Bjørn Sundt
Raluca Vernic

EAA LECTURE NOTES

Recursions for Convolutions and Compound Distributions with Insurance Applications

Springer

European Actuarial Academy
developing actuaries

EAA Lecture Notes

EAA Lecture Notes is a series supported by the European Actuarial Academy (EAA GmbH), founded on the 29 August, 2005 in Cologne (Germany) by the Actuarial Associations of Austria, Germany, the Netherlands and Switzerland. EAA offers actuarial education including examination, permanent education for certified actuaries and consulting on actuarial education.

actuarial-academy.com

EAA Lecture Notes

Wüthrich, M.V.; Bühlmann, H.; Furrer, H. **Market-Consistent Actuarial Valuation** 2007

Lütkebohmert, E. **Concentration Risk in Credit Portfolios** 2009

Bjørn Sundt • Raluca Vernic

Recursions for Convolutions and Compound Distributions with Insurance Applications

With 3 Figures and 16 Tables

Springer

Bjørn Sundt
Karenslyst allé 1A
0278 Oslo
Norway
bsundt@bluezone.no

Raluca Vernic
Dept. Mathematics and Informatics
Ovidius University of Constanta
124 Mamaia
900527 Constanta
Romania
rvernic@univ-ovidius.ro

ISSN 1865-2174
e-ISSN 1865-2182
ISBN 978-3-540-92899-7
e-ISBN 978-3-540-92900-0
DOI 10.1007/978-3-540-92900-0
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2008942789

Cover design: WMXDesign GmbH, Heidelberg

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Preface

Around 1980, recursions for aggregate claims distributions started receiving attention in the actuarial literature. Two common ways of modelling such distributions are as compound distributions and convolutions.

At first, one considered recursions for compound distributions. In particular, Panjer (1981) became crucial in this connection. He considered a class of counting distributions consisting of the negative binomial, Poisson, and binomial distributions, that is, the three most common types of counting distributions in the actuarial literature.

In the mid-eighties, De Pril turned his attention to recursions for convolutions and published several papers on that topic. These recursions can be rather time- and space-consuming, and he, and other authors, therefore developed approximations based on these recursions and error bounds for these approximations.

By extending the Panjer class, Sundt (1992) presented a framework that also covered De Pril's recursions for convolutions.

Originally, the recursions were deduced for the probability function of the distribution, but later they were adapted to other functions like cumulative distributions and stop loss transforms. As in this book we focus mainly on recursions for probability functions of distributions on the integers, we shall normally refer to such functions as distributions.

From the late nineties, the theory has been extended to multivariate distributions.

In the present book, we give a presentation of the theory. We restrict to classes of recursions that are somehow related to those of Panjer (1981), and we aim at giving a unified presentation. Although the theory has been developed mainly within the actuarial literature, the recursions can also be applied in other fields, and we have therefore tried to make the presentation relatively general. However, we use applications from insurance to illustrate and motivate the theory. Not all the recursions that we are going to present, are equally interesting in practice, but we wish to give a broad presentation of recursions within the framework of the book with emphasis on how they can be deduced.

We have tried to make the book self-contained for readers with a reasonable knowledge of probability theory. As an exception, we present Theorem 12.1 without proof.

Our main goal is to deduce recursions and give general ideas of how such recursions can be deduced. Although applications are used to illustrate and motivate the theory, the book does not aim at giving guidelines on what methods to apply in situations from practice. We consider the recursions primary as computational tools to be used for recursive evaluation of functions. However, we shall also use the recursions in analytical proofs of for instance characterisations of infinitely divisible distributions and normal distributions.

One area that we have not considered, is numerical stability. That is obviously an important aspect when considering recursions, but the tools and methodology would have been quite different from the rest of the book. Furthermore, if we should have kept the principle of making the book self-contained, then we would have needed to include a lot of deductions of results from numerical mathematics, an area where the present authors do not consider themselves as specialists. Some references are given at the end of Chap. 13.

The most compact way of presenting the theory would have been to start with the most general results and deduce the more specific results as special cases. We have chosen the opposite approach, that is, we start with the simple cases, and gradually develop the more general and complex results. This approach better reflects the historical development, and we find it more pedagogical. When knowing and understanding a proof in a simple model, it will be much easier to follow a generalised proof of an extended result. Furthermore, a reader who wants a recursion for a compound geometric distribution, should not first have to go through a complicated deduction of a messy recursion for a general multivariate distribution and get the desired result as a special case. To make it easier for the reader to see what changes have to be made in the proof when extending a result, we have sometimes copied the proof of the simple case and edited it where necessary. A problem with going from the simple to the complex, is how much to repeat. When the proof of the extension is straight forward, then we sometimes drop the proof completely or some details. We apologise to readers who find that we give too much or too little.

Each chapter is preceded by a summary. We do not give literature references in the main text of the chapter, but allocate that to a section *Further remarks and references* at the end of the chapter.

The book consists of two parts. Part I (Chaps. 1–13) is restricted to univariate distributions, whereas in Part II (Chaps. 14–20), we extend the theory to multivariate distributions. To some extent, we have tried to make Part II mirror Part I.

Chapter 1 gives an introduction to Part I. After having given a motivation for studying aggregate claims distributions in insurance, we introduce some notation and recapitulate some concepts from probability theory.

In Chap. 2, we restrict the class of counting distributions to distributions that satisfy a recursion of order one, that is, the mass of the distribution in an integer n is a factor multiplied by the mass in $n-1$; this factor will normally be in the form $a+b/n$ for some constants a and b. Within this framework, we characterise various classes of counting distributions and deduce recursions for compound distributions. We also deduce recursions for convolutions of a distribution.

