
Electronic Thesis and Dissertation Repository

March 2014

Computing and Approximation Methods for the Distribution of Multivariate Aggregate Claims

Tao Jin

The University of Western Ontario

Supervisor

Jiandong Ren

The University of Western Ontario

Graduate Program in Statistics and Actuarial Sciences

A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

© Tao Jin 2014

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>

 Part of the [Statistical Models Commons](#)

Recommended Citation

Jin, Tao, "Computing and Approximation Methods for the Distribution of Multivariate Aggregate Claims" (2014). *Electronic Thesis and Dissertation Repository*. 1914.
<https://ir.lib.uwo.ca/etd/1914>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact tadam@uwo.ca, wlsadmin@uwo.ca.

**Computing and Approximation Methods for the Distribution of
Multivariate Aggregate Claims**

(Thesis format: Monograph)

by

Tao Jin

Graduate Program in Statistical and Actuarial Sciences

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

March 22, 2014

© Tao Jin, 2014

Abstract

Insurance companies typically face multiple sources (types) of claims. Therefore, modeling dependencies among different types of risks is extremely important for evaluating the aggregate claims of an insurer. In the first part of this thesis, we consider three classes of bivariate counting distributions and the corresponding compound distributions introduced in a 1996 paper by Hesselager. We implement the recursive methods for computing the joint probability functions derived by Hesselager and then compare the results with those obtained from fast Fourier transform (FFT) methods. In applying the FFT methods, we extend the concept of exponential tilting for univariate FFT proposed by Grübel and Hermesmeier to the bivariate case. Our numerical results show that although the recursive methods yield the exact compound distributions if the floating-point representation error is ignored, they generally consume more computation time than the FFT methods. On the other hand, although FFT methods are in general very fast, they suffer from the so called alias error. However, the alias error can be effectively reduced via the introduced exponential tilting. Therefore, the FFT methods constitute viable alternatives to the recursive methods for computing the joint probabilities. In the second part of the thesis, we introduce a multivariate aggregate claims model, which allows dependencies among claim numbers as well as dependencies among claim sizes. This model makes practical sense because insurance companies typically write multi lines (types) of insurance policies and the claims from different lines of businesses are usually dependent. For example, in auto insurance, insurance companies have to pay claims due to property damages and bodily injuries. The numbers of claims from property damages and bodily injuries are typically dependent. In addition, one would expect that the sizes of the two types of claims are dependent because some accidents cause two types of claims simultaneously. For this proposed model, we derive recursive formulas for the joint probability functions of different types of claims. In addition, we show that the concepts of exponential tilting in the multivariate FFT can be ap-

plied to compute the joint probability functions of the various types of claims in the introduced multivariate aggregate claims model. Numerical examples are provided to compare the accuracy and efficiency of the two computation methods. In the third part of the thesis, we apply a moment-based technique to approximate the distribution of univariate and bivariate aggregate claims. The numerical examples presented herein indicate that the proposed approximation method constitutes another viable alternative to the recursive and FFT methods.

Key words: Bivariate counting distributions, compound distributions, recursions, fast Fourier transform, aliasing error, exponential tilting, Panjer's recursion, multivariate aggregate claims.

To my family members.

Acknowledgments

A whole host of people provided me with stimulation, inspiration and motivation during the course of my studies. Every one has contributed in helping me get to where I am today. My sincere thanks go to:

- My supervisor, Dr. Jiandong Ren.
- Dr. Serge Provost, for his support and advice over the years.
- All faculty members and staffs in the Department of Statistical and Actuarial Sciences.

Finally, I would like to thank the examination committee members for their valuable remarks and considerable recommendations.

Contents

Abstract	ii
Acknowledgements	v
List of Tables	viii
List of Figures	x
List of Acronyms & Abbreviations	xi
1 Introduction	1
2 Evaluations of certain bivariate compound distributions	5
2.1 Introduction	5
2.2 Three bivariate models and the corresponding recursive methods . . .	6
2.2.1 Model A	7
2.2.2 Model B	7
2.2.3 Model C	7
2.3 Fast Fourier transforms	9
2.4 Numerical examples	12
2.4.1 An example for model A	13
2.4.2 An example for model B	14
2.4.3 An example for model C	15
2.5 Conclusions	16
3 Evaluations of a new bivariate aggregate claims model	18
3.1 Introduction	18
3.2 Recursive formulas	20
3.2.1 Model A	21
3.2.2 Model B	23
3.2.3 Model C	27
3.3 Fast Fourier transforms	32

3.4	Numerical examples	36
3.4.1	An example for model A	36
3.4.2	An example for model B	39
3.4.3	An example for model C	41
3.5	Conclusions	43
4	Moment-based density approximations of aggregate claims	44
4.1	Introduction	44
4.2	Approximating the distribution of univariate aggregate claims	45
4.2.1	Numerical examples	48
4.3	Approximating the distribution of bivariate aggregate claims	53
4.3.1	Numerical examples	56
4.4	Conclusions	61
5	Concluding Remarks	62
	References	64
A	Recursive formulas for three models in Hesselager (1996a)	67
A.1	Recursive formulas for model A	67
A.2	Recursive formulas for model B	68
A.3	Recursive formulas for model C	69
Vita		70

List of Tables

2.1	Values of compound probabilities $g(x, y)$	13
2.2	Logarithms of the ratios of the values obtained from the two methods.	14
2.3	Computation times of the two methods (in seconds).	14
2.4	Values of compound probabilities $g(x, y)$	15
2.5	Logarithms of the ratios of the values obtained from the two methods.	15
2.6	Computation times of the two methods (in seconds).	15
2.7	Values of compound probabilities $g(x, y)$	16
2.8	Logarithms of the ratios of the values obtained from the two methods.	16
2.9	Computation times of the two methods (in seconds).	16
3.1	Model A: Some values of $g(x, y)$	37
3.2	Model A: Number of significant decimal digits agreed for $g(x, y)$ computed with recursive and FFT methods.	37
3.3	Model A: Computation times of recursive and FFT methods (in seconds).	38
3.4	Model B: Some values of $g(x, y)$	39
3.5	Model B: Number of significant decimal digits agreed for $g(x, y)$ computed with recursive and FFT methods.	40
3.6	Model B: Computation times of recursive and FFT methods (in seconds).	40
3.7	Model C: Some values of $g(x, y)$	41
3.8	Model C: Number of significant decimal digits agreed for $g(x, y)$ computed with recursive and FFT methods.	42
3.9	Model C: Computation times of recursive and FFT methods (in seconds).	42
4.1	Some values of $F_{X,Y}(x, y)$ calculated from the recursive approach and the proposed method.	58
4.2	Some values of $F_{X,Y}(x, y)$ calculated from the empirical and the approximated distribution functions obtained on the basis of simulated values.	58
4.3	Some values of $F_{X,Y}(x, y)$ calculated from the recursive approach and the proposed method.	59

4.4	Some values of $F_{X,Y}(x,y)$ calculated from the empirical and the approximated distribution functions obtained on the basis of simulated values.	60
-----	--	----

List of Figures

3.1	Model A: Plot of a measure of dependence between X and Y	38
3.2	Model A: Plot of the CDFs of aggregate claims $X+Y$ with and without dependence between X and Y	39
3.3	Model B: Plot of a measure of dependence between X and Y	40
3.4	Model B: Plot of the CDFs of aggregate claims $X+Y$ with and without dependence between X and Y	41
3.5	Model C: Plot of a measure of dependence between X and Y	42
3.6	Model C: Plot of the CDFs of aggregate claims $X+Y$ with and without dependence between X and Y	43
4.1	Exact and approximated (dashed line) density functions.	50
4.2	Exact and approximated (dashed line) distribution functions.	50
4.3	Empirical and approximated (dashed line) distribution functions obtained on the basis of simulated values.	51
4.4	Exact and approximated (dashed line) density functions.	52
4.5	Exact and approximated (dashed line) cumulative distribution functions.	52
4.6	Empirical and approximated (dashed line) distribution functions obtained on the basis of simulated values.	53
4.7	Exact and approximated (dashed line) dependency measures between X and Y	59
4.8	Exact and approximated (dashed line) dependency measures between X and Y	60

List of Acronyms & Abbreviations

PGF	Probability generating function
FFT	Fast Fourier transform
IFFT	Inverse fast Fourier transform

Chapter 1

Introduction

Numerical evaluation of compound distributions is an important task in insurance mathematics and quantitative risk management. In classical risk theory, the aggregate claims of an insurance company during a time period are modeled by the compound random variable

$$X = \sum_{i=1}^N U_i, \quad (1.1)$$

where $X = 0$ when $N = 0$, N denotes the number of claims occurring within the time period, and $\{U_i\}_{i \geq 1}$ are independent and identically distributed claim size random variables, independent of N . Panjer (1981) presented a recursive formula for computing the distribution of X for positive integer-valued U_i s, when the distribution of N belongs to the $R_1(a, b)$ class whose probability function q satisfies the recursion

$$q(n) = \left(a + \frac{b}{n}\right) q(n-1), \quad n \geq 1,$$

for some constants a and b . Sundt (1981) proved that only the binomial distribution, the Poisson distribution and the negative binomial distribution belong to this class of distributions, with each distribution being characterized by a different sign of a . Since then, the recursive method has been studied extensively in the risk theory literature and many extensions of it exist. For example, Sundt and Jewell (1981) extended Panjer's formula to the case when the claim sizes are non-negative; Sundt (1992)

generalized the $(a, b, 0)$ class to the class R_k , in which

$$q(n) = \sum_{i=1}^k \left(\alpha_i + \frac{\beta_i}{n} \right) q(n-i), \quad n = 1, 2, 3, \dots,$$

for suitable constants α_i and β_i , and recursive algorithm for the distribution of X with this more general claim number distribution was obtained; Panjer and Wang (1993) analyzed the stability of the recursive method; Willmot (1993) derived recursive formulas for compound mixed Poisson probabilities. For a comprehensive review of the recursive method, one is referred to the recent book by Sundt and Vernic (2009).

In modern insurance industry, insurers typically face multiple sources (types) of claims. Therefore, it is extremely important to model the inter-dependence among the different sources of risk. This thesis focuses on the developments in multivariate aggregate claims models. In this aspect, Hesselager (1996a) introduced and derived recursive formulas for the joint distribution of the bivariate aggregate claims random variables

$$(X, Y) = \left(\sum_{i=1}^N U_i, \sum_{j=1}^M V_j \right), \quad (1.2)$$

where $X = 0$ when $N = 0$, and $Y = 0$ when $M = 0$. The number of claims (N, M) are dependent but all the claim sizes U_i and V_j are mutually independent non-negative integer-valued random variables and are independent of (N, M) . An application of this model could be in, for example, earthquake insurance, where insurance companies need to evaluate the costs of both property damages and business interruptions. For this case, N and M could represent the numbers of claims from property damages and business interruptions respectively. Obviously they should be dependent. Three types of dependencies between (N, M) were discussed in Hesselager (1996a). For some other models of dependencies among claim numbers, see for example, Wang (1998) and Vernic (1999).

Sundt (1999) and Ambagaspitiya (1999) derived independently multivariate Pan-

jer recursion formulas for the joint distribution of the multivariate aggregate claims

$$(X_1, \dots, X_m) = \sum_{i=1}^N (U_{i1}, \dots, U_{im}), \quad (1.3)$$

where the claim number N is one-dimensional and its distribution belongs to the $R_1(a, b)$ class. However, each claim is a m -dimensional random vector and these claim vectors are mutually independent and identically distributed, and they are independent of the number of claims. For possible connections between the models in Hesselager (1996a) and in Sundt (1999), one is referred to Section 14.6 in Sundt and Vernic (2009).

In practice, both recursive methods as well as transformed based techniques like FFT are widely used. The FFT is an algorithm that can be used for inverting characteristic functions to obtain densities of discrete random variables. The FFT comes from the fields of signal processing. It is explained in detail with applications to aggregate claims calculations by Robertson (1992).

In addition, approximating the aggregate claims distribution using various methods such as the normal approximation and the Esscher approximation was crucially important historically when more accurate methods such as recursions or FFTs were computationally infeasible. Today, approximation methods are still useful where full individual claim numbers or claim sizes information is not available, they could be used to provide quick and relatively straightforward methods for estimating aggregate claims probabilities and as a check on more accurate approaches. For a detailed review, the reader is referred to Hardy (2007).

The remainder of this thesis is organized as follows. Chapter 2 considers three classes of bivariate counting distributions and the corresponding compound distributions introduced in Hesselager (1996a). We implement the recursive methods for computing the joint probability functions of the bivariate compound random variables and then compare the results with those obtained from FFT methods. In Chapter 3, a new multivariate aggregate claims model is introduced, this model allows depen-

dependencies among claim numbers as well as dependencies among claim sizes. For this proposed model, the recursive formulas are derived for the joint probability functions of different types of claims. In addition, we use multivariate FFT method to compute the joint probability functions of the various types of claims, and the results are compared with results from the derived recursive formulas for accuracy and efficiency. Chapter 4 presents a unified moment-based approach to approximate the densities of univariate and bivariate aggregate claims. The numerical examples show this approximation method is fairly accurate and it is quite useful when the information is not available on claim numbers and individual claim sizes.

Chapter 2

Evaluations of certain bivariate compound distributions

2.1 Introduction

Hesselager (1996a) introduced and derived recursive formulas for the joint distribution of the bivariate aggregate claims random variables:

$$(X, Y) = \left(\sum_{i=0}^N U_i, \sum_{i=0}^M V_i \right), \quad (2.1)$$

where (N, M) has a probability function

$$p(n, m) = \Pr(N = n, M = m),$$

and all the claim sizes U_i and V_i are mutually independent and are independent of (N, M) . Their probability functions are denoted by

$$f_1(u) = \Pr(U_i = u), \quad f_2(v) = \Pr(V_i = v), \quad (2.2)$$

on the non-negative integers.

One may also use inverse transformation to compute the joint distribution of bivariate aggregate claims random variables. Previous work on computing compound

distributions via FFT may be found in, for example, Heckman and Meyers (1983) and Bühlmann (1984) for univariate cases, and Clark and Homer (2003) for a bivariate case. Embrechts and Frei (2009) compared the recursive and the inverse transform methods in detail and concluded that the inverse transform method offers tremendous timing advantage and is a viable alternative to the recursive method.

A detailed comparison of FFT and recursive methods for computing bivariate distribution functions does not exist in the literature. Therefore, in this chapter, we consider the three bivariate aggregate claims models introduced in Hesselager (1996a). We compare the proposed recursive algorithms in the paper with the bivariate FFT methods for computing the bivariate distributions. In particular, we extend the tilting method for reducing the alias error associated with FFT method (Grübel and Hermesmeier (1999)) to two-dimensional. We conclude that with tilting, the FFT method is essentially as accurate as the recursive methods, but consumes much less computer time.

The chapter is organized as follows. Section 2.2 briefly reviews the three bivariate models in Hesselager (1996a) and the corresponding recursive formulas; Section 2.3 illustrates the FFT method and its tilting; Section 2.4 provides numerical examples to compare the accuracy and computation speed of the two methods; Section 2.5 concludes this chapter.

2.2 Three bivariate models and the corresponding recursive methods

In this section, we briefly review the three bivariate compound distributions and the corresponding recursive formulas introduced in Hesselager (1996a). For insurance applications of the three types of correlation structures, please refer to Hesselager (1996a).

2.2.1 Model A

Let $K \sim R_1(a, b)$, where the symbol “ \sim ” denotes “has the distribution”, and assume that $K = N + M$. The conditional distribution of N given K is binomial, that is:

$$\Pr(N = n|K = k) = \binom{k}{n} p_1^n p_2^{k-n}, \quad p_1 + p_2 = 1.$$

Let $P_K(s) = E[s^K]$ and $P_{N,M}(s, t) = E[s^N t^M]$ denote the probability generating functions (PGFs) of K and (N, M) respectively. Then as shown in Hesselager (1996a),

$$P_{N,M}(s, t) = P_K(p_1 s + p_2 t) \tag{2.3}$$

The recursive formulas for computing the joint distribution of (X, Y) were obtained in Hesselager (1996a). They are listed in Appendix A for completeness. We note that, to compute $g(x, y) = \Pr(X = x, Y = y)$, using the recursive methods the number of floating point operations involved is of order $O(xy(x + y))$.

2.2.2 Model B

Let $N = Z_0 + Z_1$ and let $M = Z_0 + Z_2$, where Z_0, Z_1 and Z_2 are mutually independent and $Z_j \sim R_1(a_j, b_j)$ for $j = 0, 1, 2$.

The PGF of (N, M) in this case is:

$$P_{N,M}(s, t) = P_0(st)P_1(s)P_2(t), \tag{2.4}$$

where P_j denotes the PGF of Z_j . The recursive formulas developed in Hesselager (1996a) for this case are listed in Appendix A. The number of floating point operations needed to compute $g(x, y)$ is of order $O(x^2 y^2)$.

2.2.3 Model C

Let Θ be a random variable on the support $[\sigma_1, \sigma_2]$, where $0 \leq \sigma_1 < \sigma_2 \leq \infty$. Assume that Θ has a probability density function u that satisfies

$$\frac{d}{d\theta} \log u(\theta) = \frac{\sum_{i=0}^k a_i \theta^i}{\sum_{i=0}^k b_i \theta^i},$$

for suitable constants a_i and b_i , and

$$\sum_{i=0}^k b_i \theta^i u(\theta) \rightarrow 0, \text{ when } \theta \rightarrow \sigma_1, \sigma_2.$$

One commonly used distribution that satisfies the above property is the Gamma distribution with shape parameter α and scale parameter s , which is denoted by $\text{Gamma}(\alpha, s)$ and has density function

$$u(\theta) = \frac{\theta^{\alpha-1} e^{-\theta/s}}{s^\alpha \Gamma(\alpha)}, \text{ for } \theta \geq 0 \text{ and } \alpha, s > 0. \quad (2.5)$$

In this case,

$$\frac{d}{d\theta} \log u(\theta) = \frac{(\alpha - 1) - \frac{1}{s}\theta}{\theta} = \frac{\sum_{i=0}^1 a_i \theta^i}{\sum_{i=0}^1 b_i \theta^i},$$

where $a_0 = \alpha - 1$, $a_1 = -\frac{1}{s}$, $b_0 = 0$, and $b_1 = 1$.

Given $\Theta = \theta$, N and M are conditionally independent Poisson random variables with parameters $\theta\lambda_1$ and $\theta\lambda_2$ respectively. The Poisson distribution with parameter μ , denoted by $\text{Poisson}(\mu)$, has probability function

$$p(n) = \frac{e^{-\mu} \mu^n}{n!}, \text{ } n \geq 0 \text{ and } \mu > 0.$$

Then the PGF of (N, M) is given by

$$P_{N,M}(s, t) = \int_{\sigma_1}^{\sigma_2} e^{\theta[\lambda_1(s-1) + \lambda_2(t-1)]} u(\theta) d\theta. \quad (2.6)$$

The recursive formulas derived in Hesselager(1996a) are listed in Appendix A. The number of floating point operations needed to compute $g(x, y)$ is of order $O(xy(x+y))$.

A remark is necessary here for the recursive method. In the univariate case, Panjer and Wang (1993) showed that for Poisson and negative binomial claim number distributions, the recursive formula is stable, producing relative errors that do not grow fast. For the compound binomial distribution, the negative terms in the formula can cause the successive values to blow up with alternating signs, but it does not happen frequently in practice. These results apply to the bivariate cases discussed in this chapter.

Another problem needs to be dealt with in practice is underflow/overflow, which occurs when the starting probability is smaller than the smallest number that can be represented on a computer. Panjer and Willmot (1986) suggested several different ways to overcome the problem. Their methods are applicable to the bivariate case as well.

