f}_0$ can be padded with zeros to force the r_js to be powers of two.
Step 2. Apply the one-dimensional FFT to $\mathbf{f}_j$ yielding the vector $\tilde{\mathbf{f}}_j, 1 \leq j \leq m$; then apply the multidimensional FFT to $\mathbf{f}_0$, yielding the multidimensional table $\tilde{\mathbf{f}}_0$.
Step 3. Use formula (4.1) to obtain the discrete characteristic function

$$\tilde{\varphi}_{\mathbf{S}}(\mathbf{j}) = g_{\mathbf{N}}(\tilde{\mathbf{f}}_0(\mathbf{j}), \tilde{\mathbf{f}}_1(j_1), ..., \tilde{\mathbf{f}}_m(j_m)), 0 \leq j_l \leq r_l - 1, 1 \leq l \leq m.$$

Step 4. Apply the multidimensional IFFT to $\tilde{\varphi}_{\mathbf{S}}$ to obtain the p.f. of $\mathbf{S}$.

Remark 4.2 *To find optimal r_js, one can gradually increase them (e.g., 64, 128, 256 etc.) until the differences between the solutions obtained for the current values of the r_js and the previous ones are no more significant. More details on errors are given in the next section.*

Back to the above Algorithm 1, it essentially generates two types of errors: discretization and aliasing errors.

4.2 Types of errors. Exponential tilting

The errors generated by the use of the DFT (and, in particular, of the FFT) have been investigated mainly in connection with harmonic analysis applications (i.e., images and signal processing) see, e.g., [13] and [1]. In the insurance field, even if the calculation of aggregate claims distributions using the Fourier method starts in 1983 with [10], only later on, [8] and [9] conducted a thorough study of the related errors and even proposed an improved FFT procedure based on an exponential change of measure. Concerning the same problem, [26] noted that “such typical errors can be crucial for the final result, especially when working with heavy-tailed distributions”.

4.2.1 Discretization (arithmetization) errors

The source of this types of errors is the choice of spans h_i, and the solutions to reduce it are the following:
- evaluation of upper and lower bounds (can be too pessimistic);
- successively reducing the spans until the improvment in the computed distribution is small enough.

4.2.2 Aliasing errors

This is a DFT specific error due to truncation and consists in placing below the truncation point the compound mass which lies beyond this point (*wrap-around* effect). Based on convolutions, we obtained the following upper bound for the AE corresponding to our model:

$$AE\left(\mathbf{S}\right) \leq 1 - F_{\mathbf{S}}\left(\mathbf{r} - \mathbf{1}\right).$$

To reduce the AE errors one should :
- Reasonably increasing the truncation points; if there is no available probability mass between some (smaller) point and the truncation point, the empty range will be padded with zeros;
- Apply the exponential tilting with a careful choice of the tilting parameters.

Exponential tilting
Generates an exponential decay of the distribution's tail.
The tilting operators are defined by

$$\begin{aligned} E_{\theta_j}\mathbf{f}_j &= \left[e^{-\theta_j l} f_j\left(l\right)\right]_{0\leq l\leq r_j-1}, 1 \leq j \leq m, \\ E_{\theta_1,\ldots,\theta_m}\mathbf{f}_0 &= \left[\exp\left\{-\sum_{j=1}^{m}\theta_j l_j\right\} f_0\left(l_1,\ldots,l_m\right)\right]_{0\leq l_j\leq r_j-1, 1\leq j\leq m}, \end{aligned}$$

where $\theta_j > 0, 1 \leq j \leq m$, are the tilting parameters.

FFT algorithm with exponential tilting : Algorithm 2
Step 1. Same as in Algorithm 1;
Step 2. Tilt the vectors $\mathbf{f}_j, 1 \leq j \leq m$, and the multidimensional table $\mathbf{f}_0$;
Step 3. Apply FFT to obtain $\widetilde{E_{\theta_j}\mathbf{f}_j}, 1 \leq j \leq m$, and $\widetilde{E_{\theta_1,\ldots,\theta_m}\mathbf{f}_0}$;
Step 4. Use formula (4.1) to obtain the discrete tilted characteristic function of $\mathbf{S}$, i.e. a multidimensional table of dimension $r_1 \times \ldots \times r_m$;
Step 5. Apply IFFT to this multidimensional table, then untilt the result by $E_{-\theta_1,\ldots,-\theta_m}$ to obtain the p.f. of $\mathbf{S}$.