Chapter 3 is devoted to compound mixed Poisson distributions. Such distributions are also discussed in Chap. 4 on infinitely divisible distributions. There we in particular use recursions presented in Chap. 2 to prove that an infinitely divisible distribution on the non-negative integers is infinitely divisible if and only if it can be expressed as a compound Poisson distribution.

In Chap. 5, we extend the classes of counting distributions to distributions that satisfy recursions of higher order, and we discuss properties of such distributions. We allow for infinite order of the recursions. Within that setting, the coefficients of a recursion can be expressed as a De Pril transform, a term introduced by Sundt (1995) for a function that is central in De Pril's recursions for convolutions. Many results on De Pril transforms follow as special cases of results deduced in Chap. 5 and are given in Chap. 6, which is devoted to properties and applications of De Pril transforms.

Chapter 7 is devoted to individual models, including collective approximation of such models.

In Chap. 8, we extend the theory to recursions for cumulative functions and tails, and Chap. 9 is devoted to recursions for moments of distributions.

As pointed out above, De Pril presented approximations based on his exact recursions for convolutions. These approximations consist of approximating the mass at zero and the De Pril transform of the distributions. Such approximations constitute the subject of Chap. 10. As the approximations to the distributions are not necessarily distributions themselves, we have to extend the theory to a wider class of functions.

Up to this stage, we have restricted to distributions on the non-negative integers. In Chap. 11, we extend the theory of Chap. 10 to distributions that are bounded from below, and Chap. 12 opens for negative severities.

Part I closes with Chap. 13, where we discuss how we can modify the recursions to avoid problems with numerical underflow or overflow.

In Part II, Chap. 14 mirrors Chap. 1.

In the multivariate case, we distinguish between two cases of compound distributions; those with univariate counting distribution and multivariate severity distribution and those with multivariate counting distribution and univariate severity distributions. The former case is the most straight-forward to obtain by extending the univariate case, and that is the topic of Chap. 15. From the theory presented there, we introduce the De Pril transform of a multivariate function in Chap. 16, which mirrors Chaps. 6 and 7.

Chapters 17 and 18 mirror Chaps. 9 and 10 respectively. It should be relatively straight-forward to also extend the recursions of Chap. 8 to a multivariate setting, but we have not found it worth while to pursue that in this book.

In Chap. 19, we deduce recursions for compound distributions with multivariate counting distribution and univariate severity distributions. Such distributions are also considered in Chap. 20, which mirrors Chap. 3.

We are grateful to all the people who have stimulated us and given us feedback on earlier versions of the manuscript. In particular, we would like to thank Montserrat Guillén for hosting both of us in 2006 and Bjørn Sundt in 2007 on fruitful research leave at the University of Barcelona. We would also like to thank Penelope Brading for helpful advice on linguistic issues.

Bjørn Sundt is grateful to his dear father who always cared and stimulated him during the work on this book, but sadly passed away at the end of the proof-reading period, so that he did not get the opportunity to see the finished book.

Raluca Vernic thanks her parents for unconditional support.

Bjørn Sundt and Raluca Vernic
February 2009

Course Outline

In the following, we indicate some ideas on application of the book as textbook for a one semester course. Preferably, the course should be given on the Master level. Some knowledge of non-life insurance and risk theory will give the students a better motivation.

As the book more aims at giving a general understanding of how to develop and apply recursions than teaching specific methods, the lecturer should himself consider what would be natural for him to include in the course, possibly supplementing with material not contained in the book. Furthermore, the contents of the course should be adapted to the level of the students. Hence, the course structures presented in the following should be considered only as suggestions.

A reasonable course structure could be:

- Chapter 2. Sections 2.1–2.2, 2.3.1–2.3.2, 2.4, 2.7.
- Chapter 3. Sections 3.1–3.4.
- Chapter 4. Example 4.3 can be dropped.
- Chapter 5. Sections 5.1–5.3, Theorem 5.1 can be dropped.
- Chapter 6.
- Chapter 7.
- Chapter 9. Sections 9.1.1, 9.1.3 and 9.2.
- Chapter 10.
- Chapter 13.
- Chapter 15. Sections 15.1–15.5.
- Chapter 16. Sections 16.1–16.4 and 16.7.
- Chapter 18

For a shorter course, the material from Chap. 9 and/or Part II could be dropped.

Instead of including the material from Part II, one could give a more extensive treatment of the univariate case, including, in addition to the material indicated above:

- Chapter 2. Section 2.3.3.
- Chapter 5. Section 5.4.
- Chapter 11.
- Chapter 12.

The material could be presented in the same sequence as in the book.

Material from Chaps. 1 and 14 should be included to the extent needed for understanding the rest of the course, and it should be presented when needed as background for other material; going through the whole of Chap. 1 at the beginning of the course could make the students lose interest before they get really started. Much of this material would typically be known to students from before.

Part I
Univariate Distributions

Chapter 1
Introduction

Summary

In the present chapter, we introduce some concepts that we shall apply later in the book. We restrict to concepts needed in the univariate setting of Part I; in Chap. 14, we introduce concepts that are needed only in the multivariate setting of Part II. Many of the concepts will be known to most readers, e.g. convolutions, compound distributions, moments, generating functions, etc. However, it is our aim to make this book reasonably self-contained. Furthermore, even to readers who are already familiar with the concepts, it might be useful to see them introduced within the notational framework that we will apply in this book.

Although the recursive methods that we shall present in this book, are applicable also in other areas than insurance, we shall use aggregate claims distributions in an insurance context to motivate the methods. Therefore, we give a description of this area in Sect. 1.1.

In Sect. 1.2, we list notation and concepts that will be used in this book. In particular, we give an overview of notation that will be applied in different meanings.

It will be useful to have some general notation for classes of distributions and other functions that will be considered in this book. Such notation will be introduced in Sect. 1.3.