2.3 Fast Fourier transforms

Given the joint PGF $P_{N,M}(s, t)$ of the claim numbers (N, M) , because the claim sizes U and V are assumed to be independent, the characteristic function of (X, Y) is simply

$$\phi_{XY}(s, t) = P_{N,M}(\phi_U(s), \phi_V(t)), \quad (2.7)$$

where $\phi_U(s)$ and $\phi_V(t)$ are the characteristic functions of claim sizes U and V . Therefore, the distribution of (X, Y) may be obtained by inverting the characteristic function. We next briefly introduce the bivariate FFT and the inverse fast Fourier transform (IFFT).

Let $f(x, y)$ denote a function defined on the integer values of $x = 0, 1, \dots, n - 1$, and $y = 0, 1, \dots, m - 1$. Let $\mathbb{f}_{n \times m}$ denotes the matrix* of probabilities with its ij th element f_{ij} being $f(i, j)$. Then its discrete Fourier transform $\tilde{\mathbb{f}}_{n \times m}$ has ij th element

$$\left. \begin{aligned} \tilde{f}_{ij} &= \sum_{c=0}^{m-1} \sum_{r=0}^{n-1} f_{rc} \exp\left(\frac{2\pi i}{n} ri\right) \exp\left(\frac{2\pi i}{m} cj\right) \\ i &= 0, 1, \dots, n - 1, \quad j = 0, 1, \dots, m - 1. \end{aligned} \right\} \quad (2.8)$$

The inverse mapping is

$$\left. \begin{aligned} f_{ij} &= \frac{1}{mn} \sum_{c=0}^{m-1} \sum_{r=0}^{n-1} \tilde{f}_{rc} \exp\left(-\frac{2\pi i}{n} ri\right) \exp\left(-\frac{2\pi i}{m} cj\right) \\ i &= 0, 1, \dots, n - 1, \quad j = 0, 1, \dots, m - 1. \end{aligned} \right\} \quad (2.9)$$

*For convenience, in this chapter we let the row and column indices of the matrix start from 0.

As in the one-dimensional case, to take the speedy advantage of FFT, one needs to choose n and m to be powers of 2. For our application, we may apply the following steps to calculate the bivariate aggregate claims distribution:

1. Set truncation points n and m for the claim sizes U and V and obtain the discretized claim size distributions:

$$\mathbf{f}_1 = \{f_1(0), f_1(1), \dots, f_1(n-1)\} \text{ and } \mathbf{f}_2 = \{f_2(0), f_2(1), \dots, f_2(m-1)\},$$

where $n = 2^{r_1}$ and $m = 2^{r_2}$ for some integers r_1 and r_2 . Notice that if one or both of the claim size distributions have bounded supports, the vector of probabilities can be appropriately padded with zeros in order to force m or n to be the power of 2.

2. Apply one-dimensional FFT to the two vectors of claim size distributions to obtain two vectors:

$$\tilde{\mathbf{f}}_1 = \{\tilde{f}_1(0), \tilde{f}_1(1), \dots, \tilde{f}_1(n-1)\} \text{ and } \tilde{\mathbf{f}}_2 = \{\tilde{f}_2(0), \tilde{f}_2(1), \dots, \tilde{f}_2(m-1)\}.$$

3. Use formulas (2.3), (2.4), (2.6), and (2.7) to obtain the matrix $\tilde{\phi}_{XY}$ with the ij th element $P_{N,M}(\tilde{f}_1(i), \tilde{f}_2(j))$ for $i = 0, \dots, n-1$ and $j = 0, \dots, m-1$.
4. Apply the IFFT (2.9) to $\tilde{\phi}_{XY}$ to obtain the probability function of (X, Y) .

As discussed in Gröbel and Hermesmeier (1999), the truncation of claim size distribution in the first step and the “wrap around” effect caused by the discrete Fourier transform introduce an *aliasing error*, where the compound mass that lies at the truncation point and beyond will be wrapped around and erroneously appears below the truncation point. This problem may be alleviated by choosing large enough truncation points n and m .

However, Gröbel and Hermesmeier (1999) introduced a more efficient way to alleviate the problem. The authors showed that the aliasing errors can be eliminated for

all practical purposes by a suitable change of measure, known as *exponential tilting* to the claim size distribution, which forces its tail to decrease at an exponential rate. In particular, let

$$E_\theta \mathbf{f} = [e^{-\theta j} f(j)]_{j=0,1,\dots,n-1},$$

where n is the truncation point and $\theta > 0$ is the tilting parameter. Because the operator E_θ commutes with convolutions, the distribution of the aggregate claims may be obtained by applying the operation $E_{-\theta}$ to the IFFT of $P_N(\widetilde{E_\theta \mathbf{f}})$, where P_N is the PGF of the claim number N and $\widetilde{E_\theta \mathbf{f}}$ denotes the FFT of $E_\theta \mathbf{f}$.

The tilting method may be applied to the bivariate models discussed in this chapter. This is due to the fact that the tilting operator commutes with convolutions, and in all three models considered in the chapter, as in the univariate model, the joint distribution of the aggregate claims is obtained through convolutions. That is,

$$g(x, y) = \sum_{i \geq 0} \sum_{j \geq 0} p(i, j) f_1^{*i}(x) f_2^{*j}(y),$$

where f^{*k} denote the k th convolution of f .

Therefore, when applying tilting, the distributions of the bivariate compound random variables can be computed with the following steps:

1. Set truncation points n and m for the claim sizes U and V and obtain the discretized claim size distributions:

$$\mathbf{f}_1 = \{f_1(0), f_1(1), \dots, f_1(n-1)\} \text{ and } \mathbf{f}_2 = \{f_2(0), f_2(1), \dots, f_2(m-1)\},$$

where $n = 2^{r_1}$ and $m = 2^{r_2}$ for some integers r_1 and r_2 .

2. Tilt these two sequences:

$$\mathbf{f}_1 \mapsto E_{\theta_1} \mathbf{f}_1 = [e^{-\theta_1 j} f_1(j)]_{j=0,1,\dots,n-1},$$

$$\mathbf{f}_2 \mapsto E_{\theta_2} \mathbf{f}_2 = [e^{-\theta_2 j} f_2(j)]_{j=0,1,\dots,m-1}.$$

Notice that in practice an excessively large tilting parameter θ may result in underflow or overflow problems. As suggested in Grübel and Hermesmeier (1999), choosing $\theta_1 = 10/n$ and $\theta_2 = 10/m$ generally do not lead to numerical difficulties assuming that double precision (64 bit) calculations are used.

3. Apply FFT to the two sequences respectively, resulting in two sequences $\widetilde{E_{\theta_1} \mathbf{f}_1}$ and $\widetilde{E_{\theta_2} \mathbf{f}_2}$.
4. Use formulas (2.3), (2.4), (2.6), and (2.7) to obtain the characteristic function of (X, Y) , resulting in the $n \times m$ matrix:

$$\begin{bmatrix} P_{N,M}(\widetilde{E_{\theta_1} \mathbf{f}_1}(0), \widetilde{E_{\theta_2} \mathbf{f}_2}(0)) & \cdots & P_{N,M}(\widetilde{E_{\theta_1} \mathbf{f}_1}(0), \widetilde{E_{\theta_2} \mathbf{f}_2}(m-1)) \\ P_{N,M}(\widetilde{E_{\theta_1} \mathbf{f}_1}(1), \widetilde{E_{\theta_2} \mathbf{f}_2}(0)) & \cdots & P_{N,M}(\widetilde{E_{\theta_1} \mathbf{f}_1}(1), \widetilde{E_{\theta_2} \mathbf{f}_2}(m-1)) \\ \vdots & \ddots & \vdots \\ P_{N,M}(\widetilde{E_{\theta_1} \mathbf{f}_1}(n-1), \widetilde{E_{\theta_2} \mathbf{f}_2}(0)) & \cdots & P_{N,M}(\widetilde{E_{\theta_1} \mathbf{f}_1}(n-1), \widetilde{E_{\theta_2} \mathbf{f}_2}(m-1)) \end{bmatrix}.$$

5. Apply IFFT to the above matrix. Then untilt each column by applying $E_{-\theta_1}$ and untilt each row by applying $E_{-\theta_2}$. For $i \geq 0$ and $j \geq 0$, the ij th element of the resulting matrix gives $g_{X,Y}(i, j)$.

2.4 Numerical examples

In this section, in order to compare the accuracy and computation speed of the recursive and FFT methods for computing the joint distribution of bivariate compound distributions, we present one numerical example for each of the three models introduced in Section 2.2. In all examples, we assume that claim sizes are Pareto distributed. Before applying the recursive/FFT methods, the distributions are discretized using the standard rounding method (p. 233 in Klugman *et al.* 2008). Because both methods are subject to the same discretization errors, we can focus on the comparisons of errors introduced by the recursive/FFT methods.

The computations are carried out using the free software R. R function FFT is used to carry out the Fourier transforms and their inverses. R function *discretize()* in the *actuar* package is used to discretize the claim size distributions.

2.4.1 An example for model A

Let $K \sim \text{Poisson}(15)$, $N|K = k \sim \text{Binomial}(k, 0.3)$ where the distribution of $\text{Binomial}(m, q)$ has the probability function

$$p(n) = \binom{m}{n} q^n (1 - q)^{m-n}, \quad m = 0, 1, \dots, \quad 0 \leq n \leq m \text{ and } 0 \leq q \leq 1.$$

Let $X \sim \text{Pareto}(3, 5)$, $Y \sim \text{Pareto}(4, 3)$, where the density function of $\text{Pareto}(a, b)$ distribution is

$$f(x) = \frac{ab^a}{(x+b)^{a+1}}, \quad x > 0, \quad a, b > 0.$$

The joint probability mass function at some selected points is computed using the recursive method, the FFT method with truncation points $n = m = 2^{10}$, the FFT method with truncation points $n = m = 2^{12}$, and the FFT method with truncation points $n = m = 2^{12}$ and tilting parameter $\theta = 10/n$. Table 2.1 presents the resulting probabilities. Table 2.2 illustrates the accuracy improvements from exponential tiltings by calculating the logarithms of the ratio of the values obtained by FFT and those obtained by recursion. Table 2.3 reports the computation time of different methods on a personal computer with Intel Core2 Quad CPU Q8200@2.33G and 8GB memory.

(x, y)	Recursion	FFT(2^{10})	FFT(2^{12})	FFT(tilting)
(100,100)	3.656681e-05	3.657364e-05	3.656684e-05	3.656681e-05
(400,100)	1.222787e-06	1.223075e-06	1.222788e-06	1.222787e-06
(400,300)	2.146102e-08	2.146606e-08	2.146104e-08	2.146102e-08
(600,300)	3.535786e-09	3.536731e-09	3.535790e-09	3.535786e-09
(600,600)	2.892395e-11	2.893176e-11	2.892399e-11	2.892395e-11

Table 2.1: Values of compound probabilities $g(x, y)$.

(x, y)	FFT(2^{10})	FFT(2^{12})	FFT(tilting)
(100,100)	8.102916e-05	2.507947e-07	1.253626e-15
(400,100)	1.019978e-04	3.479440e-07	4.339474e-15
(400,300)	1.019077e-04	3.478247e-07	2.859231e-14
(600,300)	1.160764e-04	5.512267e-07	6.793205e-13
(600,600)	1.171996e-04	5.524577e-07	4.554824e-09

Table 2.2: Logarithms of the ratios of the values obtained from the two methods.

(x, y)	Recursion	FFT(2^{10})	FFT(2^{12})	FFT(tilting)
(100,100)	1.00	1.28	20.86	19.39
(400,100)	4.63	1.28	20.86	19.39
(400,300)	15.04	1.28	20.86	19.39
(600,300)	25.03	1.28	20.86	19.39
(600,600)	53.91	1.28	20.86	19.39

Table 2.3: Computation times of the two methods (in seconds).

Since the recursion method obtains the exact values of the compound distribution if the errors from the floating point representation are ignored, the differences between the FFT and the recursion methods are essentially due to aliasing errors. From the first two tables we can see that as the truncation points increases, the alias errors decrease. Moreover, aliasing errors are reduced dramatically by applying tilting.

As reported in Embrechts and Frei (2009), the recursive method appears to require more CPU time than the FFT method.

2.4.2 An example for model B

Let $Z_0 \sim \text{Poisson}(2)$, $Z_1 \sim \text{Poisson}(3)$, $Z_2 \sim \text{Poisson}(5)$; $X \sim \text{Pareto}(3, 5)$, and $Y \sim \text{Pareto}(4, 3)$. As for Model A, the values of probabilities, the logarithms of the ratios of the probabilities calculated with the recursive and the FFT methods, and the computational speeds are reported in Tables 2.4, 2.5 and 2.6 respectively.

(x, y)	Recursion	FFT(2^{10})	FFT(2^{12})	FFT(tilting)
(100,100)	2.545090e-05	2.545801e-05	2.545092e-05	2.545090e-05
(400,100)	1.225507e-06	1.225848e-06	1.225508e-06	1.225507e-06
(400,300)	9.833320e-09	9.836364e-09	9.833330e-09	9.833320e-09
(600,300)	1.590431e-09	1.590992e-09	1.590433e-09	1.590431e-09
(600,600)	1.941624e-11	1.942271e-11	1.941627e-11	1.941624e-11

Table 2.4: Values of compound probabilities $g(x, y)$.

(x, y)	FFT(2^{10})	FFT(2^{12})	FFT(tilting)
(100,100)	1.212200e-04	3.614823e-07	9.643275e-16
(400,100)	1.210978e-04	4.016661e-07	1.157193e-15
(400,300)	1.344243e-04	4.367114e-07	1.515923e-13
(600,300)	1.532655e-04	6.575989e-07	1.031059e-12
(600,600)	1.447259e-04	6.335427e-07	4.215072e-10

Table 2.5: Logarithms of the ratios of the values obtained from the two methods.

(x, y)	Recursion	FFT(2^{10})	FFT(2^{12})	FFT(tilting)
(100,100)	38.69	1.79	30.17	28.14
(400,100)	673.1	1.79	30.17	28.14
(400,300)	2269.36	1.79	30.17	28.14
(600,300)	5467.79	1.79	30.17	28.14
(600,600)	12662.18	1.79	30.17	28.14

Table 2.6: Computation times of the two methods (in seconds).

2.4.3 An example for model C

Let $\Theta \sim \text{Gamma}(3, 5)$, $\lambda_1 = 2$, $\lambda_2 = 3$, $X \sim \text{Pareto}(3, 5)$, and $Y \sim \text{Pareto}(4, 3)$. For this model, the comparisons are shown in Tables 2.7, 2.8 and 2.9.

(x, y)	Recursion	FFT(2^{10})	FFT(2^{12})	FFT(tilting)
(100,100)	2.656440e-06	2.786862e-06	2.656454e-06	2.656440e-06
(400,100)	1.056183e-06	1.189863e-06	1.056186e-06	1.056183e-06
(400,300)	2.838312e-06	2.943274e-06	2.838359e-06	2.838312e-06
(600,300)	2.264384e-06	2.335480e-06	2.264409e-06	2.264384e-06
(600,600)	7.946966e-07	9.796661e-07	7.947693e-07	7.946966e-07

Table 2.7: Values of compound probabilities $g(x, y)$.

(x, y)	FFT(2^{10})	FFT(2^{12})	FFT(tilting)
(100,100)	2.081542e-02	2.290490e-06	6.268129e-15
(400,100)	5.175788e-02	1.318289e-06	4.821637e-15
(400,300)	1.577045e-02	7.208424e-06	1.880439e-14
(600,300)	1.342611e-02	4.766635e-06	1.446491e-14
(600,600)	9.087674e-02	3.973305e-05	8.900743e-14

Table 2.8: Logarithms of the ratios of the values obtained from the two methods.

(x, y)	Recursion	FFT(2^{10})	FFT(2^{12})	FFT(tilting)
(100,100)	2.72	1.18	19.70	17.63
(400,100)	12.62	1.18	19.70	17.63
(400,300)	40.70	1.18	19.70	17.63
(600,300)	66.44	1.18	19.70	17.63
(600,600)	150.13	1.18	19.70	17.63

Table 2.9: Computation times of the two methods (in seconds).

The results from model B and model C also show that the alias errors of the FFT method decrease with higher truncation points and that the errors are essentially eliminated through tilting.

2.5 Conclusions

As previously mentioned, the recursive method yields the exact compound distribution if the floating point representation error is ignored. In addition, contrary to the FFT method, one can stop the recursion at any point of interest instead of calculating

values on all the lattice points.

However, the recursive method can only be applied on certain claim number distributions, and it seems to be computationally expensive to obtain high quantiles of the compound distributions. On the other hand, FFT method can deal with arbitrary claim number distributions as long as the joint pgf is known. In addition, it is computationally very efficient. In particular, comparing the number of the floating point operations required by the recursive and the FFT methods for the three models discussed, we have $O(xy(x+y))$ versus $O(xy(\log_2 x + \log_2 y))$ in model A and model C, and $O(x^2 y^2)$ versus $O(xy(\log_2 x + \log_2 y))$ in model B. Further, the aliasing errors suffered by the FFT methods can essentially be eliminated by exponential tilting.

Chapter 3

Evaluations of a new bivariate aggregate claims model

3.1 Introduction

In Chapter 1, two bivariate aggregate claims models from Hesselager (1996a) and Sundt (1999) are introduced. Notice that in both models, either the claim numbers or the claim sizes are dependent, but not both. However, in practice, as discussed in Cummins and Wiltbank (1983), Frees and Valdez (2008) and Section 14.6 in Sundt and Vernic (2009), both claim numbers and claim sizes can be dependent. For instance, consider property damage and bodily injury claims in a portfolio of auto insurance policies; because some accidents cause both types of claims, it is reasonable to assume that the numbers of the two types of claims are dependent; in addition, the sizes of the two types of claims are also dependent when an accident causes both types of claims. To this end, we propose the following aggregate claims model

$$(X, Y) = \left(\sum_{i=1}^{N_1} U_i + \sum_{k=1}^{N_3} L_k, \sum_{j=1}^{N_2} V_j + \sum_{k=1}^{N_3} Q_k \right), \quad (3.1)$$

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

$$p(n_1, n_2, n_3) = \Pr(N_1 = n_1, N_2 = n_2, N_3 = n_3). \quad (3.2)$$

The claim sizes $\{U_i\}_{i \geq 1}$ and $\{V_j\}_{j \geq 1}$ are mutually independent and are independent of the claim numbers (N_1, N_2, N_3) and claim sizes $\{L_k, Q_k\}_{k \geq 1}$. Their probability functions are denoted by

$$f_1(u) = \Pr(U_i = u) \quad \text{and} \quad f_2(v) = \Pr(V_j = v) \quad (3.3)$$

on the non-negative integers. The claim size vectors $\{L_k, Q_k\}_{k \geq 1}$ are mutually independent and identically distributed and independent of claim numbers (N_1, N_2, N_3) and claim sizes U_i and V_j . They have the same probability function

$$f_3(l, q) = \Pr(L_k = l, Q_k = q), \quad k \geq 1, \quad (3.4)$$

on the non-negative integers.

This chapter focuses on the computation of the probability function of the bivariate aggregate claims (X, Y) defined in (3.1), which is denoted by $g(x, y)$.