AE with exponential tilting satisfies :

$$AE_{\theta_1,\ldots,\theta_m}^{Tilt}\left(\mathbf{S}\right) \leq e^{-\sum_{j=1}^{m}\theta_j\left(r_j-1\right)} AE\left(\mathbf{S}\right).$$

To avoid other numerical errors, recommended rough value of $\sum\limits_{j=1}^{m}\theta_j r_j \simeq 20$.
In the **univariate** case : $\theta = 20/r$;
In the **bivariate** case : $\theta_j = 10/r_j$,;
In the **trivariate** case (our numerical examples): $\theta_j = 7/r_j$.

Chapter 5

Monte Carlo method - a particular study case

A well-known method such as Monte Carlo simulation is quite often used to analyze the risks in a project development. This chapter aims to present the method using an ongoing ship design project for a petroleum chemical tanker.
First of all it is important to know that the primary focus during the development of a ship basic/detailed design is to pay attention to an important aspect called risk. Taking into consideration that building a ship implies a lot of risks, an important duty is to prevent them by maximizing the probability and consequences of positive events and, in the same time, by minimizing the probability and consequences of adverse events related to the project's objectives. Shipbuilding industry is for centuries a worldwide subject, see for example Hoving [12], Chida and Davies [4], Briggs Vernon [2] and involves a lot of aspects such as those that we will present in the following. Building new ships means substantial costs, a great consumption of raw matter and materials, significant human resources as well as launching new dedicated techniques and technologies. Consequently, to start a new project, a risk analysis is required to reflect as accurate as possible the implications and consequences of all the factors which are about to participate in the task achievement.

5.1 Monte Carlo simulation for a ship design project

In this study case we consider a short part of a ship design project which was planned to start on January 4, 2016 and scheduled to be completed on July 18, 2017 at a cost of $120.000. The software Primavera Risk Analysis was used.

5.1.1 Simulation 1

We assigned a triunghiular distribution to the duration uncertainty by applying percentages to the minimum, likely and maximum the total duration/ activity. After running the Monte Carlo simulation, the possibility of finishing on time is 90%. Also, we can observe that the P-80 date is now July 04, 2017 for a Triangle distribution, compared with May 30, 2017 for the Uniform distribution.

We can observe that if we choose a uniform distribution the probability to finish in time is more optimistic, 98%. This is due to the shape of the Uniform distribution, but we would expect a Triangular distribution to be more realistic for such a study.

The costs for Simulation 1 are almost similarly in both cases: even though we use a Uniform or a Triangle distribution, there are 94% chances for the Uniform distribution and 96% chances in case we chose a Triangle distribution to maintain the initial budget.

5.1.2 Simulation 2

In this part, we use a Risk Register. Every risk is defined by five attributes such as: probability, schedule, cost, performance and score. Every attribute has assigned a factor scale (low, medium, high) in accordance with its project impact factor.
We assume risks such as:

- Changing the input data
- Leaving employees
- Class certification
- Logistic database

We can easily observe that in this case, the P-80 date is now August 18, 2017 or almost more than one month from the scheduled date of July 18, 2017, and that the possibility of finishing on time is in this scenario only 39% .
Regarding the costs, there are only 19% chances to finish the project with the initial budget ($120.000), and a P-80 posibility to have a cost of $282.448, which means a $162.488 amount difference.