Sections 1.4 and 1.6 are devoted to convolutions and compound distributions respectively. These areas are crucial in this book as its main subject is recursions for such distributions.

Mixed distributions is the topic of Sect. 1.5. In particular, we describe the difference between an empirical and a pure Bayes approach. As motivation, we use determination of insurance premiums.

Section 1.7 covers definition and properties of moments, cumulants, (probability) generating functions, moment generating functions, characteristic functions, Laplace transforms, and cumulant generating functions.

In Sect. 1.8, we introduce some operators that will be useful for us in this book. In particular, we discuss some properties of the cumulation operator and the tail operator.

In Sect. 1.9, we discuss some properties of stop loss premiums.

Finally, in Sect. 1.10, we deduce some criteria for convergence of infinite series with positive terms.

1.1 Aggregate Claims Distributions

The topic of this book is recursions for convolutions and compound distributions on the integers. Since about 1980, this topic has received much attention in the actuar-

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_1, © Springer-Verlag Berlin Heidelberg 2009

ial literature. The setting is then usually evaluation of aggregate claims distributions, that is, the distribution of the aggregate amount of insurance claims occurred in an insurance portfolio within a given period. Studying this distribution is of crucial value for insurance companies. Whereas for most enterprises, uncertainty is an unintended and unwanted effect of the general activities of the enterprise, it is the basis for an insurance company; without uncertainty, there would be no need for insurance. The basis for an insurance company is selling its customers economic protection against uncertain economic losses. When a customer buys such a protection from the insurance company, neither he nor the company knows whether the losses will occur, when they will occur, or their magnitude. Hence, when the insurer sells such a product, he does not know how much it will cost him to deliver it. However, he will typically have statistical data from delivery of such products in the past, and from that he can estimate the probability distribution of amounts and numbers of claims, etc. These distributions can then be incorporated when modelling the distribution of the aggregate claims of a given period. The insurance contract between an insurance company and a customer is called an *insurance policy*, and the customers are called *policyholders*.

Although being a professional risk carrier, an insurance company will itself also normally buy risk protection from another professional risk carrier, a *reinsurance* company. The insurance company (the *cedant*) insures (cedes) some of the risk that it has taken over from its policyholders to a reinsurance company. There are various forms of reinsurance, and the reinsurance cover of an insurance portfolio often consists of various forms with several reinsurers. Reinsurance can increase the capacity of an insurance company, enable it to take on more and/or larger risks. Furthermore, it can be used to make the balance of the insurance company fluctuate less between years.

Let us mention some specific applications of aggregate claims distributions:

1. Evaluation of the need for reserves. A natural requirement could be that the reserves should be so large that the probability that they would be insufficient to cover the claims, should be less than a given value, e.g. 1%. Hence, we need to evaluate the upper 1% quantile of the aggregate claims distribution.
2. Calculation of *stop loss* premiums. With the simplest form of stop loss reinsurance, the reinsurer covers the part of the aggregate claims that exceeds a given limit called the *retention*.
3. Study the effect of changing the form and level of deductibles. Then one can compare the aggregate claims distribution with and without the change. In particular, we can see the effect on the reserve requirement and stop loss premiums.
4. Study how various reinsurance covers will affect the aggregate claims distribution.
5. Study how the aggregate claims distribution will be affected by changes in the composition and size of the portfolio. This is crucial for the company's sales strategies. What market segments should we concentrate on, how will the need for reinsurance be affected by an increase of the portfolio, etc.?

There are various ways to model aggregate claims distributions. Two main classes are individual and collective models.

In an *individual model*, it is assumed that the portfolio consists of a fixed number of independent policies. Then the aggregate claims of the portfolio becomes the sum of the aggregate claims of each of the policies, that is, the aggregate claims distribution of the portfolio becomes the *convolution* (see Sect. 1.4) of the aggregate claims distributions of each of the policies that comprise the portfolio.

In the individual model, the claims of the portfolio were related to the individual policies and then aggregated to the portfolio. In a *collective model*, we do not consider the individual policies. Instead, we consider the individual claims without relating them to the policies, and we assume that the amounts of these claims are independent and identically distributed. Thus, the aggregate claims of the portfolio is still a sum of independent and identically distributed random variables, but whereas in the individual model the number of such variables was the number of policies, in the collective model it is the number of claims. In the individual model, we assumed that the number of policies was given and fixed. It would be highly unrealistic to make such an assumption about the number of claims. The uncertainty about the number of claims is a part of the uncertainty that the insurance company has taken over from its policyholders in its capacity as a professional risk carrier. It is therefore natural to model the number of claims as a random variable, and it is normally assumed that it is independent of the claim amounts. Under these assumptions, the aggregate claims distribution becomes a *compound distribution* (see Sect. 1.6) with the claim number distribution as *counting distribution* and the claim amount distribution as *severity distribution*.

We can also have combinations of individual and collective models where we consider the subportfolio of "normal" policies collectively and some special policies individually. Furthermore, collective models are often applied as approximations to individual models.

A special case of the individual model is the *individual life model*, where each policy can have at most one claim with a given amount. This model can typically be applied to life assurance. Here each policy usually has a *sum assured* which is paid out when the policyholder dies. When assuming that the policies are independent, the total amount paid out from the portfolio during a given period can be modelled by the individual life model. For long term policies, there will normally be built up a reserve for each policy. Then the loss suffered by the company when a policyholder dies, is the sum assured minus the reserve of the policy. This amount is called the *sum at risk* of the policy. The aggregate loss of the company during the period is the sum of the sums at risk of the policyholders who die during the period, and can be modelled by the individual life model.