Let $P_{X_1, \dots, X_d}(s_1, \dots, s_d)$ denote the PGF of a d -dimensional non-negative integer-valued vector $\{X_1, \dots, X_d\}$, that is

$$P_{X_1, \dots, X_d}(s_1, \dots, s_d) = \mathbb{E} \left[s_1^{X_1} \cdots s_d^{X_d} \right]. \quad (3.5)$$

Then, for model (3.1), conditioning on the claim numbers, we have that

$$\begin{aligned} P_{X,Y}(s, t) &= \sum_{x,y=0}^{\infty} g(x, y) s^x t^y = \mathbb{E} \left[s^{\sum_{i=1}^{N_1} U_i + \sum_{k=1}^{N_3} L_k} t^{\sum_{j=1}^{N_2} V_j + \sum_{k=1}^{N_3} Q_k} \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[s^{\sum_{i=1}^{N_1} U_i + \sum_{k=1}^{N_3} L_k} t^{\sum_{j=1}^{N_2} V_j + \sum_{k=1}^{N_3} Q_k} \mid N_1, N_2, N_3 \right] \right] \\ &= \mathbb{E} \left[(\mathbb{E}[s^{U_1}])^{N_1} (\mathbb{E}[t^{V_1}])^{N_2} (\mathbb{E}[s^{L_1} t^{Q_1}])^{N_3} \right] \\ &= P_{N_1, N_2, N_3}(P_U(s), P_V(t), P_{L,Q}(s, t)). \end{aligned} \quad (3.6)$$

The remaining parts of the chapter focus on the computations of the joint distribution of the two types of claims X and Y . In Section 3.2, we derive recursive formulas for calculating the joint distribution of (X, Y) . Recursive methods can give accurate values of the joint distributions, however, the required computations can be intensive. Therefore, in Section 3.3, we study the FFT method for computing the joint distribution and we generalize the tilting method in Grübel and Hermesmeier (1999) to our multi-dimensional aggregate claims model and show how to apply the generalized tilting method to calculate the joint distribution of (X, Y) . In Section 3.4, we use numerical examples to compare the accuracy and computational speed of the recursive formulas and the FFT methods. We conclude that FFT method consumes much less computer time than recursive method, and the aliasing errors can be reduced by exponential tilting. This result agrees with that obtained in Embrechts and Frei (2009). To illustrate the actuarial applications of the model, we also use the examples in Section 3.4 to demonstrate the interdependencies between the two types of claims (X, Y) and their effects on the risk of the total claims $X + Y$.

3.2 Recursive formulas

In this section, we present recursive formulas for the probability function $g(x, y)$ of the bivariate aggregate claims (X, Y) , for $x, y \geq 0$, assuming that the dependency structures of claim numbers (N_1, N_2, N_3) follow those introduced in Models A, B and C in Hesselager (1996a). Note that we assume the distributions of claim numbers belong to the $R_1(a, b)$ class since this class of distribution is widely used for practical purposes, however, the recursive formulas can be similarly derived for the more generalized R_k class introduced in Chapter 1.

The discrete claim size distributions $f_1(\cdot)$, $f_2(\cdot)$, $f_3(\cdot, \cdot)$ have supports on non-negative integers and may have any form, including those obtained by discretizations of any continuous distributions.

3.2.1 Model A

Let K be the total number of accidents and assume that its distribution belongs to the $R_1(a, b)$ class. Assume that $K = N_1 + N_2 + N_3$ and given $K = k$, the conditional distribution of (N_1, N_2, N_3) is trinomial with parameters k and (p_1, p_2, p_3) , that is

$$\Pr(N_1 = n_1, N_2 = n_2, N_3 = n_3 | K = k) = \frac{k!}{n_1!n_2!n_3!} p_1^{n_1} p_2^{n_2} p_3^{n_3}, \quad (3.7)$$

with $n_1 + n_2 + n_3 = k$, $p_1 + p_2 + p_3 = 1$ and $0 \leq p_i \leq 1$ for $i = 1, 2, 3$.

Then we have

$$\begin{aligned} P_{N_1, N_2, N_3}(s, t, u) &= \mathbb{E} [\mathbb{E}[s^{N_1} t^{N_2} u^{N_3} | K]] = \mathbb{E} [(p_1 s + p_2 t + p_3 u)^K] \\ &= P_K(p_1 s + p_2 t + p_3 u) \end{aligned} \quad (3.8)$$

and

$$\begin{aligned} P_{X, Y}(s, t) &= P_{N_1, N_2, N_3}(P_U(s), P_V(t), P_{L, Q}(s, t)) \\ &= P_K(p_1 P_U(s) + p_2 P_V(t) + p_3 P_{L, Q}(s, t)). \end{aligned} \quad (3.9)$$

Let $I(\cdot)$ denote an indicator function such that $I(A)$ is equal to one if the event A occurs and zero otherwise. Then we have the following theorem.

Theorem 3.1. *Under the conditions in model A, the following starting value and recursive formulas hold:*

$$g(0, 0) = P_K(p_1 f_1(0) + p_2 f_2(0) + p_3 f_3(0, 0)), \quad (3.10)$$

$$\begin{aligned} g(x, y) &= \frac{1}{1 - a(p_1 f_1(0) + p_2 f_2(0) + p_3 f_3(0, 0))} \left[p_1 \sum_{u=1}^x \left(a + \frac{bu}{x} \right) f_1(u) g(x - u, y) \right. \\ &\quad + ap_2 \sum_{v=1}^y f_2(v) g(x, y - v) \\ &\quad \left. + p_3 \sum_{u=0}^x \sum_{v=0}^y I(u + v > 0) \left(a + \frac{bu}{x} \right) f_3(u, v) g(x - u, y - v) \right], \quad x \geq 1, y \geq 0, \end{aligned} \quad (3.11)$$

$$\begin{aligned}
g(x, y) = & \frac{1}{1 - a(p_1 f_1(0) + p_2 f_2(0) + p_3 f_3(0, 0))} \left[p_2 \sum_{v=1}^y \left(a + \frac{bv}{y} \right) f_2(v) g(x, y - v) \right. \\
& + ap_1 \sum_{u=1}^x f_1(u) g(x - u, y) \\
& \left. + p_3 \sum_{v=0}^y \sum_{u=0}^x I(v + u > 0) \left(a + \frac{bv}{y} \right) f_3(u, v) g(x - u, y - v) \right], \quad x \geq 0, y \geq 1.
\end{aligned} \tag{3.12}$$

Proof. The initial value $g(0, 0)$ is straightforwardly derived using the definition of PGF.

In particular, based on (3.9), model A can be represented as

$$(X, Y) = \sum_{i=1}^K (C_i, D_i), \tag{3.13}$$

where the pairs (C_i, D_i) are independent and identically distributed with probability function

$$f(u, v) = I(v = 0)p_1 f_1(u) + I(u = 0)p_2 f_2(v) + p_3 f_3(u, v). \tag{3.14}$$

For this situation, Sundt (1999, Section 4B) derived the following recursive formulas

$$g(x, y) = \sum_{u=0}^x \left(a + \frac{bu}{x} \right) \sum_{v=0}^y f(u, v) g(x - u, y - v), \quad x \geq 1, y \geq 0, \tag{3.15}$$

$$g(x, y) = \sum_{v=0}^y \left(a + \frac{bv}{y} \right) \sum_{u=0}^x f(u, v) g(x - u, y - v), \quad x \geq 0, y \geq 1. \tag{3.16}$$

By inserting the above (3.14) into formulas (3.15) and (3.16) and collecting all items containing $g(x, y)$ to the left-hand sides, the recursive formulas (3.11) and (3.12) in the theorem are obtained. \square

To calculate $g(x, y)$ for any (x, y) , we can first use (3.10) and (3.12) for $x = 0$, then use (3.11) for $x \geq 1$.

In terms of computational intensity, the major difference between equation (3.11) and the classical univariate recursive formula (see for example, equation 6.3 in Sundt

and Jewell 1981) is that in the former bivariate case, a double summation (one with respect to x and one to y) is involved to compute $g(x, y)$, whereas in the latter univariate case, a single summation with respect to x is needed to compute the probability function of a univariate aggregate claims up to value x ; the number of floating point operations required to compute this single summation is of order $O(x \cdot x) = O(x^2)$, see for example Klugman *et al.* (2008, p. 227). Therefore we may deduce that the number of floating point operations needed to compute $g(x, y)$ using (3.11) is of order $O(xy \cdot xy) = O(x^2y^2)$.

3.2.2 Model B

Let $N_1 = Z_1 + Z_0$, $N_2 = Z_2 + Z_0$ and $N_3 = Z_3 + Z_0$, where Z_1, Z_2, Z_3 and Z_0 are mutually independent and $Z_j \sim R_1(a_j, b_j)$ for $j = 0, 1, 2, 3$. The PGF of (N_1, N_2, N_3) is

$$\begin{aligned} P_{N_1, N_2, N_3}(s, t, u) &= \mathbb{E} [s^{N_1} t^{N_2} u^{N_3}] = \mathbb{E} [s^{Z_1+Z_0} t^{Z_2+Z_0} u^{Z_3+Z_0}] \\ &= \mathbb{E} [s^{Z_1} t^{Z_2} u^{Z_3} (stu)^{Z_0}] = P_{Z_1}(s) P_{Z_2}(t) P_{Z_3}(u) P_{Z_0}(stu). \end{aligned} \quad (3.17)$$

Then, by (3.6)

$$\begin{aligned} P_{X,Y}(s, t) &= P_{N_1, N_2, N_3}(P_U(s), P_V(t), P_{L,Q}(s, t)) \\ &= P_{Z_1}(P_U(s)) P_{Z_2}(P_V(t)) P_{Z_3}(P_{L,Q}(s, t)) P_{Z_0}(P_U(s) P_V(t) P_{L,Q}(s, t)). \end{aligned} \quad (3.18)$$

Let f^{*2} denote the 2nd convolution of f , and define two constants C_1 and C_2 as

$$\begin{aligned} C_1 &= 1 - a_1 f_1(0) - a_3 f_3(0, 0) + a_1 a_3 f_1(0) f_3(0, 0) - a_0 f_1(0) f_2(0) f_3(0, 0) \\ &\quad + a_0 a_1 f_1^2(0) f_2(0) f_3(0, 0) - a_0 a_1 a_3 f_1^2(0) f_2(0) f_3^2(0, 0) + a_0 a_3 f_1(0) f_2(0) f_3^2(0, 0), \end{aligned}$$

$$\begin{aligned} C_2 &= 1 - a_2 f_2(0) - a_3 f_3(0, 0) + a_2 a_3 f_2(0) f_3(0, 0) - a_0 f_2(0) f_1(0) f_3(0, 0) \\ &\quad + a_0 a_2 f_2^2(0) f_1(0) f_3(0, 0) - a_0 a_2 a_3 f_2^2(0) f_1(0) f_3^2(0, 0) + a_0 a_3 f_2(0) f_1(0) f_3^2(0, 0). \end{aligned}$$

Then we have the following theorem.

Theorem 3.2. *Under the conditions in model B, the following starting value and recursive formulas hold:*

$$g(0, 0) = P_{Z_1}(f_1(0))P_{Z_2}(f_2(0))P_{Z_3}(f_3(0, 0))P_{Z_0}(f_1(0)f_2(0)f_3(0, 0)), \quad (3.19)$$

$$\begin{aligned}
g(x, y) = & \frac{1}{C_1} \left\{ \sum_{u=1}^x \left(a_1 + \frac{b_1 u}{x} \right) f_1(u) g(x-u, y) \right. \\
& + \sum_{u=0}^x \sum_{v=0}^y I(u+v > 0) \left(a_3 + \frac{b_3 u}{x} \right) f_3(u, v) g(x-u, y-v) \\
& - \sum_{u=1}^x \left[\sum_{i=0}^u \sum_{v=0}^y I(i+v+x-u > 0) \frac{a_1(a_3 u + b_3 i)}{x} f_3(i, v) g(u-i, y-v) \right] f_1(x-u) \\
& + \sum_{u=1}^x \sum_{v=0}^y \left[\sum_{i=0}^u \left(\sum_{j=0}^v I(i+j+x+y-u-v > 0) \frac{a_0(u-i)}{x} [f_1(i) - a_1 f_1^{*2}(i)] \times \right. \right. \\
& \left. \left. f_2(j) g(u-i, v-j) - \frac{a_3(a_1 + b_1) i}{x} f_1(i) g(u-i, v) \right) \right] f_3(x-u, y-v) \\
& + \sum_{u=1}^x \sum_{v=0}^y \left[\sum_{i=0}^u \sum_{j=0}^v \left(\frac{(a_0 + b_0) u}{x} f_1(i) - \frac{2a_1(a_0 + b_0) u + (a_0 b_1 - a_1 b_0) i}{2x} f_1^{*2}(i) \right) \times \right. \\
& \left. f_2(j) f_3(u-i, v-j) \right] g(x-u, y-v) \\
& + \sum_{u=1}^x \sum_{v=0}^y \left[\sum_{i=0}^u \sum_{j=0}^v I(i+j+x+y-u-v > 0) \left(\frac{2a_0 a_1 a_3 u + a_3(a_0 b_1 + a_1 b_0) i}{2x} \times \right. \right. \\
& \left. \left. f_1^{*2}(i) - \frac{a_3(a_0 u + b_0 i)}{x} f_1(i) \right) f_2(j) g(u-i, v-j) \right] f_3^{*2}(x-u, y-v) \\
& + (2a_0 a_3 + a_0 b_3 + a_3 b_0) \sum_{m=1}^x \sum_{n=0}^y \left[\sum_{u=1}^m \sum_{v=0}^n \left[\sum_{i=0}^u \sum_{j=0}^v \frac{u-i}{x} [a_1 f_1^{*2}(i) - f_1(i)] \times \right. \right. \\
& \left. \left. f_2(j) f_3(u-i, v-j) \right] g(m-u, n-v) \right] f_3(x-m, y-n) \left. \right\}, \quad x \geq 1, y \geq 0,
\end{aligned} \quad (3.20)$$

$$g(x, y) = \frac{1}{C_2} \left\{ \sum_{v=1}^y \left(a_2 + \frac{b_2 v}{y} \right) f_2(v) g(x, y-v) \right.$$

$$\begin{aligned}
& + \sum_{v=0}^y \sum_{u=0}^x I(v+u > 0) \left(a_3 + \frac{b_3 v}{y} \right) f_3(u, v) g(x-u, y-v) \\
& - \sum_{v=1}^y \left[\sum_{j=0}^v \sum_{u=0}^x I(j+u+y-v > 0) \frac{a_2(a_3 v + b_3 j)}{y} f_3(u, j) g(x-u, v-j) \right] f_2(y-v) \\
& + \sum_{v=1}^y \sum_{u=0}^x \left[\sum_{j=0}^v \left(\sum_{i=0}^u I(j+i+y+x-v-u > 0) \frac{a_0(v-j)}{y} [f_2(j) - a_2 f_2^{*2}(j)] \times \right. \right. \\
& \quad \left. \left. f_1(i) g(u-i, v-j) - \frac{a_3(a_2 + b_2)j}{y} f_2(j) g(u, v-j) \right) \right] f_3(x-u, y-v) \\
& + \sum_{v=1}^y \sum_{u=0}^x \left[\sum_{j=0}^v \sum_{i=0}^u \left(\frac{(a_0 + b_0)v}{y} f_2(j) - \frac{2a_2(a_0 + b_0)v + (a_0 b_2 - a_2 b_0)j}{2y} f_2^{*2}(j) \right) \times \right. \\
& \quad \left. f_1(i) f_3(u-i, v-j) \right] g(x-u, y-v) \\
& + \sum_{v=1}^y \sum_{u=0}^x \left[\sum_{j=0}^v \sum_{i=0}^u I(j+i+y+x-v-u > 0) \left(\frac{2a_0 a_2 a_3 v + a_3(a_0 b_2 + a_2 b_0)j}{2y} \times \right. \right. \\
& \quad \left. \left. f_2^{*2}(j) - \frac{a_3(a_0 v + b_0 j)}{y} f_2(j) \right) f_1(i) g(u-i, v-j) \right] f_3^{*2}(x-u, y-v) \\
& + (2a_0 a_3 + a_0 b_3 + a_3 b_0) \sum_{n=1}^y \sum_{m=0}^x \left[\sum_{v=1}^n \sum_{u=0}^m \left[\sum_{j=0}^v \sum_{i=0}^u \frac{v-j}{y} [a_2 f_2^{*2}(j) - f_2(j)] \times \right. \right. \\
& \quad \left. \left. f_1(i) f_3(u-i, v-j) \right] g(m-u, n-v) \right] f_3(x-m, y-n) \Big\}, \quad x \geq 0, y \geq 1.
\end{aligned} \tag{3.21}$$

Proof. Similar to the proof of Theorem 3.1, the initial value $g(0, 0)$ is obtained by setting $s, t = 0$ in (3.18).

Differentiating (3.18) with respect to s and using the identity for $Z_j \sim R_1(a_j, b_j)$, $j = 0, 1, 2, 3$,

$$(1 - a_j s) P'_{Z_j}(s) = (a_j + b_j) P_{Z_j}(s),$$

one obtains

$$\frac{\partial P_{X,Y}(s, t)}{\partial s} = P_{X,Y}(s, t) \frac{(a_1 + b_1) P'_U(s)}{1 - a_1 P_U(s)} + P_{X,Y}(s, t) \frac{(a_3 + b_3) \frac{\partial P_{L,Q}(s, t)}{\partial s}}{1 - a_3 P_{L,Q}(s, t)} +$$

$$P_{X,Y}(s,t) \frac{(a_0 + b_0)P_V(t) \left[P'_U(s)P_{L,Q}(s,t) + P_U(s) \frac{\partial P_{L,Q}(s,t)}{\partial s} \right]}{1 - a_0 P_U(s)P_V(t)P_{L,Q}(s,t)}.$$

Rearranging terms in the above equation results in

$$\begin{aligned} \frac{\partial P_{X,Y}(s,t)}{\partial s} = & (a_1 + b_1)P'_U(s)P_{X,Y}(s,t) \\ & - (2a_0a_1 + a_0b_1 + a_1b_0)P'_U(s)P_U(s)P_V(t)P_{L,Q}(s,t)P_{X,Y}(s,t) \\ & - a_3(a_1 + b_1)P'_U(s)P_{L,Q}(s,t)P_{X,Y}(s,t) \\ & + (2a_3a_0a_1 + a_3a_0b_1 + a_3a_1b_0)P'_U(s)P_U(s)P_V(t)P_{L,Q}^2(s,t)P_{X,Y}(s,t) \\ & + (a_3 + b_3) \frac{\partial P_{L,Q}(s,t)}{\partial s} P_{X,Y}(s,t) \\ & - (2a_0a_3 + a_0b_3 + a_3b_0)P_U(s)P_V(t) \frac{\partial P_{L,Q}(s,t)}{\partial s} P_{L,Q}(s,t)P_{X,Y}(s,t) \\ & - a_1(a_3 + b_3)P_U(s) \frac{\partial P_{L,Q}(s,t)}{\partial s} P_{X,Y}(s,t) \\ & + (2a_1a_0a_3 + a_1a_0b_3 + a_1a_3b_0)P'_U(s)P_V(t) \frac{\partial P_{L,Q}(s,t)}{\partial s} P_{L,Q}(s,t)P_{X,Y}(s,t) \\ & + (a_0 + b_0)P'_U(s)P_V(t)P_{L,Q}(s,t)P_{X,Y}(s,t) \\ & - a_3(a_0 + b_0)P'_U(s)P_V(t)P_{L,Q}^2(s,t)P_{X,Y}(s,t) \\ & + (a_0 + b_0)P_U(s)P_V(t) \frac{\partial P_{L,Q}(s,t)}{\partial s} P_{X,Y}(s,t) \\ & - a_1(a_0 + b_0)P'_U(s)P_V(t) \frac{\partial P_{L,Q}(s,t)}{\partial s} P_{X,Y}(s,t) \\ & + a_3P_{L,Q}(s,t) \frac{\partial P_{X,Y}(s,t)}{\partial s} + a_1P_U(s) \frac{\partial P_{X,Y}(s,t)}{\partial s} \\ & - a_1a_3P_U(s)P_{L,Q}(s,t) \frac{\partial P_{X,Y}(s,t)}{\partial s} + a_0P_U(s)P_V(t)P_{L,Q}(s,t) \frac{\partial P_{X,Y}(s,t)}{\partial s} \\ & - a_0a_3P_U(s)P_V(t)P_{L,Q}^2(s,t) \frac{\partial P_{X,Y}(s,t)}{\partial s} \\ & - a_0a_1P'_U(s)P_V(t)P_{L,Q}(s,t) \frac{\partial P_{X,Y}(s,t)}{\partial s} \\ & + a_0a_1a_3P'_U(s)P_V(t)P_{L,Q}^2(s,t) \frac{\partial P_{X,Y}(s,t)}{\partial s}. \end{aligned}$$

Equating the coefficients of $s^{x-1}t^y$ on both sides of above equation and collecting all items containing $g(x,y)$ to the left-hand side yields (3.20). Equation (3.21) follows by symmetry. \square

Although the formulas look formidable, they are actually easy to be translated into a computer program. In terms of computational intensity, a six-fold summation (three with respect to x and three to y) is needed to compute $g(x, y)$. Using similar rationale for determining the computational intensity for equation (3.11), because of the recursive nature of the formula, using equation (3.20), the number of floating point operations needed to compute $g(x, y)$ is of order $O(x^3y^3 \cdot xy) = O(x^4y^4)$.