Chapter 6

Conclusions

6.1 General conclusions

In this PhD thesis original results have been presented in accordance with the study of the aggregate claims distributions of a portfolio insurance, portfolio which may contain policies for: fires, floods, earthquakes, traffic accidents, and so on.
We focused firstly on the extension of two bivariate models to a multivariate setting, and, moreover, we created our own multivariate aggregate claims model, with its own dependencies.
Our goal was to obtain an exact recursive formula for the probability function of the multivariate compound distribution of both models.
Why did we do this, and what did we aim by doing this? The answer is obvious, we wanted to contribute to the improvement of the actuarial science in order to assist the actuaries to a better analysis of an insurance scenario, especially when a portfolio is simultaneously affected by different types of claims. It is important to mention here that all these aspects were proved not only in classical mathematical ways, but also with numerical examples.
Furthermore, due to the fact that I personally activate as a planner engineer in the shipbuilding industry, I consider that it was a good idea to complete my thesis with a study case from this field of activity. I wanted to emphasize the strong link between planning and the mathematical science, and affirm on that way that : "Math is everywhere!".
I showed how important a simulation method such as Monte Carlo could be and how this can help us make a relevant risk analysis of a ship design project.
In my opinion, I believe that making this mixture between the classical stochastic theories and real life circumstances makes my research interesting to read and undoubtedly original.
It is very important to mention that the entire research of these thesis is the result

of a strong collaboration with my scientific advisor, PhD Professor Vasile Preda and with PhD Associate Professor Raluca Vernic who, due to their academic experience, coordinated my ideas in the best way possible in order to make a good and efficient work.

Bibliography

[1] Briggs, W.L. and Henson, V.E. (1995). The DFT: An Owners' Manual for the Discrete Fourier Transform. Siam.

[2] Briggs Vernon, L. (2015). History of Shipbuilding on North River, Plymouth County, Massachusetts: With Genealogies of the Shipbuilders, and Accounts of the Industries Upon Its Tributaries, 1640 to 1872 (Classic Reprint), Forgotten Books.

[3] Bühlmann, H. (1984). Numerical evaluation of the compound Poisson distribution: recursion or fast Fourier transform? Scand. Actuar. J., 116-126.

[4] Chida, T., Davies, O. (2013). The Japanese Shipping and Shipbuilding Industries: A History of their Modern Growth,Bloomsbury Academic Collections.

[5] Dlugokecki, V., Fanguy, D., Hepintstall, L.(2009). Leading the way for midtier shipyards to implement design for production methodologies, Journal of Ship Production, Vol. 25, No. 2, p. 99-108.

[6] Dorp van J.R. ; Duffey, M. R. (1997). Statistical dependence in risk analysis for project networks using Monte Carlo methods, Elsevier International.

[7] Embrechts,P., Grübel R. and Pitts, S. M. (1993). *Some applications of the fast Fourier transform algorithm in insurance mathematics. This paper is dedicated to Professor WS Jewell on the occasion of his 60th birthday*, Stat. Neerl.,59-75, 47.

[8] Grubel, R. and Hermesmeier, R. (1999). Computation of compound distributions I: aliasing errors and exponential tilting. ASTIN Bulletin 29 (2), 197-214.

[9] Grubel, R. and Hermesmeier, R. (2000). Computation of compound distributions II: discretization errors and Richardson extrapolation. ASTIN Bulletin 30 (2), 309-331.

[10] Heckman, P. E. and Meyers, G. G. (1983). The calculation of aggregate loss distributions from claim severity and claim count distributions. Proceedings of the Casualty Actuarial Society LXX, 22-61.

[11] Hesselager, O. (1996a). A recursive procedure for calculation of some mixed compound Poisson distributions. Scandinavian Actuarial Journal, 1, pp. 54-63.

[12] Hoving, A.J.,Wildeman, D.(2012). Nicolaes Witsen and Shipbuilding in the Dutch Golden Age,Texas AM University Press.

[13] Jerri, A.(1992). Integral and discrete transforms with applications and error analysis (Vol. 162). CRC.

[14] Jin, T.,Ren, J. (2014). Recursions and fast Fourier transforms for a new bivariate aggregate claims model. Scandinavian Actuarial Journal, 8, pp. 729-752.