A probabilistic model will always be an approximation to reality where one has to find a balance between realism and mathematical simplicity; a completely realistic model would have insurmountable complexity and so many parameters that it would be impossible to estimate all of them reliably. Furthermore, we would normally not have sufficient knowledge to set up a completely realistic model. In particular, some of the independence assumptions we have made, would not always be realistic; a windstorm could destroy many houses, and in a cold winter with icy roads, there could be many automobile accidents of a similar kind. Since the mid-nineties, there has been an increasing interest in modelling dependence between insurance claims and risks. We shall allow for such dependences in Part II where we study recursions for multivariate distributions.

At the beginning of this section, we made the restriction that the claim amounts should be integer-valued. By changing the monetary unit, the theory trivially extends to claim amounts in the set $\{hi\}_{i=-\infty}^{\infty}$ for some constant $h > 0$. Such distributions are called *lattice distributions* or *arithmetic distributions*.

As so much of the literature on recursions for convolutions and compound distributions has appeared in the framework of aggregate claims distributions, it seems natural to use that framework as a reference point in the presentation in this book. However, the recursions should also be applicable in other settings both within and outside actuarial science.

In the discussion above, we have assumed that all the distributions we have discussed, are univariate. That will be the framework in Part I of the book. In Part II, we shall extend the theory to multivariate distributions.

1.2 Some Notation and Conventions

Let us now introduce some notation and conventions that will be applied in the book:

1. In Part I, we shall mainly concentrate on distributions on the integers. For convenience, we shall normally mean such distributions when referring to discrete distributions or the discrete case. Furthermore, our main attention will be on the probability function of the distribution, and then it will be convenient to associate the distribution with its probability function. Hence, when referring to a distribution, we shall normally mean its probability function. Such functions will be denoted by lower case italics. On the rare occasions when we encounter other sorts of univariate distributions, this convention becomes awkward. Although somewhat inconsistent, we associate such a distribution with its cumulative distribution function, which we denote by a capital italic. Hence, the case of the letter indicates what we mean.
2. To avoid thinking too much about regularity conditions, etc., we will always tacitly assume the existence of the quantities that we work with.
3. When giving a function an argument outside the range for which it has been defined, we tacitly assume that the value of the function for that argument is equal to zero.
4. A summation $\sum_{i=a}^{b}$ is assumed to be equal to zero when $b < a$.
5. If x is a real number, then we let $[x]$ denote the largest integer less than or equal to x and $\{x\}$ the smallest integer greater than or equal to x, and we let $x_+ = \max(x, 0)$ and $x_- = \min(x, 0)$.
6. If u is a univariate function, then we let $u(x+) = \lim_{y\downarrow x} u(y)$ and $u(x-) = \lim_{y\uparrow x} u(y)$. Analogously, if w_θ is a function that depends on a parameter θ, then we let $w_{\gamma+} = \lim_{\theta\downarrow\gamma} w_\theta$ and $w_{\gamma-} = \lim_{\theta\uparrow\gamma} w_\theta$.

7. We shall normally indicate the number of an element in a sequence by a subscript. However, if that number does not contain necessary information in a context, then, for convenience, we sometimes drop it. For instance, if $Y_1, Y_2, \ldots$ are identically distributed random variables, then we might write $\mathrm{E}\,Y$ instead of $\mathrm{E}\,Y_i$ as $\mathrm{E}\,Y_i$ has the same value for all values of i.
8. For a sequence $x_1, x_2, \ldots$, we let $\mathbf{x}_n$ denote the $n \times 1$ vector of the n first elements, that is, $\mathbf{x}_n = (x_1, x_2, \ldots, x_n)'$.
9. If Υ is an operator on functions on some countable set, then we define $\Upsilon^t f = \Upsilon \Upsilon^{t-1} f$ for $t = 1, 2, \ldots$ with initial value $\Upsilon^0 f = f$.
10. We use the Halmos *iff* in the meaning *if and only if*.
11. A □ is used to mark the end of a proof or an example.

Sometimes we shall use the same notation in different meanings, but we hope that this will not create confusion. In the following, we list some cases.

1. For a sequence $x_1, x_2, \ldots$, we let $x_{\bullet n} = \sum_{j=1}^{n} x_j$. However, in a two-way classification with numbers x_{ij}, we let $x_{\bullet j} = \sum_i x_{ij}$ and $x_{i\bullet} = \sum_j x_{ij}$.
2. We shall apply the letter π for a parameter in some parametric classes of distributions. On the rare occasions when we apply π in its traditional meaning as the Ludolph number $3.14\ldots$, then that will be pointed out explicitly.
3. We shall denote the rth derivative of a function f by $f^{(r)}$, but in the context of approximations, we use that notation for an rth order approximation (see Chap. 10), and we shall also use $^{(r)}$ in the context $n^{(r)} = \prod_{j=0}^{r-1}(n-j)$.
4. We shall apply Γ for the *Gamma function*

$$\Gamma(x) = \int_0^\infty y^{x-1} \mathrm{e}^{-y}\, \mathrm{d}y, \qquad (x > 0) \tag{1.1}$$

   but also for the cumulation operator (see Sect. 1.8).
5. We let I denote the indicator function, that is, $I(A)$ is equal to one if the condition A is satisfied, and zero otherwise. However, we also sometimes apply I for a parameter.
6. We shall indicate complementary sets by $\sim$, that is, if A and B are sets, then $A \sim B$ is the set of elements that are contained in A, but not in B. However, we shall also apply $\sim$ to denote asymptotic equality, that is, by

$$f(x) \sim g(x), \qquad (x \to a)$$

   we shall mean that $\lim_{x \to a} f(x)/g(x) = 1$.

Other conventions and notation will be introduced in the following sections.