Notice that in the trivariate Poisson case where $a_j = 0$ and $b_j = \lambda_j$, the formulas in Theorem 3.2 simplify dramatically, yielding for $x \geq 1, y \geq 0$,

$$g(x, y) = \frac{\lambda_1}{x} \sum_{u=1}^x u f_1(u) g(x-u, y) + \frac{\lambda_3}{x} \sum_{u=1}^x \sum_{v=0}^y u f_3(u, v) g(x-u, y-v) + \frac{\lambda_0}{x} \sum_{u=1}^x \sum_{v=0}^y \left[\sum_{i=0}^u \sum_{j=0}^v u f_1(i) f_2(j) f_3(u-i, v-j) \right] g(x-u, y-v), \quad (3.22)$$

and for $x \geq 0, y \geq 1$,

$$g(x, y) = \frac{\lambda_2}{y} \sum_{v=1}^y v f_2(v) g(x, y-v) + \frac{\lambda_3}{y} \sum_{v=1}^y \sum_{u=0}^x v f_3(u, v) g(x-u, y-v) + \frac{\lambda_0}{y} \sum_{v=1}^y \sum_{u=0}^x \left[\sum_{j=0}^v \sum_{i=0}^u v f_2(j) f_1(i) f_3(u-i, v-j) \right] g(x-u, y-v). \quad (3.23)$$

3.2.3 Model C

Let Θ be a random variable on the support $[\sigma_1, \sigma_2]$, where $0 \leq \sigma_1 < \sigma_2 \leq \infty$. Assume that Θ has a probability density function u that satisfies

$$\frac{d}{d\theta} \log u(\theta) = \frac{\sum_{i=0}^k a_i \theta^i}{\sum_{i=0}^k b_i \theta^i} \quad (3.24)$$

for suitable constants a_i and b_i , and

$$\sum_{i=0}^k b_i \theta^i u(\theta) \rightarrow 0, \text{ when } \theta \rightarrow \sigma_1, \sigma_2. \quad (3.25)$$

Many distributions also have density functions satisfying these properties. For example, Hesselager (1996b) introduced a generalized inverse Gaussian distribution,

which is denoted by $\text{GIG}(\mu, \beta, \alpha)$, and has density function

$$u(\theta) = \frac{\mu^{-\alpha}}{2K_\alpha(\mu/\beta)} \theta^{\alpha-1} \exp\left\{-\frac{\theta^2 + \mu^2}{2\beta\theta}\right\}, \text{ for } \theta > 0 \text{ and } \mu, \beta > 0, \alpha \in \mathbb{R}, \quad (3.26)$$

where $K_\alpha(\cdot)$ is the modified Bessel function of the third kind. This density function satisfies

$$\frac{d}{d\theta} \log u(\theta) = \frac{\mu^2 + 2\beta(\alpha - 1)\theta - \theta^2}{2\beta\theta^2} = \frac{\sum_{i=0}^2 a_i \theta^i}{\sum_{i=0}^2 b_i \theta^i},$$

where $a_0 = \mu^2$, $a_1 = 2\beta(\alpha - 1)$, $a_2 = -1$, and $b_0 = b_1 = 0$, $b_2 = 2\beta$.

Another commonly used distribution that satisfies the above properties is the Gamma distribution, as we mentioned in Section 2.2.3. Note that the above conditions are satisfied by any finite mixture of Erlang distributions whose density function is

$$u(\theta) = \sum_{k=1}^m q_k e_k(s), \quad m = 1, 2, \dots,$$

where $e_k(s)$ is the gamma density function with shape parameter k and a constant scale parameter, and the set of mixing weights $\{q_k: k = 1, 2, \dots, m\}$ are nonnegative and sum up to one. Under the choice of a gamma mixing distribution, a simpler recursion than the recursion of Theorem 3.3 is presented in Subsection 7.4 of Sundt and Vernic (2004) and Section 20.3 of Sundt and Vernic (2009).

We assume that conditional on $\Theta = \theta$, N_1 , N_2 and N_3 are independent and follow Poisson distributions with parameters $\theta\lambda_1$, $\theta\lambda_2$ and $\theta\lambda_3$ respectively. Then the PGF of (N_1, N_2, N_3) is given by

$$P_{N_1, N_2, N_3}(s, t, u) = \int_{\sigma_1}^{\sigma_2} e^{\theta[\lambda_1(s-1) + \lambda_2(t-1) + \lambda_3(u-1)]} u(\theta) d\theta, \quad (3.27)$$

and the PGF of (X, Y) is given by

$$P_{X, Y}(s, t) = \int_{\sigma_1}^{\sigma_2} e^{\theta[\lambda_1(P_U(s)-1) + \lambda_2(P_V(t)-1) + \lambda_3(P_{L, Q}(s, t)-1)]} u(\theta) d\theta \quad (3.28)$$

$$= \int_{\sigma_1}^{\sigma_2} P_{X, Y|\theta}(s, t) u(\theta) d\theta, \quad (3.29)$$

where

$$P_{X, Y|\theta}(s, t) = \mathbb{E} [s^X t^Y | \Theta = \theta] = e^{\theta[\lambda_1(P_U(s)-1) + \lambda_2(P_V(t)-1) + \lambda_3(P_{L, Q}(s, t)-1)]} \quad (3.30)$$

is the conditional PGF of (X, Y) given $\Theta = \theta$.

To set up the recursions for the probability function of (X, Y) , we need an auxiliary function

$$h_i(x, y) = \int_{\sigma_1}^{\sigma_2} \theta^i g_\theta(x, y) u(\theta) d\theta, \quad (3.31)$$

where $g_\theta(x, y)$ is the conditional probability of (X, Y) given θ . Notice that our target quantity is $g(x, y) = h_0(x, y)$.

Theorem 3.3. *Under the conditions in model C, the following starting value and recursive formulas hold:*

$$h_i(0, 0) = \int_{\sigma_1}^{\sigma_2} \theta^i e^{-\lambda[1-\tilde{f}(0)]\theta} u(\theta) d\theta, \quad (3.32)$$

where $\lambda = \lambda_1 + \lambda_2 + \lambda_3$ and $\tilde{f}(0) = \frac{\lambda_1 f_1(0) + \lambda_2 f_2(0) + \lambda_3 f_3(0,0)}{\lambda}$.

For $i = 0, 1, \dots, k-1$ and $x \geq 1, y \geq 0$,

$$h_i(x, y) = \frac{\lambda_1}{x} \sum_{u=1}^x u f_1(u) h_{i+1}(x-u, y) + \frac{\lambda_3}{x} \sum_{u=1}^x \sum_{v=0}^y u f_3(u, v) h_{i+1}(x-u, y-v). \quad (3.33)$$

For $i = 0, 1, \dots, k-1$ and $x \geq 0, y \geq 1$,

$$h_i(x, y) = \frac{\lambda_2}{y} \sum_{v=1}^y v f_2(v) h_{i+1}(x, y-v) + \frac{\lambda_3}{y} \sum_{v=1}^y \sum_{u=0}^x v f_3(u, v) h_{i+1}(x-u, y-v). \quad (3.34)$$

For $x \geq 1, y \geq 0$ or $x \geq 0, y \geq 1$,

$$\begin{aligned} c_k h_k(x, y) = & \lambda_1 \sum_{u=1}^x f_1(u) \sum_{i=0}^k b_i h_i(x-u, y) + \lambda_2 \sum_{v=1}^y f_2(v) \sum_{i=0}^k b_i h_i(x, y-v) \\ & + \lambda_3 \sum_{v=1}^y f_3(0, v) \sum_{i=0}^k b_i h_i(x, y-v) + \lambda_3 \sum_{u=1}^x \sum_{v=0}^y f_3(u, v) \sum_{i=0}^k b_i h_i(x-u, y-v) \\ & + \sum_{i=0}^{k-1} [(i+1)b_{i+1} - c_i] h_i(x, y), \end{aligned} \quad (3.35)$$

where $c_i = \lambda [1 - \tilde{f}(0)] b_i - a_i$.

Proof. Obviously,

$$g_\theta(0, 0) = e^{\theta[\lambda_1(f_1(0)-1)+\lambda_2(f_2(0)-1)+\lambda_3(f_3(0,0)-1)]},$$

that is

$$g_\theta(0, 0) = e^{-\lambda[1-\tilde{f}(0)]\theta}. \quad (3.36)$$

Multiplying both sides of (3.36) by $\theta^i u(\theta)$ and integrating over θ yields (3.32).

Differentiating (3.30) with respect to s results in

$$\frac{\partial P_{X,Y|\theta}(s, t)}{\partial s} = \theta\lambda_1 P_{X,Y|\theta}(s, t)P'_U(s) + \theta\lambda_3 P_{X,Y|\theta}(s, t)\frac{\partial P_{L,Q}(s, t)}{\partial s}. \quad (3.37)$$

Expanding both sides of (3.37) polynomially and then comparing the coefficients of $s^{x-1}t^y$ for $x \geq 1$, we obtain

$$xg_\theta(x, y) = \theta\lambda_1 \sum_{u=1}^x u f_1(u)g_\theta(x-u, y) + \theta\lambda_3 \sum_{u=1}^x \sum_{v=0}^y u f_3(u, v)g_\theta(x-u, y-v). \quad (3.38)$$

Multiplying both sides of (3.38) by $\theta^i u(\theta)$, integrating over θ , and then dividing by x , one obtains (3.33). Equation (3.34) follows by symmetry.

Rewriting (3.24) as $u(\theta) \sum_{i=0}^k a_i \theta^i = u'(\theta) \sum_{i=0}^k b_i \theta^i$, multiplying both sides by $P_{X,Y|\theta}(s, t)$, and then integrating over θ , we obtain

$$\begin{aligned} \int_{\sigma_1}^{\sigma_2} \sum_{i=0}^k a_i \theta^i P_{X,Y|\theta}(s, t) u(\theta) d\theta &= \int_{\sigma_1}^{\sigma_2} \sum_{i=0}^k b_i \theta^i P_{X,Y|\theta}(s, t) u'(\theta) d\theta \\ &= \left[\sum_{i=0}^k b_i \theta^i P_{X,Y|\theta}(s, t) u(\theta) \right] \Big|_{\sigma_1}^{\sigma_2} - \int_{\sigma_1}^{\sigma_2} \sum_{i=1}^k i \theta^{i-1} b_i u(\theta) P_{X,Y|\theta}(s, t) d\theta \\ &\quad - \int_{\sigma_1}^{\sigma_2} \sum_{i=0}^k b_i \theta^i u(\theta) \frac{dP_{X,Y|\theta}(s, t)}{d\theta} d\theta. \end{aligned} \quad (3.39)$$

Following from (3.30),

$$\frac{dP_{X,Y|\theta}(s, t)}{d\theta} = P_{X,Y|\theta}(s, t) [\lambda_1 P_U(s) + \lambda_2 P_V(t) + \lambda_3 P_{L,Q}(s, t) - \lambda], \quad (3.40)$$

applying (3.25) and (3.40), equation (3.39) gives

$$\int_{\sigma_1}^{\sigma_2} \sum_{i=0}^k a_i \theta^i P_{X,Y|\theta}(s, t) u(\theta) d\theta = - \int_{\sigma_1}^{\sigma_2} \sum_{i=0}^{k-1} (i+1) \theta^i b_{i+1} u(\theta) P_{X,Y|\theta}(s, t) d\theta$$

$$- \int_{\sigma_1}^{\sigma_2} \sum_{i=0}^k b_i \theta^i u(\theta) P_{X,Y|\theta}(s,t) [\lambda_1 P_U(s) + \lambda_2 P_V(t) + \lambda_3 P_{L,Q}(s,t) - \lambda] d\theta. \quad (3.41)$$

After comparing the coefficients of $s^x t^y$ on both sides of (3.41) and rearranging terms, we obtain equation (3.35). \square

Remark 1: Notice that (3.28) can be written as

$$P_{X,Y}(s,t) = \int_{\sigma_1}^{\sigma_2} e^{\lambda\theta \left[\frac{1}{\lambda} (\lambda_1 P_U(s) + \lambda_2 P_V(t) + \lambda_3 P_{L,Q}(s,t)) - 1 \right]} u(\theta) d\theta, \quad (3.42)$$

which shows that model C can be represented as

$$(X, Y) = \sum_{i=1}^N (C_i, D_i), \quad (3.43)$$

where conditional on $\Theta = \theta$, N follows the Poisson($\lambda\theta$) distribution, and the pairs (C_i, D_i) are independent and identically distributed with probability function

$$f(u, v) = \frac{1}{\lambda} (I(v=0)\lambda_1 f_1(u) + I(u=0)\lambda_2 f_2(v) + \lambda_3 f_3(u, v)). \quad (3.44)$$

For this situation, Sundt and Vernic (2004, Section 4A and 4C) and Sundt and Vernic (2009, Section 20.4.1 and 20.4.2) showed that

$$h_i(0, 0) = \int_{\sigma_1}^{\sigma_2} \theta^i e^{-\lambda\theta(1-f(0,0))} u(\theta) d\theta, \quad i \geq 0, \quad (3.45)$$

$$h_i(x, y) = \frac{\lambda}{x} \sum_{u=1}^x \sum_{v=0}^y u f(u, v) h_{i+1}(x-u, y-v), \quad x \geq 1, y \geq 0, i \geq 0, \quad (3.46)$$

$$h_i(x, y) = \frac{\lambda}{y} \sum_{v=1}^y \sum_{u=0}^x v f(u, v) h_{i+1}(x-u, y-v), \quad x \geq 0, y \geq 1, i \geq 0, \quad (3.47)$$

$$\begin{aligned} c_k h_k(x, y) &= \lambda \sum_{u=1}^x \sum_{v=0}^y f(u, v) \sum_{i=0}^k b_i h_i(x-u, y-v) + \lambda \sum_{v=1}^y f(0, v) \sum_{i=0}^k b_i h_i(x, y-v) \\ &\quad - \sum_{i=0}^{k-1} [c_i - (i+1)b_{i+1}] h_i(x, y), \quad x \geq 1, y \geq 0 \text{ or } x \geq 0, y \geq 1. \end{aligned} \quad (3.48)$$

Inserting (3.44) into formulas (3.45), (3.46), (3.47) and (3.48) yields (3.32), (3.33), (3.34) and (3.35) respectively.

Remark 2: Notice that, as (3.33) and (3.34) actually hold for all $i \geq 0$, one can just use these two equations to compute the joint probabilities, but when the values of x or y get large, the computation intensity increases tremendously. Using equation (3.35) allows us to limit i to the maximum value k . This reduces computation time dramatically. Similar to Model A, the number of floating point operations needed to compute $g(x, y)$ is of order $O(x^2y^2)$, because a double summation (one with respect to x and one to y) is involved.

3.3 Fast Fourier transforms

In this section, we discuss the use of the bivariate FFT and its IFFT to compute the joint probability functions of the aggregate claims.

For our application, the characteristic function of (X, Y) is given by

$$\begin{aligned}\phi_{X,Y}(s, t) &= \mathbb{E} \left[e^{is(\sum_{i=1}^{N_1} U_i + \sum_{k=1}^{N_3} L_k) + it(\sum_{j=1}^{N_2} V_j + \sum_{k=1}^{N_3} Q_k)} \right] \\ &= P_{N_1, N_2, N_3}(\phi_U(s), \phi_V(t), \phi_{L,Q}(s, t)),\end{aligned}\tag{3.49}$$

where $\phi_U(s)$, $\phi_V(t)$ and $\phi_{L,Q}(s, t)$ are the characteristic functions of claim sizes U , V and (L, Q) . Then the distribution of (X, Y) may be obtained by the following procedure 1.

Procedure 1:

1. Set truncation points for the claim sizes U , V , and (L, Q) to r , w and (r, w) respectively to obtain the truncated claim size distributions

$$\mathbf{f}_1 = \{f_1(0), f_1(1), \dots, f_1(r-1)\}, \quad \mathbf{f}_2 = \{f_2(0), f_2(1), \dots, f_2(w-1)\},$$

and

$$\mathbb{f}_3 = \begin{bmatrix} f_3(0,0) & \cdots & f_3(0,w-1) \\ f_3(1,0) & \cdots & f_3(1,w-1) \\ \vdots & \ddots & \vdots \\ f_3(r-1,0) & \cdots & f_3(r-1,w-1) \end{bmatrix},$$

where $r = 2^{r_1}$ and $w = 2^{r_2}$ for some positive integers r_1 and r_2 . Notice that if the claim size distributions have bounded supports, the vector or the matrix of probabilities can be appropriately padded with zeros in order to force r or w to be powers of two.

2. Apply one-dimensional FFT to \mathbf{f}_1 and \mathbf{f}_2 to obtain two vectors $\tilde{\mathbf{f}}_1$ and $\tilde{\mathbf{f}}_2$. And apply two-dimensional FFT to the claim size matrix \mathbb{f}_3 to get

$$\tilde{\mathbb{f}}_3 = \begin{bmatrix} \tilde{f}_3(0,0) & \cdots & \tilde{f}_3(0,w-1) \\ \tilde{f}_3(1,0) & \cdots & \tilde{f}_3(1,w-1) \\ \vdots & \ddots & \vdots \\ \tilde{f}_3(r-1,0) & \cdots & \tilde{f}_3(r-1,w-1) \end{bmatrix}.$$

3. Use formula (3.49) to obtain the matrix $\tilde{\phi}_{X,Y}$ with the ij th element $P_{N_1, N_2, N_3}(\tilde{f}_1(i), \tilde{f}_2(j), \tilde{f}_3(i, j))$ for $i = 0, \dots, r-1$ and $j = 0, \dots, w-1$.
4. Apply the IFFT (2.9) to $\tilde{\phi}_{X,Y}$ to obtain the probability function of (X, Y) .

As mentioned in Section 2.3, Grübel and Hermesmeier (1999) introduced a more efficient way known as *exponential tilting* to alleviate the problem associated with the aliasing errors. We next show that the tilting method may be applied to the multivariate model presented in this chapter by properly choosing the tilting parameters. To do this, we need some notations as defined in Grübel and Hermesmeier (1999). First, we define the convolution product $c = a * b$ of two sequences $a = (a_r)_{r \in \mathbb{N}_0}$, $b = (b_r)_{r \in \mathbb{N}_0}$ as

$$c_r = \sum_{i=0}^r a_i b_{r-i} \quad \text{for all } r \in \mathbb{N}_0.$$

Then because for all $r \geq 0$,

$$\left((E_\theta a) * (E_\theta b) \right)_r = \sum_{i=0}^r (E_\theta a)_i (E_\theta b)_{r-i} = e^{-\theta r} \sum_{i=0}^r a_i b_{r-i} = \left(E_\theta (a * b) \right)_r,$$

we have $E_\theta(a) * E_\theta(b) = E_\theta(a * b)$.