[15] Jin, T., Provost, S.B.,Ren, J. (2016). Recursions and fast Fourier transforms for a new bivariate aggregate claims model. Scandinavian Actuarial Journal, 3, pp. 216-245.

[16] Johnson, N.L., Kotz, S., Balakrishnan, N. (1997). Discrete Multivariate Distributions. Wiley, New York.

[17] Klugman, S.A.,Panjer, H.H.,Willmot, G.E. (2004). Loss Models: From Data to Decisions, 2nd edn, Wiley-Interscience. Hoboken.

[18] Kolic, D., Fafandjel, N., Calic, B.(2010).Determining how to apply the design for production concept in shipyards through risk analysis, Engineering Review, 30-1, 63-72.

[19] Panjer, H. H. (1981). Recursive evaluation of a family of compound distributions, ASTIN Bull., 12(1), 22-26.

[20] Panjer, H.H. (1990). The aggregate claims distribution and stop-loss reinsurance. Transactions of the Society of Actuaries, 32, pp. 523-545

[21] Robe-Voinea, E.G. and Vernic, R. (2015a). On a multivariate aggregate claims model with multivariate Poisson counting distribution. To appear in Proceedings of the Romanian Academy - Series A.

[22] Robe-Voinea, E.G. and Vernic, R.(2015b). On the recursive evaluation of a certain multivariate compound distribution. 8th Congress of Romanian Mathematicians, June 28-July 1, 2015, "Alexandru Ioan Cuza" University of Iasi.

[23] Robe-Voinea, E.G. and Vernic, R. (2015b). Another approach to the evaluation of a certain multivariate compound distribution. To appear in Analele Ştiintifice ale Universitatii Ovidius

[24] Robe-Voinea, E.G. and Vernic, R.,(2016). Fast Fourier Transform for multivariate aggregate claims, *Computational and Applied, Springer International Publishing*, DOI 10.1007/s40314-016-0336-6, Print ISSN 0101-8205, Online ISSN 1807-0302.

[25] Robe-Voinea, E.G. and Vernic, R. (2015c). On the recursive evaluation of a certain multivariate compound distribution. To appear in Acta Mathematicae Applicatae Sinica (English Series), Springer International Publishing.

[26] Schaller, P. and Temnov, G. (2008). Efficient and precise computation of convolutions: applying fft to heavy tailed distributions. Computational Methods in Applied Mathematics 8 (2), 187-200.

[27] Storch, R.L., Lim, S.(1999).Improving flow to achieve lean manufacturing in shipbuilding, Production Planning and Control, Vol. 10, No.2, p.127-137.

[28] Sundt, B. (1992). On some extensions of Panjer's class of counting distributions. ASTIN Bulletin 22, 1, 61-80.

[29] Sundt, B.(1999). On multivariate Panjer recursions. ASTIN Bulletin, 29, 1, 29-45.

[30] Sundt, B.(2000a). Multivariate compound Poisson distributions and infinite divisibility. ASTIN Bull. 30, 305-308.

[31] Sundt, B.(2000b). On multivariate Vernic recursions. ASTIN Bull. 30, 111-122.

[32] Sundt, B., Vernic, R.(2004). Recursions for compound mixed multivariate Poisson distributions. Bltter der DGVFM, 26, 4, pp. 665-691.

[33] Sundt, B.,Vernic, R.(2009). Recursions for Convolutions and Compound Distributions with Insurance Applications, EAA Lectures Notes. Springer-Verlag Berlin.

[34] Vernic, R.(1999). Recursive evaluation of some bivariate compound distributions. ASTIN Bull. 29, 315-325.

[35] Williams, R.E.(1980). Computing the probability function for aggregate claims. Proceedings of the Canadian Institute of Actuaries, XI, pp. 39-47.

[36] Willmot, G.E.(1986). Mixed compound Poisson distributions. ASTIN Bulletin, 16, pp. 59-79.

[37] Winston, W.L.(2004) Introduction to Probability Models Operations Research, Vol. 2, 4th edition, Thomson Learning, Canada.