1.3 Classes of Distributions and Functions

We let $\mathcal{P}_1$ denote the class of all univariate distributions on the integers. In Part I, the subscript 1 may seem as a nuisance; the reason for having it is that in Part II, we will replace it with m when we introduce similar notation for classes of m-variate distributions.

For all integers l, we let $\mathcal{P}_{1l}$ denote the class of all distributions $f \in \mathcal{P}_1$ for which $f(x) = 0$ for all integers $x < l$. We obviously have $\mathcal{P}_{1,l+1} \subset \mathcal{P}_{1l}$. We also introduce $\mathcal{P}_{1\underline{l}} = \mathcal{P}_{1l} \sim \mathcal{P}_{1,l+1}$, that is, the class of all distributions in $\mathcal{P}_{1l}$ with a positive mass at l.

Let

$$\mathcal{P}_{1\_} = \bigcup_{l=-\infty}^{\infty} \mathcal{P}_{1\underline{l}} = \bigcup_{l=-\infty}^{\infty} \mathcal{P}_{1l},$$

that is, the set of all distributions in $\mathcal{P}_1$ whose support is bounded from below.

As we are also going to approximate distributions with functions that are not necessarily distributions themselves, that is, they are not necessarily non-negative and do not necessarily sum to one, we shall also introduce analogous classes of such functions. Whereas we denote classes of distributions with $\mathcal{P}$, we shall apply $\mathcal{F}$ for analogous classes of functions, so that $\mathcal{F}_1$ will denote the class of all functions on the integers.

For all integers l, we let $\mathcal{F}_{1l}$ denote the set of all functions $f \in \mathcal{F}_1$ for which $f(x) = 0$ for all integers $x < l$, and we let $\mathcal{F}_{1\underline{l}}$ denote the set of functions $f \in \mathcal{F}_{1l}$ with a positive mass at l. Note that we do not have $\mathcal{F}_{1\underline{l}}$ equal to $\mathcal{F}_{1l} \sim \mathcal{F}_{1,l+1}$ as the latter class also contains the functions in $\mathcal{F}_{1l}$ with a negative mass at l.

Finally, we introduce $\mathcal{F}_{1\_} = \bigcup_{l=-\infty}^{\infty} \mathcal{F}_{1\underline{l}}$.

1.4 Convolutions

The *convolution* of a finite number of distributions is the distribution of the sum of that number of independent random variables having these distributions.

If $f, g \in \mathcal{P}_1$, then we denote their convolution by $f * g$. If X and Y are independent random variables with distribution f and g respectively, then, for all integers z,

$$\begin{aligned}(f * g)(z) &= \Pr(X + Y = z) = \Pr\left(\bigcup_{x=-\infty}^{\infty} ((X = x) \cap (Y = z - x))\right) \\ &= \sum_{x=-\infty}^{\infty} \Pr((X = x) \cap (Y = z - x)) = \sum_{x=-\infty}^{\infty} \Pr(X = x) \Pr(Y = z - x) \\ &= \sum_{x=-\infty}^{\infty} f(x) g(z - x).\end{aligned}$$

When $f, g \in \mathcal{P}_{10}$, this reduces to the finite summation

$$(f * g)(z) = \sum_{x=0}^{z} f(x) g(z - x) \qquad (z = 0, 1, 2, \ldots) \tag{1.2}$$

as $f(x) = 0$ when $x < 0$, and $g(z - x) = 0$ when $x > z$; in particular, this gives

$$(f * g)(0) = f(0)g(0). \tag{1.3}$$

The *M-fold convolution* of a distribution is defined as the distribution of the sum of M independent random variables with the original distribution. The M-fold convolution of f is denoted by f^{M*}, and we have $f^{M*} = f^{(M-1)*} * f$ for $M = 1, 2, \ldots$ with f^{0*} being the distribution concentrated in zero. In particular, we have $f^{1*} = f$. When $f \in \mathcal{P}_{10}$, (1.3) gives

$$f^{M*}(0) = f(0)^M. \tag{1.4}$$

In terms of cumulative distributions we define the convolution $F * G$ of two univariate distributions F and G by the Lebesgue–Stieltjes integral

$$(F * G)(z) = \int_{(-\infty,\infty)} G(z - x)\,\mathrm{d}F(x), \qquad (-\infty < z < \infty)$$

and we have $F^{M*} = F^{(M-1)*} * F$ for $M = 1, 2, \ldots$ with F^{0*} being the distribution concentrated in zero.

We define the convolution of two functions $f, g \in \mathcal{F}_1$ by (1.2); the M-fold convolution f^{M*} is defined accordingly with f^{0*} being concentrated in zero with mass one.

1.5 Mixed Distributions

Let us assume that in an automobile insurance portfolio with independent policies, we have reason to believe that for each policy, the number of claims occurring during a policy year has a parametric distribution with a real-valued parameter θ where a higher value of θ indicates that we can expect more claims. If we also assumed the value of this parameter to be the same for all the policies, then it could be estimated from portfolio data. However, we know that some drivers are good and are likely to have few claims whereas others are bad and are likely to have lots of claims. It is then natural to assume that a good driver has a low value of θ and a bad driver has a high value. When we gradually get experience with the individual policy, the claims data can give an indication on whether it is a good driver or a bad driver, that is, whether the value of θ is low or high. In such situations, it can be reasonable to apply *experience rating*, that is, the premium depends on the past claims experience of the individual policy. However, even for an old policy, it would be rather hazardous to base the estimate of θ solely on the claims experience of that policy, and for a new policy, we do not have any claims experience at all, so what should we do then? On the one hand, we know that there are differences between the policies reflected in their values of θ. On the other hand, we have a portfolio with lots of such policies; wouldn't there be any value in all the claims experience we have got from the portfolio when we are going to rate a new policy? One way to relate data from each policy to other policies, is to consider the parameters θ as values of independent and identically distributed random variables. In this setting, the unconditional distribution is called a *mixed distribution*. The random parameter is called the *mixing variable* and its distribution the *mixing distribution*. The portfolio data could be applied to estimate the mixing distribution and its parameters, in particular, its mean. The mixing distribution represents the distribution of the random parameters among all possible policies. We know that the policies have different parameter values; when drawing a policy at random from a representative sample of policies, then the mixing distribution represents the probability that its parameter value will be within some given interval. In this example, the mixing distribution has a frequentist interpretation; it represents the risk structure of the portfolio and is therefore often called the *structure distribution*. Its parameters are called *structure parameters*, and the random parameter is called a *structure variable*. This setting is within the framework of *empirical Bayes theory*; each unit (policy) has a random risk parameter, and we can estimate the distribution of this parameter from a portfolio of independent units.