Similarly, define the convolution product $\mathbb{C} = \mathbb{A} * \mathbb{B}$ of two matrices $\mathbb{A} = (a_{ij})_{i,j \in \mathbb{N}_0}$, $\mathbb{B} = (b_{ij})_{i,j \in \mathbb{N}_0}$ by

$$c_{ij} = \sum_{r=0}^i \sum_{w=0}^j a_{r,w} b_{i-r,j-w} \quad \text{for all } i, j \in \mathbb{N}_0,$$

and then define the tilting operator E_{θ_1, θ_2} by

$$(E_{\theta_1, \theta_2} \mathbb{A})_{ij} = e^{-\theta_1 i} e^{-\theta_2 j} a_{ij} \quad \text{and} \quad (E_{\theta_1, \theta_2} \mathbb{B})_{ij} = e^{-\theta_1 i} e^{-\theta_2 j} b_{ij} \quad \text{for all } i, j \in \mathbb{N}_0.$$

Then we have

$$\begin{aligned} \left((E_{\theta_1, \theta_2} \mathbb{A}) * (E_{\theta_1, \theta_2} \mathbb{B}) \right)_{ij} &= \sum_{r=0}^i \sum_{w=0}^j (E_{\theta_1, \theta_2} \mathbb{A})_{r,w} (E_{\theta_1, \theta_2} \mathbb{B})_{i-r,j-w} \\ &= e^{-\theta_1 i} e^{-\theta_2 j} \sum_{r=0}^i \sum_{w=0}^j a_{r,w} b_{i-r,j-w} \\ &= \left(E_{\theta_1, \theta_2} (\mathbb{A} * \mathbb{B}) \right)_{ij}, \end{aligned}$$

which means $(E_{\theta_1, \theta_2} \mathbb{A}) * (E_{\theta_1, \theta_2} \mathbb{B}) = E_{\theta_1, \theta_2} (\mathbb{A} * \mathbb{B})$.

For the multivariate model in this chapter, we can express the joint probability as

$$g(x, y) = \sum_{n_1 \geq 0} \sum_{n_2 \geq 0} \sum_{n_3 \geq 0} p(n_1, n_2, n_3) \sum_{i=0}^x \sum_{j=0}^y f_1^{*n_1}(i) f_2^{*n_2}(j) f_3^{*n_3}(x-i, y-j), \quad (3.50)$$

where f^{*k} denotes the k th convolution of f . If we tilt f_1 and f_2 with tilting parameters θ_1 and θ_2 to get $E_{\theta_1} f_1$ and $E_{\theta_2} f_2$, and tilt f_3 with tilting parameters θ_3, θ_4 to get $E_{\theta_3, \theta_4} f_3$, and denote the resulting aggregate claims distribution by $\bar{g}(x, y)$, then

$$\bar{g}(x, y) = \sum_{n_1 \geq 0} \sum_{n_2 \geq 0} \sum_{n_3 \geq 0} p(n_1, n_2, n_3) \times$$

$$\begin{aligned}
& \sum_{i=0}^x \sum_{j=0}^y (E_{\theta_1} f_1)^{*n_1}(i) (E_{\theta_2} f_2)^{*n_2}(j) (E_{\theta_3, \theta_4} f_3)^{*n_3}(x-i, y-j) \\
&= \sum_{n_1 \geq 0} \sum_{n_2 \geq 0} \sum_{n_3 \geq 0} p(n_1, n_2, n_3) \times \\
& \quad \sum_{i=0}^x \sum_{j=0}^y e^{-\theta_1 i} f_1^{*n_1}(i) e^{-\theta_2 j} f_2^{*n_2}(j) e^{-\theta_3(x-i)} e^{-\theta_4(y-j)} f_3^{*n_3}(x-i, y-j). \quad (3.51)
\end{aligned}$$

It is obvious that to make the tilting commute with convolutions, one needs to set $\theta_3 = \theta_1$ and $\theta_4 = \theta_2$. This yields $\bar{g}(x, y) = e^{-\theta_1 x} e^{-\theta_2 y} g(x, y)$ and so $g(x, y) = E_{-\theta_1, -\theta_2} \bar{g}(x, y)$.

Therefore, when applying exponential tilting, we can use the following steps to compute the joint distribution of (X, Y) .

Procedure 2:

1. Same as step 1 in procedure 1.
2. Tilt the two sequences $\mathbf{f}_1, \mathbf{f}_2$ and the matrix \mathbb{f}_3 ,

$$\mathbf{f}_1 \mapsto E_{\theta_1} \mathbf{f}_1 = [e^{-\theta_1 j} f_1(j)]_{j=0,1,\dots,r-1},$$

$$\mathbf{f}_2 \mapsto E_{\theta_2} \mathbf{f}_2 = [e^{-\theta_2 j} f_2(j)]_{j=0,1,\dots,w-1},$$

$$\mathbb{f}_3 \mapsto E_{\theta_1, \theta_2} \mathbb{f}_3 = [e^{-\theta_1 i} e^{-\theta_2 j} f_3(i, j)]_{i=0,1,\dots,r-1; j=0,1,\dots,w-1}.$$

To avoid the problems of underflow or overflow, Grübel and Hermesmeier (1999) suggested choosing $\theta = 20/r$ in univariate case. For our bivariate case, we choose $\theta_1 = 10/r$ and $\theta_2 = 10/w$.

3. Apply FFT to the two sequences and the matrix, resulting in two sequences $\widetilde{E_{\theta_1} \mathbf{f}_1}$, $\widetilde{E_{\theta_2} \mathbf{f}_2}$ and one matrix $\widetilde{E_{\theta_1, \theta_2} \mathbb{f}_3}$.
4. Use formulas (3.8), (3.17), (3.27), and (3.49) to obtain the tilted characteristic function of (X, Y) , resulting in the $r \times w$ matrix with the ij th element $P_{N_1, N_2, N_3} \left(\widetilde{E_{\theta_1} \mathbf{f}_1}(i), \widetilde{E_{\theta_2} \mathbf{f}_2}(j), \widetilde{E_{\theta_1, \theta_2} \mathbb{f}_3}(i, j) \right)$ for $i = 0, \dots, r-1$ and $j = 0, \dots, w-1$.

5. Apply IFFT to the above matrix. Untilting the obtained matrix by $E_{-\theta_1, -\theta_2}$ yields $g_{X,Y}(i, j)$.

As introduced in Klugman *et al.* (2008, p. 242), the number of floating point operations needed for FFT method on a vector with length n is of order $O(n \log_2 n)$. Thus for the computation of $g(x, y)$, the number of floating point operations is of order $O(xy \log_2 y + yx \log_2 x) = O(xy(\log_2 x + \log_2 y))$.

3.4 Numerical examples

In this section, we present numerical examples for models introduced in Section 3.2 to compare the accuracy and speed of the recursive and FFT methods for computing the joint distribution of the multivariate compound distribution. All calculations are done using the free statistical software R. We assume that the claim sizes U and V follow Pareto distributions. Let the claim sizes (L, Q) follow a bivariate Pareto distribution introduced by Lindley and Singpurwalla (1986). The bivariate Pareto distribution, denoted by $\text{BiPareto}(\gamma_1, \gamma_2, \beta)$, has density function

$$f(l, q) = \frac{(\beta + 1)\beta}{\gamma_1 \gamma_2} \left(\frac{l}{\gamma_1} + \frac{q}{\gamma_2} + 1 \right)^{-(\beta+2)}, \quad l, q > 0 \text{ and } \gamma_1, \gamma_2, \beta > 0. \quad (3.52)$$

To apply the recursive and FFT methods, the distributions of U , V and (L, Q) are discretized using the standard rounding method (Klugman *et al.* 2008, p. 233). R function `fft()` is used to carry out the Fourier transforms and their inverses.

3.4.1 An example for model A

Let $K \sim \text{Poisson}(8)$, conditional on $K = k$, the distribution of (N_1, N_2, N_3) is trinomial with parameters k and $(p_1 = 0.2, p_2 = 0.3, p_3 = 0.5)$. Let $U \sim \text{Pareto}(3, 5)$, $V \sim \text{Pareto}(4, 3)$, and $(L, Q) \sim \text{BiPareto}(2, 4, 3)$. The joint probability mass function $g(x, y)$ at some selected points are computed using the recursive method, the FFT

method with truncation points $r = w = 2^8$, the FFT method with truncation points $r = w = 2^9$, and the FFT method with truncation points $r = w = 2^9$ and tilting parameters $\theta_1 = \theta_2 = 10/r$.

To illustrate the accuracy of different methods, we list in Table 3.1 some actual probability values calculated with different methods and in Table 3.2 the number of significant decimal digits agreed. Table 3.3 compares the computation time of the different methods on a personal computer with Intel Core2 Quad CPU Q8200@2.33G and 8GB memory. Since the recursion method obtains the exact values of the compound distribution if the errors from the floating point representation are ignored, the differences between the FFT and the recursion methods are essentially due to aliasing errors. From the results we can see that, as the truncation points increase, the alias errors decrease. Moreover, aliasing errors are reduced significantly by tilting. As reported in Embrechts and Frei (2009), the recursive method appears to spend more CPU time than the FFT method. We also note that each of the computation times reported here is a result of average of those for 100 runs of the corresponding computer program.

Quantity	Recursion	FFT(2^8)	FFT(2^9)	FFT(2^9 with tilting)
$g(10, 10)$	4.7603012e-05	4.8705431e-05	4.7639162e-05	4.7603013e-05
$g(20, 30)$	7.9430590e-05	8.1122257e-05	7.9499398e-05	7.9430593e-05
$g(30, 30)$	7.2078212e-05	7.4164653e-05	7.2154758e-05	7.2078215e-05

Table 3.1: Model A: Some values of $g(x, y)$.

Quantity	FFT(2^8)	FFT(2^9)	FFT(2^9 with tilting)
$g(10, 10)$	1	3	7
$g(20, 30)$	0	3	7
$g(30, 30)$	1	2	7

Table 3.2: Model A: Number of significant decimal digits agreed for $g(x, y)$ computed with recursive and FFT methods.

To illustrate the dependency between X and Y , we plot in Figure 3.1 a dependency

Quantity	Recursion	FFT(2^8)	FFT(2^9)	FFT(2^9 with tilting)
$g(10, 10)$	0.0915	0.0523	0.3660	0.4912
$g(20, 30)$	1.8157	0.0523	0.3660	0.4912
$g(30, 30)$	4.0479	0.0523	0.3660	0.4912

Table 3.3: Model A: Computation times of recursive and FFT methods (in seconds).

measure introduced by Coles, Heffernan and Tawn (2000) defined by $\chi(p) = \Pr(X > \pi_p(X)|Y > \pi_p(Y))$, where $\pi_p(X)$ and $\pi_p(Y)$ represent the $100p\%$ percentile of X and Y respectively. The measure $\chi(p) > 1 - p$ indicates positive dependencies. This figure clearly shows the heavy dependency between the two types of risks.

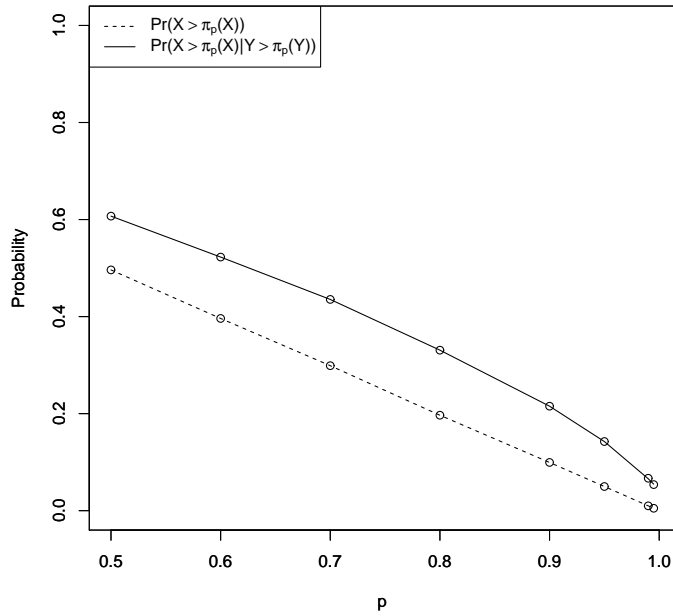


Fig. 3.1: Model A: Plot of a measure of dependence between X and Y .

To demonstrate the effect of dependency on the risk of aggregate claims $X + Y$, we plot the cumulative distribution functions (CDFs) for $X + Y$ and $X' + Y'$ where X' and Y' have the same marginal distribution with X and Y but are independent. The results are showed in Figure 3.2, it can be seen that the dependency between X

and Y results in a heavier tail for the aggregate claims $X + Y$.

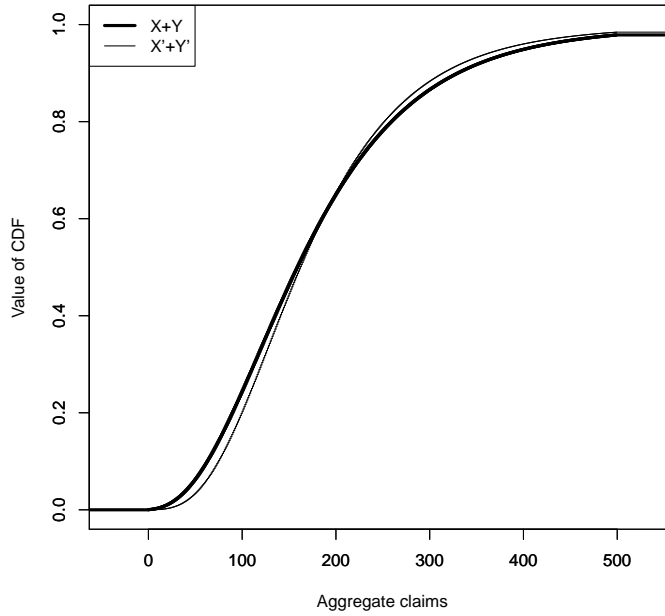


Fig. 3.2: Model A: Plot of the CDFs of aggregate claims $X + Y$ with and without dependence between X and Y .

3.4.2 An example for model B

Let $Z_0 \sim \text{Poisson}(3)$, $Z_1 \sim \text{Poisson}(2)$, $Z_2 \sim \text{Poisson}(4)$, $Z_3 \sim \text{Poisson}(5)$, $U \sim \text{Pareto}(3, 5)$, $V \sim \text{Pareto}(4, 3)$, and $(L, Q) \sim \text{BiPareto}(2, 4, 3)$. For this example, we provide three tables 3.4, 3.5, 3.6 and two figures 3.3, 3.4. They provide similar information to those for example A.

Quantity	Recursion	FFT(2^8)	FFT(2^9)	FFT(2^9 with tilting)
$g(10, 10)$	4.8156806e-07	9.2478228e-06	6.1918132e-07	4.8156811e-07
$g(20, 30)$	2.0814650e-06	9.5419547e-06	2.1958211e-06	2.0814654e-06
$g(30, 30)$	2.3532538e-06	9.4440615e-06	2.4617731e-06	2.3532543e-06

Table 3.4: Model B: Some values of $g(x, y)$.

Quantity	FFT(2^8)	FFT(2^9)	FFT(2^9 with tilting)
$g(10, 10)$	0	0	6
$g(20, 30)$	0	1	7
$g(30, 30)$	0	1	6

Table 3.5: Model B: Number of significant decimal digits agreed for $g(x, y)$ computed with recursive and FFT methods.

Quantity	Recursion	FFT(2^8)	FFT(2^9)	FFT(2^9 with tilting)
$g(10, 10)$	0.9787	0.0581	0.3768	0.4875
$g(20, 30)$	82.4113	0.0581	0.3768	0.4875
$g(30, 30)$	254.9361	0.0581	0.3768	0.4875

Table 3.6: Model B: Computation times of recursive and FFT methods (in seconds).

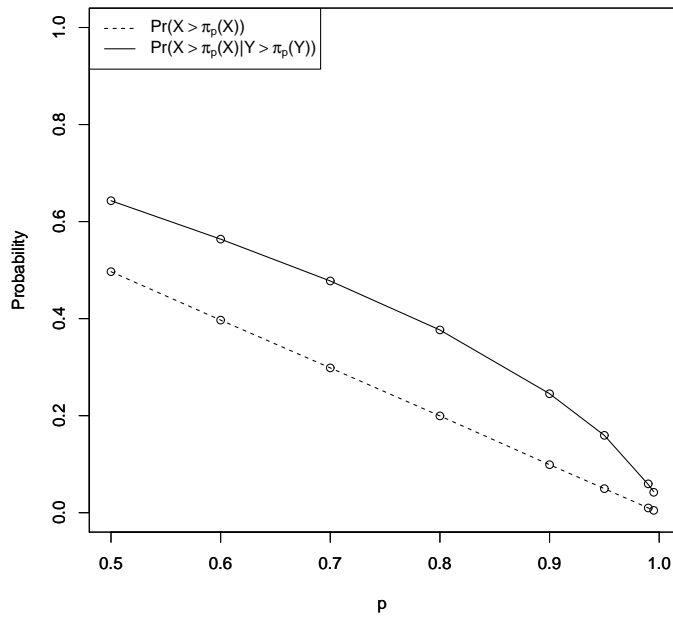


Fig. 3.3: Model B: Plot of a measure of dependence between X and Y .

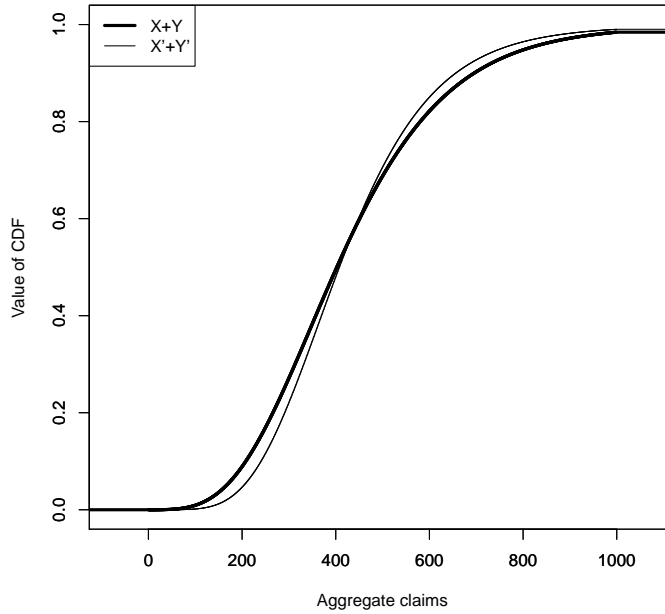


Fig. 3.4: Model B: Plot of the CDFs of aggregate claims $X + Y$ with and without dependence between X and Y .

3.4.3 An example for model C

Let: $\Theta \sim \text{GIG}(2, 1, 2)$, $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 0.5$, $U \sim \text{Pareto}(3, 5)$, $V \sim \text{Pareto}(4, 3)$, and $(L, Q) \sim \text{BiPareto}(2, 4, 3)$. For this model, the results are shown in tables 3.7, 3.8, 3.9 and figures 3.5, 3.6.

Quantity	Recursion	FFT(2^8)	FFT(2^9)	FFT(2^9 with tilting)
$g(20, 20)$	3.4587685e-05	3.9388355e-05	3.4766397e-05	3.4587688e-05
$g(25, 30)$	3.5972251e-05	4.0773325e-05	3.6160165e-05	3.5972255e-05
$g(40, 40)$	3.2485592e-05	3.7264073e-05	3.2674854e-05	3.2485597e-05

Table 3.7: Model C: Some values of $g(x, y)$.

Quantity	FFT(2^8)	FFT(2^9)	FFT(2^9 with tilting)
$g(20, 20)$	1	2	7
$g(25, 30)$	0	1	7
$g(40, 40)$	1	2	7

Table 3.8: Model C: Number of significant decimal digits agreed for $g(x, y)$ computed with recursive and FFT methods.

Quantity	Recursion	FFT(2^8)	FFT(2^9)	FFT(2^9 with tilting)
$g(20, 20)$	1.9588	2.0513	8.2618	8.3115
$g(25, 30)$	6.1082	2.0513	8.2618	8.3115
$g(40, 40)$	25.1801	2.0513	8.2618	8.3115

Table 3.9: Model C: Computation times of recursive and FFT methods (in seconds).

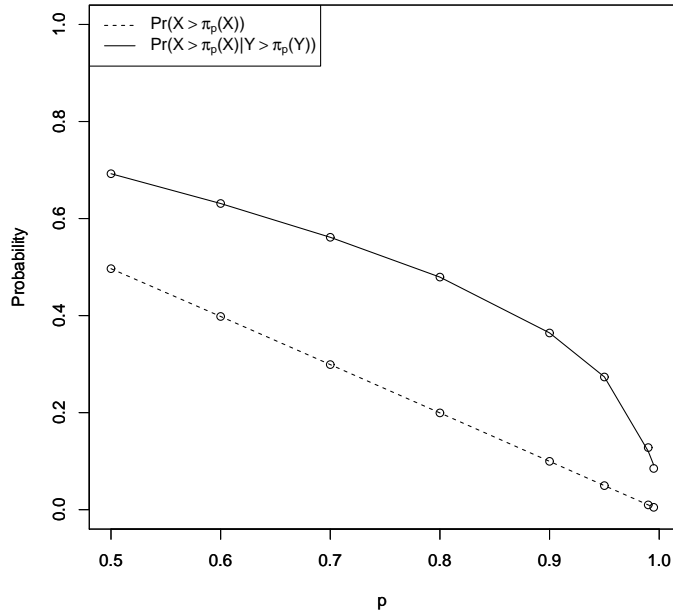


Fig. 3.5: Model C: Plot of a measure of dependence between X and Y .

Results from model B and model C also show that the alias errors of the FFT method decrease with higher truncation points and the errors are reduced through tilting.

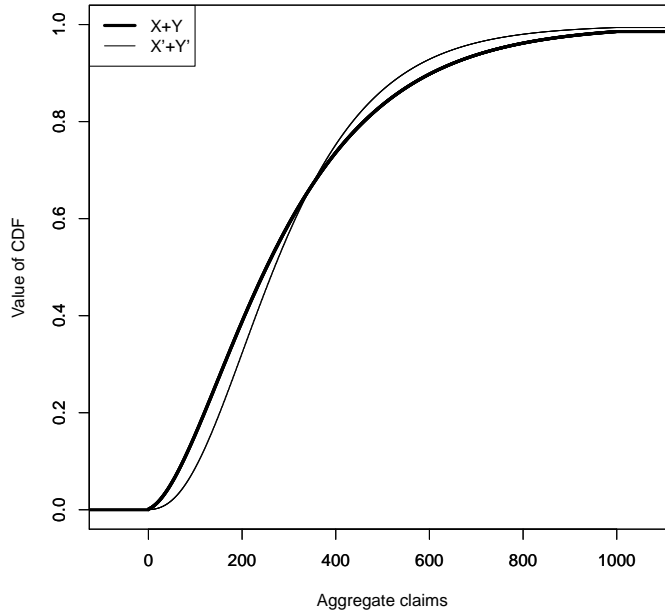


Fig. 3.6: Model C: Plot of the CDFs of aggregate claims $X + Y$ with and without dependence between X and Y .

3.5 Conclusions

We proposed a new bivariate aggregate claims model and derived recursive formulas for calculating the joint probabilities. The recursive formulas give exact joint probabilities if the floating point representation error is ignored. The correctness of the recursive formulas was verified by comparing the results from both recursive and FFT methods. As showed in Chapter 2, the aliasing errors suffered by the FFT method can be reduced effectively by exponential tilting. Comparing the numbers of the floating point operations required by the recursive and the FFT methods for the discussed three models, we have $O(x^2y^2)$ vs. $O(xy(\log_2x + \log_2y))$ in model A and model C, and $O(x^4y^4)$ vs. $O(xy(\log_2x + \log_2y))$ in model B.

Chapter 4

Moment-based density

approximations of aggregate claims

4.1 Introduction

In this chapter, we apply a moment-based method introduced by Provost (2005) to approximate the density functions of univariate and bivariate aggregate claims random variables. We provide some illustrative examples where the approximate distribution functions are compared with those obtained from the recursive method. The resulting approximations turned out to be quite accurate. Since the methodology only involves solving systems of linear equations, it is much less computer intensive than the recursive or FFT methods. It should be pointed out that the proposed approach can also be utilized in conjunction with observed data, in which case the sample moments would be used in lieu of the exact moments. To our knowledge, this constitutes, in the context of risk theory, the first attempt to approximate the density function of bivariate aggregate claims random vectors.

The chapter is organized as follows. Section 4.2 introduces the proposed moment-based density approximation method, explains how it can be applied to univariate aggregate claims and presents some numerical examples. Section 4.3 extends the

methodology to bivariate compound random variables and also includes illustrative numerical examples. Section 4.4 contains some concluding remarks.

4.2 Approximating the distribution of univariate aggregate claims

Provost (2005) proposed a unified approach to density approximation and showed that the resulting approximants are mathematically equivalent to those obtained by making use of orthogonal polynomials, such as the Legendre, Laguerre, Jacobi, and Hermite polynomials, and their associated weight functions. This conceptually simple semiparametric technique eliminates some of the complications associated with the use of orthogonal polynomials while yielding identical density approximants. For the purpose of approximating the distribution of univariate aggregate claims, the methodology can be stated as follows.

Let $f_S(s)$ be the density function of a continuous random variable S defined on $(0, \infty)$ and $g(s)$ be a $\text{Gamma}(\alpha, \theta)$ distributed base density function given by

$$g(s) = \frac{s^{\alpha-1} e^{-s/\theta}}{\theta^\alpha \Gamma(\alpha)}, \quad \text{for } s > 0 \text{ and } \alpha, \theta > 0, \quad (4.1)$$

where α is the shape parameter and θ is the scale parameter. Then, the exact density function of S can be approximated by

$$\tilde{f}_{S,t}(s) = g(s) \sum_{r=0}^t c_r s^r, \quad s > 0, \quad (4.2)$$

where t is a suitably selected positive integer and the c_r 's are real polynomial coefficients. Note that (4.2) gives a proper probability density function and the coefficients $c_r, r = 1, \dots, t$, do not have to be positive. As such, it belongs to a family of distributions that are generated by some combinations (more general than mixtures) of gamma random variables. Since the moments of the random variable S can usually be easily determined, one can obtain numerical values for the parameters α and θ in

$g(s)$ by matching the first two moments of the base gamma distribution to those of S , that is,

$$\alpha\theta = \mu_S(1) \quad \text{and} \quad \alpha(\alpha + 1)\theta^2 = \mu_S(2), \quad (4.3)$$

where $\mu_S(a)$ is the a^{th} moment of S , so that

$$\alpha = \frac{\mu_S^2(1)}{\mu_S(2) - \mu_S^2(1)} \quad \text{and} \quad \theta = \frac{\mu_S(2) - \mu_S^2(1)}{\mu_S^2(1)}. \quad (4.4)$$

Similarly, the coefficients c_r 's can be determined by solving the system of linear equations resulting from matching the first t moments obtained from $\tilde{f}_{S,t}(s)$ to those of S , that is,

$$\int_0^\infty s^a \tilde{f}_{S,t}(s) ds = \int_0^\infty s^a \left(g(s) \sum_{r=0}^t c_r s^r \right) ds = \mu_S(a), \quad a = 0, 1, \dots, t. \quad (4.5)$$

Equivalently, (4.5) can be written as

$$\sum_{r=0}^t c_r \int_0^\infty s^{a+r} g(s) ds = \sum_{r=0}^t c_r m_{a+r} = \mu_S(a), \quad a = 0, 1, \dots, t, \quad (4.6)$$

where $m_{a+r} = \theta^{a+r} \Gamma(\alpha + a + r) / \Gamma(\alpha) = \theta^{a+r} (\alpha + a + r - 1) \cdots \alpha$. Thus, one has

$$\begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_t \end{pmatrix} = \begin{pmatrix} m_0 & \cdots & m_t \\ m_1 & \cdots & m_{t+1} \\ \vdots & \ddots & \vdots \\ m_t & \cdots & m_{2t} \end{pmatrix}^{-1} \begin{pmatrix} \mu_S(0) \\ \mu_S(1) \\ \vdots \\ \mu_S(t) \end{pmatrix}. \quad (4.7)$$

To determine the degree of the adjustment, one may plot $\tilde{f}_{S,t}(s)$ for various values of t and select t such that no significant differences are observed between $\tilde{f}_{S,t}(s)$ and $\tilde{f}_{S,t+1}(s)$; alternatively, one could choose t such that

$$\int_0^\infty \left(\tilde{f}_{S,t}(s) - \tilde{f}_{S,t+\Delta t}(s) \right)^2 ds < \varepsilon,$$

where Δt is a positive integer increment such as 2 or 3 and ε is a certain predetermined tolerance level.

Essentially, this approximation technique involves the determination of the parameters of a gamma base density and a suitable polynomial adjustment. Some previously published results support the selection of a gamma distribution as an appropriate initial approximation. For example, Papush *et al.* (2001) made use of gamma, normal and lognormal random variables to approximate certain aggregate claims distributions under seven scenarios when no separate information on the claim numbers and sizes was available, and concluded that, in each case, the gamma distribution provides a much better fit than the normal or lognormal; moreover, Sundt (1982) showed that, under some special conditions, the distribution of the aggregate claims behaves asymptotically as a gamma-type distribution in its tail when the distribution of the number of claims is negative binomial.

Furthermore, Tijms (1994, p. 163-164) showed that any positive continuous distribution can be arbitrarily closely approximated by a mixture of gamma distributions with integer shape parameters (sometimes called Erlang distributions). That is, the density function of a positive continuous distribution can be expressed as

$$f_S(s) = \sum_{k=1}^{\infty} q_k h_k(s), \quad (4.8)$$

where $h_k(s)$ is the gamma(k, θ) density function and the set of mixing weights $\{q_k: k = 1, 2, \dots\}$ are nonnegative and sum up to one. However, in practice, it is difficult to determine the values of the mixing weights and the scale parameter θ so that the distribution of the mixture be sufficiently close to the target distribution. The approximation method proposed in this chapter could be viewed as a variation of Tijms's theorem involving a finite mixture of gamma densities wherein the gamma density parameters and polynomial coefficients are obtained by applying a moment matching approach.

4.2.1 Numerical examples

In this section, we approximate the density function of the aggregate claims random variable X in (1.1) under two sets of distributional assumptions. Denoting by $P_N(z) = \mathbb{E}(z^N)$ the probability generating function of N and, $M_U(k) = \mathbb{E}(e^{kU})$ the moment generating function of U , the moment generating function of X is given by

$$M_X(k) = P_N(M_U(k)). \quad (4.9)$$

Then, the a^{th} moment of X is obtained by differentiation as follows:

$$\mu_X(a) = \left. \frac{d^a M_X(k)}{dk^a} \right|_{k=0}. \quad (4.10)$$

Since our approximation method applies to continuous distributions while the distribution of X contains a probability mass at 0, we shall approximate the distribution of S , which is, in fact, that of X given $X > 0$. The density function of S is then

$$f_S(s) = \frac{f_X(s)}{\tau}, \quad s > 0, \quad (4.11)$$

where $\tau = 1 - \Pr(X = 0) = 1 - \Pr(N = 0)$. In addition, $\mu_S(0) = \mu_X(0) = 1$, and

$$\mu_S(a) = \frac{\mu_X(a)}{\tau}, \quad a = 1, 2, \dots \quad (4.12)$$

The cumulative distribution function of X , that is, $F_X(x) = \Pr(X \leq x)$, $x \geq 0$, can thus be expressed as

$$F_X(x) = \tau F_S(x) + \Pr(N = 0), \quad (4.13)$$

where $F_S(x)$ is the distribution function of S .

In each example, the number of claims N is assumed to follow a Poisson distribution with parameter λ , denoted by $\text{Poisson}(\lambda)$. In the first example, the individual claim U is gamma distributed, while in the second one, U follows an inverse Gaussian distribution, denoted by $\text{IG}(\mu, \theta)$, with density function

$$f_U(u) = \left(\frac{\theta}{2\pi u^3} \right)^{\frac{1}{2}} \exp\left(-\frac{\theta(u - \mu)^2}{2\mu^2 u} \right), \quad \text{for } u > 0 \text{ and } \mu, \theta > 0.$$

According to Klugman *et al.* (2008, p. 220-221), when the individual claim distributions are gamma or inverse Gaussian, both are closed under convolutions and an analytic form of the density function $f_X(x)$, $x > 0$, is available. In the following subsections, we compare graphically the approximate distributions obtained by applying the proposed approach with their exact counterparts. Additionally, we generated sample values of the aggregate claims by simulation and compared the plots of the approximated distribution functions based on the sample moments of the underlying distribution with the simulated empirical distribution functions. The calculations were carried out by making use of the symbolic computational software package *Mathematica*. The code is available from the authors upon request.

4.2.1.1 Example 1

Let $N \sim \text{Poisson}(3)$, $U \sim \text{gamma}(3, 2)$, and the order of the polynomial adjustment be $t = 15$. The moment generating function of the aggregate claims being in this case

$$M_X(k) = \exp \left[3 \left(\frac{1}{(1-2k)^3} - 1 \right) \right], \quad (4.14)$$

the moments $\mu_S(a)$, $a = 1, 2, \dots$, can be easily evaluated from (4.10) and (4.12). From (4.4), the base density function is $\text{gamma}(2.684349, 7.056877)$ distributed.

The plots of the exact and approximated density and distribution functions are respectively presented in Figures 4.1 and 4.2. Figure 4.3 compares the approximated distribution function obtained from an application of the proposed methodology in conjunction with the sample moments calculated from simulated values on the basis of 100, 500, 1,000 and 10,000 replications, with the empirical distribution function. We selected $t = 15$ in this example, as increasing t to 16 did not produce a noticeable improvement (graphically).

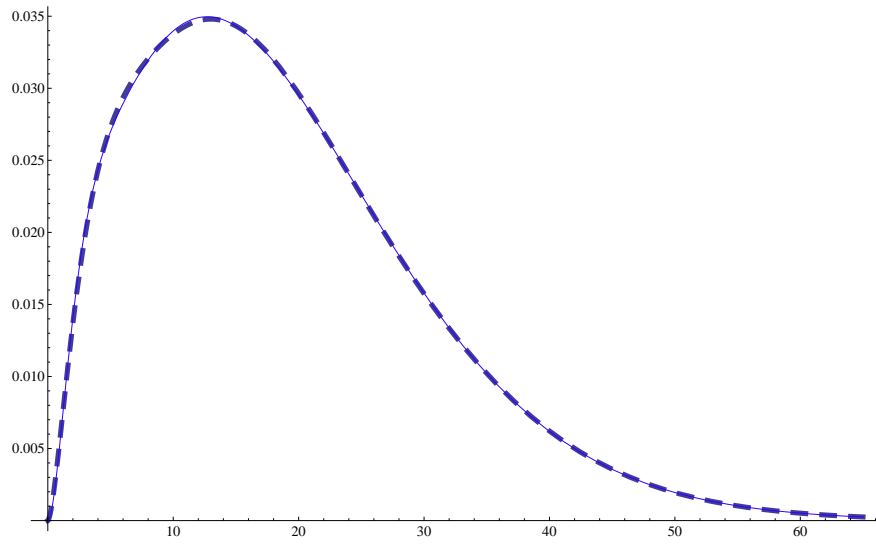


Fig. 4.1: Exact and approximated (dashed line) density functions.

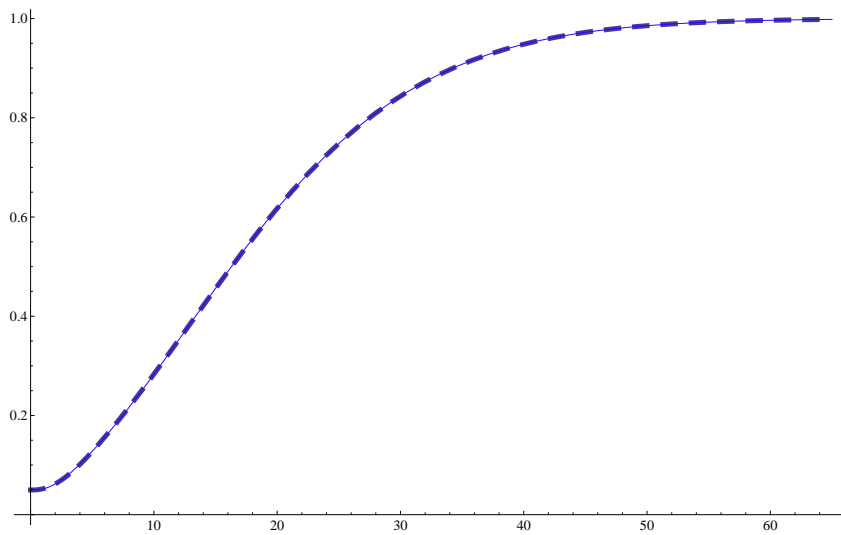


Fig. 4.2: Exact and approximated (dashed line) distribution functions.

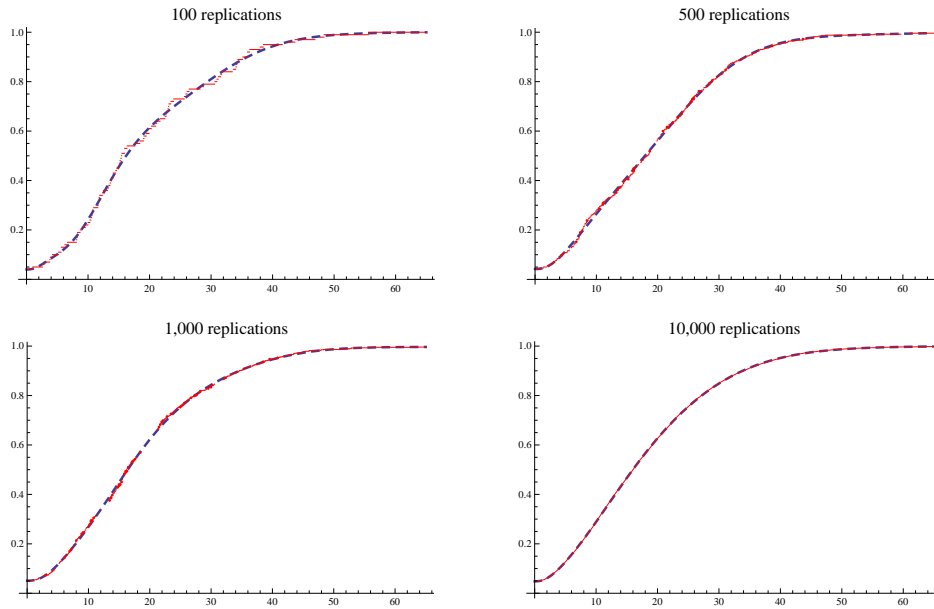


Fig. 4.3: Empirical and approximated (dashed line) distribution functions obtained on the basis of simulated values.

4.2.1.2 Example 2

Let $N \sim \text{Poisson}(15)$, $U \sim \text{IG}(3, 2)$ and $t = 13$. The plots of the exact and approximated density and distribution functions are respectively presented in Figures 4.4 and 4.5. Figure 4.6 compares the approximated distribution function based on the sample moments with the empirical distribution function.