A mixed distribution can also arise within a pure *Bayesian* (subjectivistic) context. A Bayesian would have a subjective opinion on the likeliness of the value of the parameter, and he would represent that opinion by assigning a distribution to the parameter.

In an insurance context, there should be room for both ways of modeling. In an automobile insurance portfolio, there could be lots of similar policies from which one could estimate the structure parameters, so that one can afford to be a frequentist. On the other hand, when setting the premium for a rather special type of ship with which one does not have any experience, one has to rely on subjective judgement.

Let U denote the mixing distribution, A its range, and Θ the mixing variable of our policy. The conditional claim number distribution given that $\Theta = \theta$ is p_θ for each $\theta \in A$. Then the unconditional distribution of the policy is the mixed distribution p given by

$$p(n) = \int_A p_\theta(n)\,\mathrm{d}U(\theta) = \mathrm{E}\,p_\Theta(n). \qquad (n = 0, 1, 2, \ldots)$$

More generally, if F_θ is a univariate distribution for all values of $\theta \in A$, then we can construct the mixed distribution F given by

$$F(x) = \int_A F_\theta(x)\,\mathrm{d}U(\theta) = \mathrm{E}\,F_\Theta(x). \qquad (-\infty < x < \infty)$$

The distributions F_θ would usually belong to a parametric class. If these distributions have a particular name, then one would name the mixed distribution by putting *mixed* in front of that name. For instance, if all the F_θs are Poisson distributions, then one would call F a mixed Poisson distribution.

Above, we modelled dependence between claim numbers from different years for a fixed policy by assigning a random risk parameter to this policy and assuming that these claim numbers are conditionally independent given this parameter. In a similar way, we can apply mixed distributions for modelling dependence between policies within an insurance portfolio. Let us mention two cases:

1. We can model dependence between policies within a group of policies, say, policies from the same district, by assigning a random risk parameter to this group and assuming that these policies are conditionally independent given this parameter.
2. Some phenomena, e.g. hurricanes, can affect many policies within a year. We can model such dependence between policies within that year by assigning a random risk parameter to the year and assuming that the policies are conditionally independent given this parameter.

1.6 Compound Distributions and Functions

The compound distribution $p \vee H$ with counting distribution $p \in \mathcal{P}_{10}$ and univariate severity distribution H is the distribution of $X = Y_{\bullet N}$ where $Y_1, Y_2, \ldots$ are independent and identically distributed with distribution H and independent of N, which has distribution p. In an insurance context, N could be the number of claims occurring in an insurance portfolio during a specified period and Y_i the ith of these claims for $i = 1, 2, \ldots, N$. We have

$$\begin{aligned}(p \vee H)(x) &= \Pr(X \le x) = \Pr\left(\bigcup_{n=0}^{\infty}((Y_{\bullet n} \le x) \cap (N = n))\right) \\ &= \sum_{n=0}^{\infty} \Pr((Y_{\bullet n} \le x) \cap (N = n)) = \sum_{n=0}^{\infty} \Pr(N = n) \Pr(Y_{\bullet n} \le x | N = n) \\ &= \sum_{n=0}^{\infty} p(n) \Pr(Y_{\bullet n} \le x) = \sum_{n=0}^{\infty} p(n) H^{n*}(x),\end{aligned}$$

that is,

$$p \vee H = \sum_{n=0}^{\infty} p(n) H^{n*}. \tag{1.5}$$

From this expression, we see that a compound distribution can be interpreted as a mixed distribution with the counting distribution as mixing distribution.

The counting distribution of a compound distribution typically belongs to a parametric class where it has a particular name. In that case, one would name the compound distribution by putting *compound* in front of that name. For instance, if the counting distribution is a Poisson distribution, then one would call the compound distribution a compound Poisson distribution.

If $h \in \mathcal{P}_1$, then

$$p \vee h = \sum_{n=0}^{\infty} p(n) h^{n*}. \tag{1.6}$$

When $h \in \mathcal{P}_{10}$, insertion of (1.4) gives

$$(p \vee h)(0) = \sum_{n=0}^{\infty} p(n) h(0)^n. \tag{1.7}$$

When $h \in \mathcal{P}_{11}$, then each Y_i is greater than or equal to one, so that $\sum_{i=1}^{n} Y_i \geq n$. This implies that $h^{n*}(x) = \Pr(Y_{\bullet n} = x) = 0$ when $n > x$. Hence,

$$(p \vee h)(x) = \sum_{n=0}^{x} p(n) h^{n*}(x), \qquad (x = 0, 1, 2, \dots) \tag{1.8}$$

so that we avoid the infinite summation in (1.6). In particular, we obtain that

$$(p \vee h)(0) = p(0). \tag{1.9}$$

For $p \in \mathcal{F}_{10}$ and $h \in \mathcal{F}_1$, we define $p \vee h$ by (1.6), provided that the summation exists.