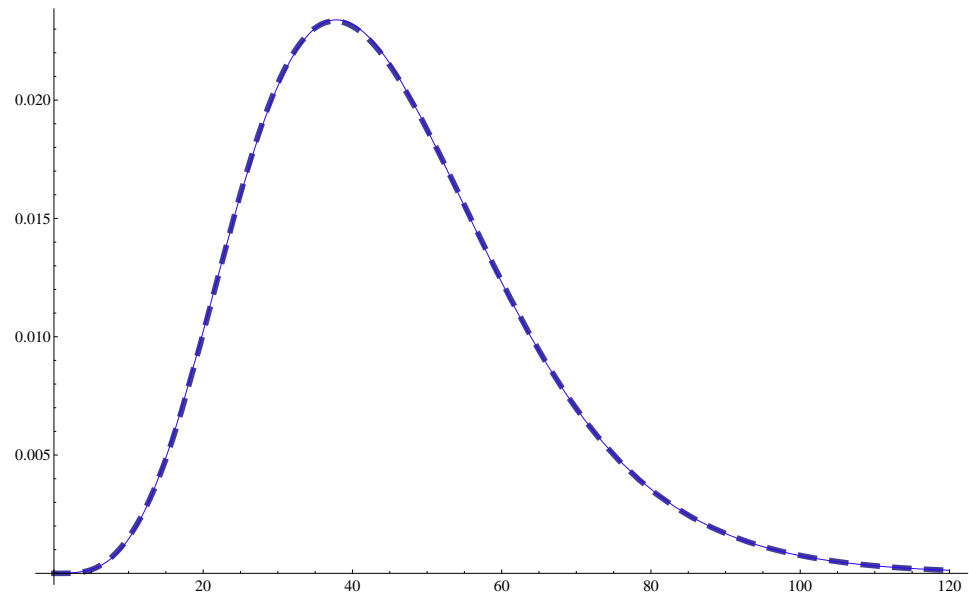


Fig. 4.4: Exact and approximated (dashed line) density functions.

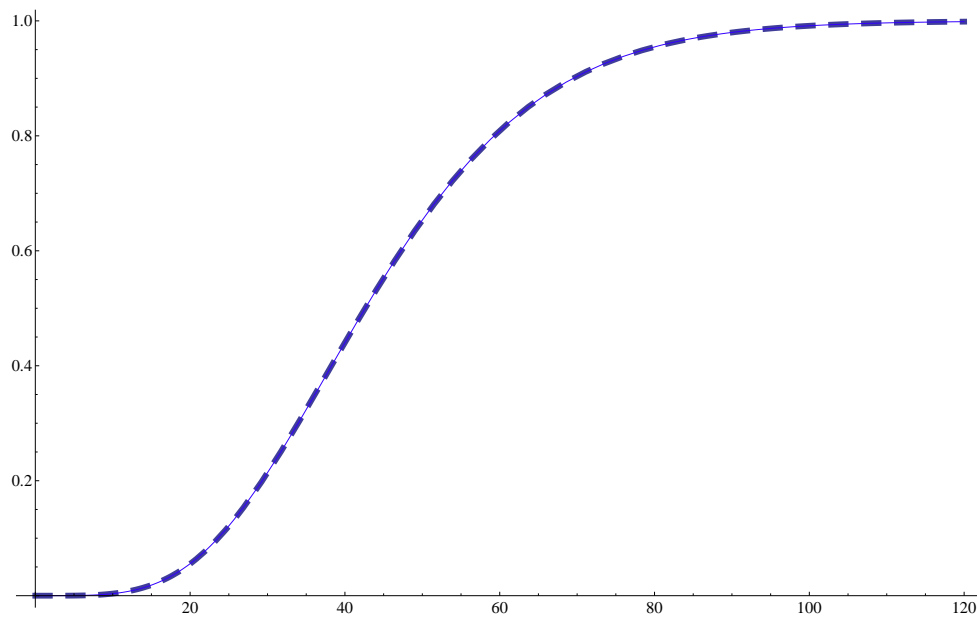


Fig. 4.5: Exact and approximated (dashed line) cumulative distribution functions.

It is seen from both examples that the proposed approximation methodology proves very accurate. However, it should be pointed out that the proposed method-

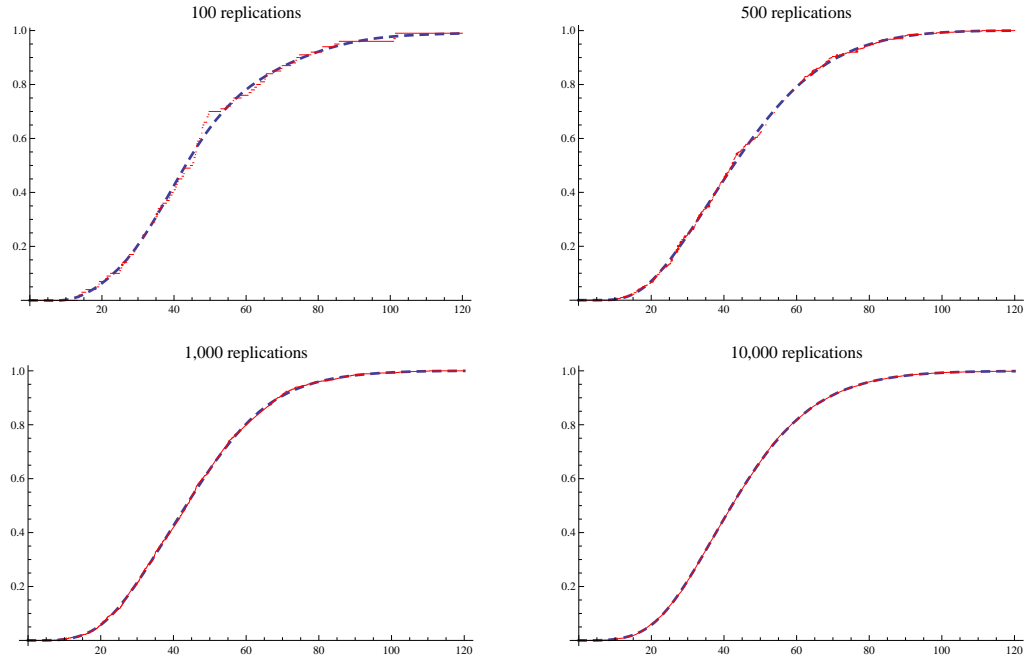


Fig. 4.6: Empirical and approximated (dashed line) distribution functions obtained on the basis of simulated values.

ology would not be applicable when the target distribution has an extremely heavy tail since then only a limited number of moments would be available.

4.3 Approximating the distribution of bivariate aggregate claims

Let $f_{S_1, S_2}(s_1, s_2)$, $s_1, s_2 > 0$, be the joint density function of the continuous random variables S_1 and S_2 . In this section, we extend the technique introduced in Section 4.2 to approximate $f_{S_1, S_2}(s_1, s_2)$. The methodology can be described as follows.

First, a pair of uncorrelated random variables (V, Z) are produced from (S_1, S_2) by applying the linear transformation,

$$\begin{pmatrix} V \\ Z \end{pmatrix} = \mathbb{C}^{-\frac{1}{2}} \begin{pmatrix} S_1 \\ S_2 \end{pmatrix}, \quad (4.15)$$

where $\mathbb{C}^{-\frac{1}{2}}$ is the inverse of the symmetric square root of the covariance matrix of (S_1, S_2) , that is,

$$\mathbb{C}^{-\frac{1}{2}} = \begin{pmatrix} \text{Var}(S_1) & \text{Cov}(S_1, S_2) \\ \text{Cov}(S_1, S_2) & \text{Var}(S_2) \end{pmatrix}^{-\frac{1}{2}} \equiv \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix}. \quad (4.16)$$

By setting the bivariate base density function as

$$g(v, z) = g_1(v) g_2(z) \quad (4.17)$$

where $g_1(v)$ and $g_2(z)$ are the approximated marginal density functions of V and Z , which are assumed to be $\text{gamma}(\alpha_1, \theta_1)$ and $\text{gamma}(\alpha_2, \theta_2)$ distributed respectively, the density function of (V, Z) can be approximated by

$$\tilde{f}_{V,Z,t}(v, z) = g(v, z) \sum_{r+q \leq t} c_{r,q} v^r z^q, \quad v, z > 0, \quad (4.18)$$

where r and q are nonnegative integers, t is an appropriately selected positive integer and the $c_{r,q}$'s are polynomial coefficients to be determined. One could also take $(0, \dots, t)$, as the range of each of the indices r and q , but it was observed that for a given t , this did not result in a noticeable improvement in accuracy. Note that by making use of (4.17), one assumes that the uncorrelated pair (V, Z) is also independently distributed. In general, uncorrelation does not imply independence; however the polynomial adjustment should address most of the remaining dependence relationships between V and Z .

The parameters (α_1, θ_1) and (α_2, θ_2) of the components of the base gamma density functions are obtained by matching the first two moments associated with $g_1(v)$ and $g_2(z)$ to those of V and Z . Next, we once again apply a moment-matching technique to assign numerical values to the $c_{r,q}$'s. The joint moments of (V, Z) are determined from those of (S_1, S_2) , which are assumed to be known. In light of (4.15) and making use of the notation introduced in (4.16), $\mu_{V,Z}(a, b)$, the joint moment of orders a and b of (V, Z) can be expressed as

$$\mu_{V,Z}(a, b) = \text{E}(V^a Z^b) = \text{E} \left[\left(\sum_{k=0}^a \binom{a}{k} \beta_{11}^k \beta_{12}^{a-k} S_1^k S_2^{a-k} \right) \left(\sum_{\ell=0}^b \binom{b}{\ell} \beta_{21}^\ell \beta_{22}^{b-\ell} S_1^\ell S_2^{b-\ell} \right) \right]$$

$$= \sum_{k=0}^a \sum_{\ell=0}^b \binom{a}{k} \binom{b}{\ell} \beta_{11}^k \beta_{12}^{a+\ell-k} \beta_{22}^{b-\ell} \mu_{S_1, S_2}(k + \ell, a + b - k - \ell), \quad (4.19)$$

where a and b are nonnegative integers. These joint moments of V and Z are equated to those associated with the approximate density specified by (4.18):

$$\int_0^\infty \int_0^\infty v^a z^b \left(g(v, z) \sum_{r+q \leq t} c_{r,q} v^r z^q \right) dv dz = \mu_{V,Z}(a, b), \quad \text{for } a + b \leq t. \quad (4.20)$$

This equation can be reexpressed as

$$\sum_{r+q \leq t} c_{r,q} \int_0^\infty \int_0^\infty v^{r+a} z^{q+b} g(v, z) dv dz = \sum_{r+q \leq t} c_{r,q} m_{r+a, q+b} = \mu_{V,Z}(a, b), \quad \text{for } a+b \leq t, \quad (4.21)$$

where

$$m_{r+a, q+b} = [\theta_1^{r+a} (\alpha_1 + r + a - 1) \cdots \alpha_1] [\theta_2^{q+b} (\alpha_2 + q + b - 1) \cdots \alpha_2]. \quad (4.22)$$

The values of the coefficients $c_{r,q}$ are then determined by solving the system of linear equations resulting from (4.21). The value of t can be selected in a manner similar to that proposed for the univariate case in Section 4.2.

Finally, the approximated joint density function of (S_1, S_2) is obtained from that of (V, Z) as follows:

$$\tilde{f}_{S_1, S_2, t}(s_1, s_2) = |J| \tilde{f}_{V, Z, t}(\beta_{11}s_1 + \beta_{12}s_2, \beta_{21}s_1 + \beta_{22}s_2), \quad s_1, s_2 > 0, \quad (4.23)$$

where

$$J = \begin{vmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{vmatrix} \quad (4.24)$$

is the Jacobian of the inverse transformation. In view of (4.18), the approximate density function has the following representation:

$$\begin{aligned} \tilde{f}_{S_1, S_2, t}(s_1, s_2) &= |J| g_1(\beta_{11}s_1 + \beta_{12}s_2) g_2(\beta_{21}s_1 + \beta_{22}s_2) \times \\ &\quad \sum_{r+q \leq t} c_{r,q} (\beta_{11}s_1 + \beta_{12}s_2)^r (\beta_{21}s_1 + \beta_{22}s_2)^q, \quad s_1, s_2 > 0. \end{aligned} \quad (4.25)$$

4.3.1 Numerical examples

First, some preliminary considerations are provided in connection with the proposed methodology for approximating the joint distribution of bivariate aggregate claims as specified by (2.1). Let $P_{N,W}(z_1, z_2) = \mathbf{E}(z_1^N z_2^W)$ be the probability generating function of (N, W) ; then the moment generating function of (X, Y) is given by

$$M_{X,Y}(k_1, k_2) = \mathbf{E}(e^{k_1 X + k_2 Y}) = P_{N,W}(M_U(k_1), M_V(k_2)), \quad (4.26)$$

and the joint moment of orders a and b of (X, Y) is determined as follows:

$$\mu_{X,Y}(a, b) = \left. \frac{\partial^{a+b} M_{X,Y}(k_1, k_2)}{\partial k_1^a \partial k_2^b} \right|_{k_1=0, k_2=0}. \quad (4.27)$$

As in the univariate case, the approximation methodology does not directly apply to the distribution of aggregate claims as the method requires the target density function to be continuous, while the bivariate aggregate claims (X, Y) have point masses at $X = 0$ and $Y = 0$. Accordingly, we first approximate the distribution of the random vector (S_1, S_2) , which is defined only for $X, Y > 0$. The joint density function of (S_1, S_2) is then

$$f_{S_1, S_2}(s_1, s_2) = \frac{f_{X,Y}(s_1, s_2)}{\kappa}, \quad s_1, s_2 > 0, \quad (4.28)$$

where $\kappa = 1 - \Pr(N = 0) - \Pr(W = 0) + \Pr(N = 0, W = 0)$. Thus, the joint moments of (S_1, S_2) are

$$\mu_{S_1, S_2}(a, b) = \begin{cases} \frac{\mu_{X,Y}(a,b)}{\kappa} & \text{if } a, b = 1, 2, \dots \\ \frac{\mu_{X,Y}(0,b) - \Pr(N=0)\mu_{Y|N=0}(b)}{\kappa} & \text{if } a = 0, b = 1, 2, \dots \\ \frac{\mu_{X,Y}(a,0) - \Pr(W=0)\mu_{X|W=0}(a)}{\kappa} & \text{if } a = 1, 2, \dots, b = 0 \\ 1 & \text{if } a = 0, b = 0, \end{cases} \quad (4.29)$$

where $\mu_{Y|N=0}(b)$ denotes the b^{th} conditional moment of Y given $N = 0$ and $\mu_{X|W=0}(a)$ denotes the a^{th} conditional moment of X given $W = 0$. Both $\mu_{Y|N=0}(b)$ and $\mu_{X|W=0}(a)$ may be conveniently computed when the correlation structure between N and W is given.

The joint cumulative distribution function of (X, Y) , that is, $F_{X,Y}(x, y) = \Pr(X \leq x, Y \leq y)$, $x, y \geq 0$, can then be expressed as

$$F_{X,Y}(x, y) = \kappa F_{S_1, S_2}(x, y) + \Pr(N = 0)F_{Y|N=0}(y) + \Pr(W = 0)F_{X|W=0}(x) - \Pr(N = 0, W = 0), \quad (4.30)$$

where $F_{S_1, S_2}(x, y)$ is the distribution function of (S_1, S_2) , $F_{X|W=0}(x)$ and $F_{Y|N=0}(y)$ are the conditional distribution function of X given $W = 0$ and the conditional distribution function of Y given $N = 0$ respectively. Based on (4.29), whenever $x, y > 0$, $F_{S_1, S_2}(x, y)$ can be determined by making use of the methodology described in the introduction of this section, on the basis of the joint moments of (X, Y) . The distribution functions $F_{X|W=0}(x)$ and $F_{Y|N=0}(y)$ can be readily obtained by the method outlined in Section 4.2 for the univariate aggregate claims model.

In the following numerical examples, we utilize the model introduced in Hesselager (1996a) in which $N = N_0 + N_1$ and $W = N_0 + N_2$, where N_0, N_1 and N_2 are mutually independent and $N_i \sim \text{Poisson}(\lambda_i)$ for $i = 0, 1, 2$, so that $(W|N = 0) \sim \text{Poisson}(\lambda_2)$ and $(N|W = 0) \sim \text{Poisson}(\lambda_1)$. Hesselager (1996a) derived recursive formulas to calculate the joint probabilities of (X, Y) when the individual claim random variables U and V are discretized to nonnegative integer values.

In both examples, we compare the approximated values of the distribution functions with those obtained from the recursive method, in which the individual claims are discretized using the standard rounding method (Klugman *et al.* (2008, p. 232)). The discretization interval is taken small enough so that the error involved could be ignored. In addition, we simulated 5,000 sample values of the bivariate aggregate claims and then, compared selected values of the simulated empirical distribution function with those of the approximated distribution function based on the sample moments. Furthermore, to illustrate the dependence between X and Y , we plotted a dependency measure calculated by utilizing both the recursive and the approximation methods.

4.3.1.1 Example 1

Let $N_0 \sim \text{Poisson}(7)$, $N_1 \sim \text{Poisson}(8)$, $N_2 \sim \text{Poisson}(9)$, $U \sim \text{gamma}(2, 2)$, $V \sim \text{gamma}(3, 1.5)$, and $t = 8$. Table 4.1 lists the values of $F_{X,Y}(x, y)$ obtained from the recursive approach and the proposed method at given points of the distribution. Table 4.2 lists the values of $F_{X,Y}(x, y)$ obtained from the empirical and approximated distribution functions which were determined from samples of simulated values. Figure 4.7 shows plots of the dependency measure $\chi(p)$ against p , evaluated by making use of the recursive and approximation methods, the dotted line representing the reference line for the relationship between p and $\Pr(X > \pi_p(X)) = 1 - p$. This figure clearly indicates the presence of a strong dependency between X and Y , which is well captured by both methodologies.

Quantity	Recursion	Approximation	Quantity	Recursion	Approximation
$F_{X,Y}(52, 54)$	0.106406	0.104420	$F_{X,Y}(89, 85)$	0.708083	0.705525
$F_{X,Y}(62, 59)$	0.201156	0.201456	$F_{X,Y}(92, 93)$	0.807833	0.802099
$F_{X,Y}(67, 65)$	0.304023	0.307647	$F_{X,Y}(101, 103)$	0.905475	0.902043
$F_{X,Y}(74, 69)$	0.401565	0.404960	$F_{X,Y}(112, 110)$	0.950726	0.950726
$F_{X,Y}(75, 76)$	0.512051	0.515397	$F_{X,Y}(130, 127)$	0.990449	0.992031
$F_{X,Y}(81, 80)$	0.604866	0.604980	$F_{X,Y}(150, 150)$	0.999286	0.999013

Table 4.1: Some values of $F_{X,Y}(x, y)$ calculated from the recursive approach and the proposed method.

Quantity	Empirical	Approximated	Quantity	Empirical	Approximated
$F_{X,Y}(52, 54)$	0.110800	0.110579	$F_{X,Y}(89, 85)$	0.700200	0.700284
$F_{X,Y}(62, 59)$	0.201800	0.203602	$F_{X,Y}(92, 93)$	0.804600	0.798009
$F_{X,Y}(67, 65)$	0.300600	0.304419	$F_{X,Y}(101, 103)$	0.898800	0.898436
$F_{X,Y}(74, 69)$	0.394000	0.398631	$F_{X,Y}(112, 110)$	0.948400	0.947691
$F_{X,Y}(75, 76)$	0.508800	0.507593	$F_{X,Y}(130, 127)$	0.989200	0.990421
$F_{X,Y}(81, 80)$	0.600200	0.597797	$F_{X,Y}(150, 150)$	0.999400	0.999201

Table 4.2: Some values of $F_{X,Y}(x, y)$ calculated from the empirical and the approximated distribution functions obtained on the basis of simulated values.