1.7 Some Useful Transforms

1.7.1 Definitions and General Results

For a univariate distribution F, we introduce the (*probability*) *generating function* τ_F, the *moment generating function* ω_F, the *Laplace transform* γ_F, the *characteristic function* ζ_F, and the *cumulant generating function* θ_F. To be able to interpret these functions in terms of random variables, we also introduce a random variable X with distribution F.

For real numbers s, we have

$$\tau_F(s) = \mathrm{E}\, s^X = \int_{(-\infty,\infty)} s^x \,\mathrm{d}F(x) \tag{1.10}$$

$$\omega_F(s) = \mathrm{E}\, \mathrm{e}^{sX} = \int_{(-\infty,\infty)} \mathrm{e}^{sx} \,\mathrm{d}F(x) = \tau_F(\mathrm{e}^s)$$

$$\gamma_F(s) = \mathrm{E}\, \mathrm{e}^{-sX} = \int_{(-\infty,\infty)} \mathrm{e}^{-sx} \,\mathrm{d}F(x) = \omega_F(-s) = \tau_F(\mathrm{e}^{-s})$$

$$\zeta_F(s) = \mathrm{E}\, \mathrm{e}^{\mathrm{i}sX} = \int_{(-\infty,\infty)} \mathrm{e}^{\mathrm{i}sx} \,\mathrm{d}F(x)$$

$$\theta_F(s) = \ln \omega_F(s) = \ln \tau_F(\mathrm{e}^s). \tag{1.11}$$

When they exist, each of these functions defines the distribution F uniquely. The generating function is normally applied only for distributions in $\mathcal{P}_{10}$ and the Laplace transform only for distributions on the non-negative numbers.

As $\mathrm{e}^{\mathrm{i}sx} = \cos sx + \mathrm{i}\sin sx$ is bounded, the characteristic function exists for all (measurable) distributions. This is also the case for Laplace transforms and generating functions if we restrict to distributions on the non-negative numbers and $s \geq 0$. In connection with characterisations, this is an advantage of these functions.

For any non-negative integer j and any real number c, we denote the jth order *moment* of F around c by $\mu_F(j;c)$, that is,

$$\mu_F(j;c) = \mathrm{E}(X-c)^j = \int_{(-\infty,\infty)} (x-c)^j \,\mathrm{d}F(x).$$

In particular, we have $\mu_F(0;c) = 1$ for any value of c. For convenience, we let $\mu_F(j) = \mu_F(j;0)$. This is the jth order *non-central moment* of F, whereas $\mu_F(j;\mu_F(1))$ is the jth order *central moment* of F.

For all non-negative integers j, we have $\omega_F^{(j)}(s) = \mathrm{E}\,X^j \mathrm{e}^{sX}$. In particular, with $s = 0$, this gives

$$\mu_F(j) = \omega_F^{(j)}(0). \tag{1.12}$$

Thus, if F possesses moments of all orders, then

$$\omega_F(s) = \sum_{j=0}^{\infty} \frac{\mu_F(j)}{j!} s^j. \tag{1.13}$$

For all non-negative integers j, we define the jth order *factorial moment* $\nu_F(j)$ of F by $\nu_F(j) = \mathrm{E}\,X^{(j)}$. We have $\tau_F^{(j)}(s) = \mathrm{E}\,X^{(j)} s^{X-j}$, which in particular gives

$$\nu_F(j) = \tau_F^{(j)}(1). \tag{1.14}$$

Thus, if F possesses moments of all orders, then

$$\tau_F(s+1) = \sum_{j=0}^{\infty} \frac{\nu_F(j)}{j!} s^j. \tag{1.15}$$

For all non-negative integers j, the jth order *cumulant* $\kappa_F(j)$ of F is defined by $\kappa_F(j) = \theta_F^{(j)}(0)$; in particular, we have $\kappa_F(0) = 0$. Hence, if F possesses cumulants of all orders, then

$$\theta_F(s) = \sum_{j=1}^{\infty} \frac{\kappa_F(j)}{j!} s^j.$$

The following theorem gives a recursion for the non-central moments of F in terms of its cumulants.

Theorem 1.1 *For any univariate distribution F for which ω_F exists, we have the recursion*

$$\mu_F(j) = \sum_{u=1}^{j} \binom{j-1}{u-1} \kappa_F(u)\mu_F(j-u). \qquad (j = 1, 2, \dots) \tag{1.16}$$

Proof For any univariate distribution F for which ω_F exists, (1.11) gives

$$\theta_F'(s) = \frac{\mathrm{d}}{\mathrm{d}s} \ln \omega_F(s) = \frac{\omega_F'(s)}{\omega_F(s)}.$$

Hence,

$$\begin{aligned}
\sum_{j=1}^{\infty} \frac{\mu_F(j)}{(j-1)!} s^j &= s\omega_F'(s) = s\theta_F'(s)\omega_F(s) = \sum_{u=1}^{\infty} \frac{\kappa_F(u)}{(u-1)!} s^u \sum_{j=0}^{\infty} \frac{\mu_F(j)}{j!} s^j \\
&= \sum_{j=1}^{\infty} \sum_{u=1}^{j} \frac{\kappa_F(u)}{(u-1)!} \frac{\mu_F(j-u)}{(j-u)!} s^j \\
&= \sum_{j=1}^{\infty} \left(\frac{1}{(j-1)!} \sum_{u=1}^{j} \binom{j-1}{u-1} \kappa_F(u)\mu_F(j-u) \right) s^j.
\end{aligned}$$

Comparison of coefficients gives (1.16). □

Solving (1.16) for $\kappa_F(j)$ gives the recursion

$$\kappa_F(j) = \mu_F(j) - \sum_{u=1}^{j-1} \binom{j-1}{u-1} \kappa_F(u)\mu_F(j-u). \quad (j = 1, 2, \dots) \tag{1.17}$$

In particular, we obtain $\kappa_F(1) = \mu_F(1)$ and $\kappa_F(2) = \mu_F(2) - \mu_F(1)^2$, that is, the first and second order cumulants are equal to the mean and variance.