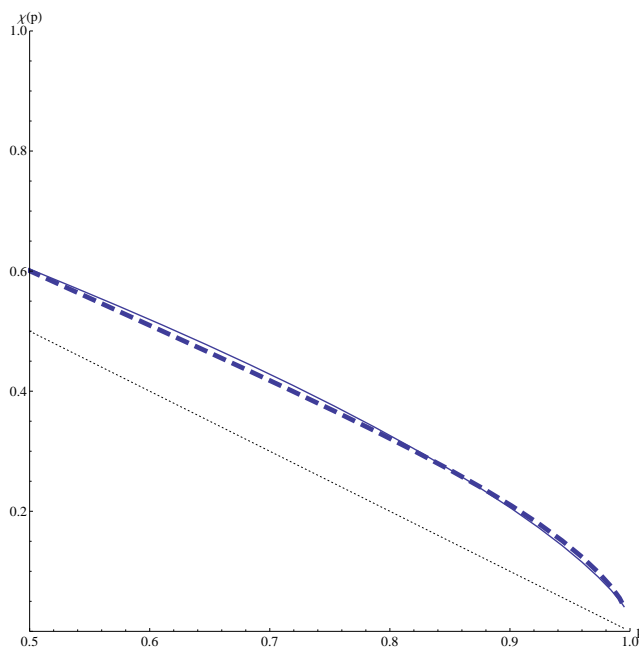


Fig. 4.7: Exact and approximated (dashed line) dependency measures between X and Y .

4.3.1.2 Example 2

Let $N_0 \sim \text{Poisson}(8)$, $N_1 \sim \text{Poisson}(9)$, $N_2 \sim \text{Poisson}(10)$, $U \sim \text{IG}(3, 4)$, $V \sim \text{IG}(4, 5)$, and $t = 8$. For this case, the results are presented in Tables 4.3 and 4.4 as well as in Figure 4.8.

Quantity	Recursion	Approximation	Quantity	Recursion	Approximation
$F_{X,Y}(48, 50)$	0.103833	0.102490	$F_{X,Y}(80, 85)$	0.708307	0.708275
$F_{X,Y}(58, 56)$	0.204674	0.202657	$F_{X,Y}(98, 91)$	0.802741	0.803056
$F_{X,Y}(62, 62)$	0.301747	0.300674	$F_{X,Y}(108, 103)$	0.903843	0.902583
$F_{X,Y}(68, 67)$	0.402988	0.402674	$F_{X,Y}(115, 113)$	0.951134	0.950217
$F_{X,Y}(71, 73)$	0.508339	0.508959	$F_{X,Y}(140, 134)$	0.990466	0.990828
$F_{X,Y}(78, 78)$	0.607973	0.609034	$F_{X,Y}(155, 160)$	0.999064	0.999090

Table 4.3: Some values of $F_{X,Y}(x, y)$ calculated from the recursive approach and the proposed method.

Quantity	Empirical	Approximated	Quantity	Empirical	Approximated
$F_{X,Y}(48, 50)$	0.103200	0.103436	$F_{X,Y}(80, 85)$	0.713000	0.713866
$F_{X,Y}(58, 56)$	0.199800	0.203058	$F_{X,Y}(98, 91)$	0.804800	0.809291
$F_{X,Y}(62, 62)$	0.304200	0.301513	$F_{X,Y}(108, 103)$	0.908200	0.908783
$F_{X,Y}(68, 67)$	0.406000	0.404530	$F_{X,Y}(115, 113)$	0.956800	0.955374
$F_{X,Y}(71, 73)$	0.511800	0.511497	$F_{X,Y}(140, 134)$	0.993000	0.992575
$F_{X,Y}(78, 78)$	0.613400	0.613464	$F_{X,Y}(155, 160)$	0.999200	0.998926

Table 4.4: Some values of $F_{X,Y}(x, y)$ calculated from the empirical and the approximated distribution functions obtained on the basis of simulated values.

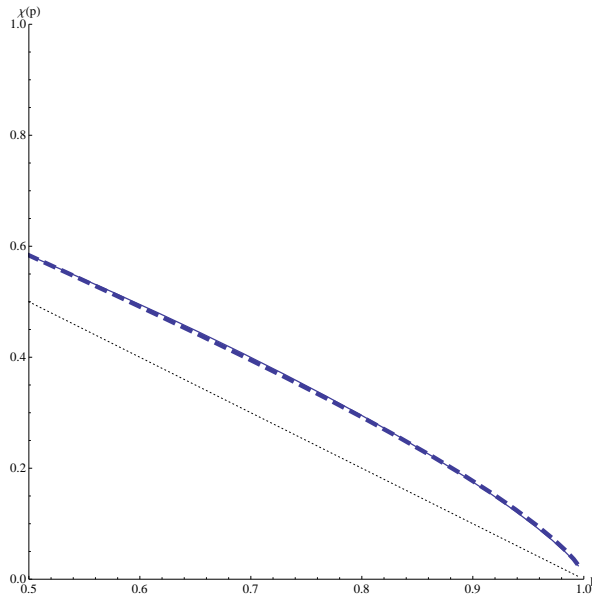


Fig. 4.8: Exact and approximated (dashed line) dependency measures between X and Y .

Moreover, upper quantiles of the bivariate compound distributions are readily evaluated by integration of the approximate density function. It is seen from both examples that, overall, the proposed approximation method can provide quite accurate values for the distribution of bivariate aggregate claims.

4.4 Conclusions

We applied a moment-based density approximation method to model the distributions of univariate and bivariate aggregate claims. The examples indicate that this methodology is reliable. Not only the proposed technique is conceptually simple and computationally efficient, but it also produces approximation results that are rather accurate. This approach can therefore be utilized for evaluating right tail quantiles of aggregate claims distributions. Additionally, given a set of observed aggregate claims, the method advocated herein can readily be applied in conjunction with the sample moments for modeling purposes.

Chapter 5

Concluding Remarks

In this thesis we first introduce three bivariate aggregate claims models in Hesselager (1996a) with the corresponding recursive formulas for calculating the joint probabilities of two types of aggregate claims. Then we show how to use FFT method with exponential tilting to compute the joint probabilities, a detailed comparison of the FFT method and the recursive method for computing bivariate distribution functions is provided using numerical examples. Next, we introduce a new bivariate aggregate claims model in which both claim numbers and claim sizes are dependent. We derive the recursive formulas for the joint probabilities of the bivariate aggregate claims for three types of correlation structures. In the numerical examples, the results from the recursive formulas are compared with the results from FFT method with tilting to verify the correctness of the derived recursive formulas. After that, we apply a moment-based method introduced by Provost (2005) to approximate the density functions of univariate and bivariate aggregate claims random variables. Using the bivariate model in Hesselager (1996a), we test this approximation method by comparing the value of joint distribution function calculating from the approximated joint density, with the results calculating from both theoretical formulas and simulations. It shows that this moment-based approximation method works very effectively.

As shown in this thesis, the recursive method provides exact joint probabilities

if the floating point representation error is ignored. Note that when the claim size random variables are continuous, the recursive approach is also an approximation in the sense that the continuous distribution is approximated using a discrete distribution, however, we can take the discretization interval small enough so that the error involved could be ignored. The recursive method could be useful when we need theoretical values of probabilities and use it to calibrate other approximation methods. The drawbacks of the recursive method are intensive computational time and limitations on the claim number distributions. In this aspect, FFT and moment-based density approximation methods could be viable alternatives to the recursive methods. The FFT method can deal with arbitrary claim number distributions as long as the joint PGF is known; the moment-based density approximation method only need some joint moments of the aggregate claims and is particularly useful when no separate information on claim numbers and individual claim sizes is available. In conclusion, the three methods discussed in this thesis have their own pros and cons, the choice of the methods depends on the different factors in different scenarios.

References

- [1] Ambagaspitiya, R. S. (1999) On the distributions of two classes of correlated aggregate claims. *Insurance: Mathematics and Economics*, 24, 301-308.
- [2] Bühlmann, H. (1984) Numerical evaluation of the compound Poisson distribution: recursion or fast Fourier transform? *Scandinavian Actuarial Journal*, 2, 116-126.
- [3] Clark, D. R. and Homer, D. L. (2003) Insurance applications of bivariate distributions. *Proceedings of the Casualty Actuarial Society*, XC, 172 & 173, 274-307.
- [4] Coles, S., Heffernan, J. and Tawn, J. (2000) Dependence measures for extreme value analyses. *Extremes*, 2(4), 339-365.
- [5] Cummins, J. and Wiltbank, L. (1983) Estimating the total claims distribution using multivariate frequency and severity distributions. *The Journal of Risk and Insurance*, 50, 377-403.
- [6] Embrechts, P. and Frei, M. (2009) Panjer recursion versus FFT for compound distributions. *Mathematical Methods of Operations Research*, 69, 497-508.
- [7] Frees, E. and Valdez, E. (2008) Hierarchical insurance claims modeling. *Journal of the American Statistical Association*, 103, 1457-1469.
- [8] Grübel, R. and Hermesmeier, R. (1999) Computation of compound distributions I: aliasing errors and exponential tilting. *ASTIN Bulletin*, 29(2), 197-214.
- [9] Hardy, M. R. (2007) Approximating the aggregate claims distribution. *Encyclopedia of Actuarial Science*, 1, 76-86.
- [10] Heckman, P. E. and Meyers G. G. (1983) The calculation of aggregate loss distributions for claim severity and claim count distributions, *Proceedings of the Casualty Actuarial Society*, LXX, 22-61.

- [11] Hesselager, O. (1996a) Recursions for certain bivariate counting distributions and their compound distributions. *ASTIN Bulletin*, 26(1), 35-52.
- [12] Hesselager, O. (1996b) A recursive procedure for calculation of some mixed compound Poisson distributions. *Scandinavian Actuarial Journal*, 1, 54-63.
- [13] Homer, David L. and Clark, David R. (2003) Rinsurance applications of bivariate distributions. *Proceedings of the Casualty Actuarial Society* , 274-307.
- [14] Klugman, S. A., Panjer, H. H. and Willmot, G. E. (2008) *Loss Models*, 3rd edition, Wiley, New Jersey.
- [15] Lindley, D. V. and Singpurwalla, N. D. (1986) Multivariate distribution for the life lengths of a system sharing a common environment. *Journal of Applied Probability*, 23, 418-431.
- [16] Panjer, H. H. (1981) Recursive evaluation of a family of compound distributions. *ASTIN Bulletin*, 12(1), 22-26.
- [17] Panjer, H. H. and Wang, S. (1993) On the stability of recursive formulas. *ASTIN Bulletin*, 23(2), 227-258.
- [18] Panjer, H. H. and Willmot, G. E. (1986) Computational aspects of recursive evaluation of compound distributions. *Insurance: Mathematics and Economics*, 5, 113-116.
- [19] Papush, D. E., Patrik, G. S. and Podgaitz, F. (2001) Approximations of the aggregate loss distribution. *CAS Forum*, Winter, 175-186.
- [20] Provost, S. B. (2005) Moment-based density approximants. *The Mathematica Journal*, 9, 727-756.
- [21] Robertson, J. (1992) The computations of aggregate loss distributions. *Proceedings of the Casualty Actuarial Society*, LXXIX, 57-133.

- [22] Sundt, B. (1982) Asymptotic behaviour of compound distributions and stop-loss premiums. *ASTIN Bulletin*, 13(2), 89-98.
- [23] Sundt, B. (1992) On some extensions of Panjer's class of counting distributions. *ASTIN Bulletin*, 22, 61-80.
- [24] Sundt, B. (1999) On multivariate Panjer recursions. *ASTIN Bulletin*, 29, 29-45.
- [25] Sundt, B. and Jewell, W. S. (1981) Further results on recursive evaluation of compound distributions. *ASTIN Bulletin*, 12, 27-39.
- [26] Sundt, B. and Vernic, R. (2004) Recursions for compound mixed multivariate Poisson distributions. *Blätter der Deutschen Gesellschaft für Versicherungs und Finanzmathematik*, XXVI, 665-691.
- [27] Sundt, B. and Vernic, R. (2009) *Recursions for Convolutions and Compound Distributions with Insurance Applications*. Springer Verlag: Berlin-Heidelberg.
- [28] Tijms, H.C. (1994) *Stochastic Models: An Algorithmic Approach*. Chichester: John Wiley.
- [29] Vernic, R. (1999) Recursive evaluation of some bivariate compound distributions. *ASTIN Bulletin*, 29, 315-325.
- [30] Wang, Shaun S. (1998) Aggregation of correlated risk portfolios: Models and Algorithms. *Proceedings of the Casualty Actuarial Society*, 85, 848-937.
- [31] Willmot, G. (1993) On recursive evaluation of mixed Poisson probabilities and related quantities. *Scandinavian Actuarial Journal*, 2, 114-133.

Appendix A

Recursive formulas for three models in Hesselager (1996a)

A.1 Recursive formulas for model A

Under the conditions in Model A, it holds that

$$g(0, 0) = P_K(p_1 f_1(0) + p_2 f_2(0)). \quad (\text{A.1})$$

For $x \geq 1$,

$$g(x, y) = \frac{1}{1 - ap_1 f_1(0) - ap_2 f_2(0)} \times \left[p_1 \sum_{u=1}^x \left(a + \frac{bu}{x} \right) f_1(u) g(x - u, y) + ap_2 \sum_{v=1}^y f_2(v) g(x, y - v) \right], \quad (\text{A.2})$$

and for $y \geq 1$,

$$g(x, y) = \frac{1}{1 - ap_1 f_1(0) - ap_2 f_2(0)} \times \left[p_2 \sum_{v=1}^y \left(a + \frac{bv}{y} \right) f_2(v) g(x, y - v) + ap_1 \sum_{u=1}^x f_1(u) g(x - u, y) \right]. \quad (\text{A.3})$$

A.2 Recursive formulas for model B

Assume that the conditions in Model B hold true. Then

$$g(0, 0) = P_{Z_0}(f_1(0)f_2(0))P_{Z_1}(f_1(0))P_{Z_2}(f_2(0)). \quad (\text{A.4})$$

For $x \geq 1$,

$$\begin{aligned} g(x, y) = & \frac{1}{1 - a_1 f_1(0) - a_0 f_1(0) f_2(0) + a_0 a_1 f_1^{*2}(0) f_2(0)} \times \\ & \left[\sum_{u=1}^x \left(a_1 + \frac{b_1 u}{x} \right) f_1(u) g(x-u, y) + \sum_{v=1}^y a_0 f_1(0) f_2(v) g(x, y-v) \right. \\ & + \sum_{u=1}^x \sum_{v=0}^y \left(a_0 + \frac{b_0 u}{x} \right) f_1(u) f_2(v) g(x-u, y-v) \\ & - \sum_{v=1}^y a_0 a_1 f_1^{*2}(0) f_2(v) g(x, y-v) \\ & \left. - \sum_{u=1}^x \sum_{v=0}^y \left(a_0 a_1 + \frac{(a_0 b_1 + b_0 a_1) u}{2x} \right) f_1^{*2}(u) f_2(v) g(x-u, y-v) \right], \quad (\text{A.5}) \end{aligned}$$

and for $y \geq 1$,

$$\begin{aligned} g(x, y) = & \frac{1}{1 - a_2 f_2(0) - a_0 f_2(0) f_1(0) + a_0 a_2 f_2^{*2}(0) f_1(0)} \times \\ & \left[\sum_{v=1}^y \left(a_2 + \frac{b_2 v}{y} \right) f_2(v) g(x, y-v) + \sum_{u=1}^x a_0 f_2(0) f_1(u) g(x-u, y) \right. \\ & + \sum_{v=1}^y \sum_{u=0}^x \left(a_0 + \frac{b_0 v}{y} \right) f_2(v) f_1(u) g(x-u, y-v) \\ & - \sum_{u=1}^x a_0 a_2 f_2^{*2}(0) f_1(u) g(x-u, y) \\ & \left. - \sum_{v=1}^y \sum_{u=0}^x \left(a_0 a_2 + \frac{(a_0 b_2 + b_0 a_2) v}{2y} \right) f_2^{*2}(v) f_1(u) g(x-u, y-v) \right]. \quad (\text{A.6}) \end{aligned}$$

A.3 Recursive formulas for model C

For $j = 1, 2$, let

$$g_\theta^{(j)}(x) = \sum_{n=0}^{\infty} \frac{(\theta\lambda_j)^n}{n!} e^{-\theta\lambda_j} f_j^{*n}(x)$$

denote the conditional probability functions of X and Y and let

$$g_\theta(x, y) = \sum_{n,m=0}^{\infty} p_\theta(n, m) f_1^{*n}(x) f_2^{*m}(y) = g_\theta^{(1)}(x) g_\theta^{(2)}(y)$$

denote the joint conditional probability function of (X, Y) . Define the auxiliary functions

$$h_i(x, y) = \int_{\sigma_1}^{\sigma_2} \theta^i g_\theta(x, y) u(\theta) d\theta,$$

then $g(x, y) = h_0(x, y)$.

Under the conditions in Model C, it holds that

$$h_i(0, 0) = \int_{\sigma_1}^{\sigma_2} \theta^i e^{-\lambda \cdot (1 - \tilde{f}(0)) \theta} u(\theta) d\theta, \quad (\text{A.7})$$

with $\lambda = \lambda_1 + \lambda_2$, and

$$\tilde{f}(0) = \frac{\lambda_1 f_1(0) + \lambda_2 f_2(0)}{\lambda_1 + \lambda_2}.$$

For $i = 0, \dots, k-1$,

$$h_i(x, y) = \lambda_1 \sum_{u=1}^x \frac{u}{x} f_1(u) h_{i+1}(x-u, y), \quad x \geq 1, \quad (\text{A.8})$$

$$h_i(x, y) = \lambda_2 \sum_{v=1}^y \frac{v}{y} f_2(v) h_{i+1}(x, y-v), \quad y \geq 1, \quad (\text{A.9})$$

and

$$\begin{aligned} c_k h_k(x, y) &= \lambda_1 \sum_{u=1}^x f_1(u) \sum_{i=0}^k b_i h_i(x-u, y) + \lambda_2 \sum_{v=1}^y f_2(v) \sum_{i=0}^k b_i h_i(x, y-v) \\ &\quad + \sum_{i=0}^{k-1} h_i(x, y) [(i+1)b_{i+1} - c_i], \end{aligned} \quad (\text{A.10})$$

where $c_i = \lambda \cdot (1 - \tilde{f}(0)) b_i - a_i$.

Vita

Name:	Tao Jin
Post-secondary Education and Degrees:	<p>The University of Western Ontario London, Ontario, Canada 2010-Present, Ph.D., Actuarial Science</p> <p>The University of Western Ontario London, Ontario, Canada 2008-2009, M.Sc., Actuarial Science</p> <p>Zhejiang University Hangzhou, Zhejiang, China 1992-1996, B.Eng., Chemical Engineering</p>
Awards and Scholarships:	<p>Queen Elizabeth II Graduate Scholarships in Science and Technology 2012-2014</p> <p>Ontario Graduate Scholarship 2011-2012</p> <p>Graduate Thesis Research Award The University of Western Ontario, London, Canada 2012-2013</p> <p>CKER travel grant Society of Actuaries 2011</p> <p>Western Graduate Research Scholarship The University of Western Ontario, London, Canada 2008-2013</p>
Realted Work Experience:	<p>Graduate Research Assistant The University of Western Ontario London, Ontario, 2008-2013</p> <p>Graduate Teaching Assistant The University of Western Ontario London, Ontario, 2008-2013</p>

Publications

1. Jin, T., Provost, S. B., Ren, J., “Moment-based density approximations for aggregate losses”. Under revision, Scandinavian Actuarial Journal, 2014.
2. Jin, T. and Ren, J., “Recursions and fast Fourier transforms for a new bivariate aggregate claims model”. Published online, DOI:10.1080/03461238.2012.762548, Scandinavian Actuarial Journal, 2013.
3. Jin, T. and Ren, J., “Recursions and fast Fourier transforms for certain bivariate compound distributions”. The Journal of Operational Risk. 5(4), 19-33, 2010.