1.7.2 Convolutions

Let X and Y be independent random variable with distribution F and G respectively. Then

$$\zeta_{F*G}(s) = \mathrm{E}\,\mathrm{e}^{\mathrm{i}s(X+Y)} = \mathrm{E}\,\mathrm{e}^{\mathrm{i}sX}\,\mathrm{E}\,\mathrm{e}^{\mathrm{i}sY},$$

that is,

$$\zeta_{F*G} = \zeta_F \zeta_G. \tag{1.18}$$

Analogously, we get

$$\omega_{F*G} = \omega_F \omega_G \tag{1.19}$$

$$\tau_{F*G} = \tau_F \tau_G \tag{1.20}$$

$$\gamma_{F*G} = \gamma_F \gamma_G,$$

but

$$\theta_{F*G} = \ln \omega_{F*G} = \ln \omega_F \omega_G = \ln \omega_F + \ln \omega_G = \theta_F + \theta_G,$$

that is,

$$\kappa_{F*G}(j) = \kappa_F(j) + \kappa_G(j). \qquad (j = 0, 1, 2, \dots) \tag{1.21}$$

For all non-negative integers j and constants c and d, we have

$$\begin{aligned}\mu_{F*G}(j; c+d) &= \mathrm{E}(X+Y-c-d)^j = \mathrm{E}((X-c)+(Y-d))^j \\ &= \sum_{i=0}^{j} \binom{j}{i} \mathrm{E}(X-c)^i \, \mathrm{E}(Y-d)^{j-i},\end{aligned}$$

that is,

$$\mu_{F*G}(j; c+d) = \sum_{i=0}^{j} \binom{j}{i} \mu_F(i; c) \mu_G(j-i; d). \tag{1.22}$$

1.7.3 Discrete Distributions

For $f \in \mathcal{P}_1$, we have

$$\tau_f(s) = \sum_{x=-\infty}^{\infty} s^x f(x); \qquad \omega_f(s) = \sum_{x=-\infty}^{\infty} \mathrm{e}^{sx} f(x) \tag{1.23}$$

$$\gamma_f(s) = \sum_{x=-\infty}^{\infty} \mathrm{e}^{-sx} f(x); \qquad \zeta_f(s) = \sum_{x=-\infty}^{\infty} \mathrm{e}^{\mathrm{i}sx} f(x); \qquad \theta_f(s) = \ln \omega_f(s). \tag{1.24}$$

For all non-negative integers j, we introduce the moments

$$\mu_f(j; c) = \sum_{x=-\infty}^{\infty} (x-c)^j f(x); \qquad \mu_f(j) = \sum_{x=-\infty}^{\infty} x^j f(x), \tag{1.25}$$

the factorial moments

$$\nu_f(j) = \sum_{x=-\infty}^{\infty} x^{(j)} f(x) \tag{1.26}$$

and the cumulants

$$\kappa_f(j) = \theta_f^{(j)}(0). \tag{1.27}$$

If $f \in \mathcal{P}_{10}$, then, for all non-negative integers x,

$$\tau_f^{(x)}(s) = \mathrm{E}\, X^{(x)} s^{X-x} = \sum_{y=x}^{\infty} y^{(x)} s^{y-x} f(y).$$

In particular, we obtain

$$\tau_f^{(x)}(0) = \mathrm{E}\, X^{(x)} I(X = x) = x! f(x),$$

so that

$$f(x) = \frac{\tau_f^{(x)}(0)}{x!}. \tag{1.28}$$

If, in addition, $f(x) = 0$ for all integers x greater than some positive integer r, then $\nu_f(j) = 0$ for $j = r+1, r+2, \ldots$.

1.7.4 Compound Distributions

We shall now consider the compound distribution $p \vee H$ with counting distribution $p \in \mathcal{P}_{10}$ and univariate severity distribution H. Let $Y_1, Y_2, \ldots$ be independent and identically distributed random variables with distribution H independent of N, which has distribution p. Then

$$\zeta_{p\vee H}(s) = \mathrm{E}\, \mathrm{e}^{\mathrm{i}sY_{\bullet N}} = \mathrm{E}\,\mathrm{E}\left[\prod_{j=1}^{N} \mathrm{e}^{\mathrm{i}sY_j} \middle| N\right] = \mathrm{E}(\mathrm{E}\, \mathrm{e}^{\mathrm{i}sY})^N,$$

that is,

$$\zeta_{p\vee H}(s) = \tau_p(\zeta_H(s)). \tag{1.29}$$

Analogously, we have

$$\tau_{p\vee H}(s) = \tau_p(\tau_H(s)) \tag{1.30}$$

$$\omega_{p\vee H}(s) = \tau_p(\omega_H(s)) \tag{1.31}$$

$$\gamma_{p\vee H}(s) = \tau_p(\gamma_H(s)).$$

We have

$$\mu_{p\vee H}(1) = \mathrm{E}\, Y_{\bullet N} = \mathrm{E}\,\mathrm{E}[Y_{\bullet N}|N] = \mathrm{E}[N\, \mathrm{E}\, Y] = \mathrm{E}\, N\, \mathrm{E}\, Y,$$

that is,

$$\mu_{p\vee H}(1) = \mu_p(1)\mu_H(1), \tag{1.32}$$

which could also be deduced by derivation of (1.29).

For $h \in \mathcal{P}_{10}$, we can write (1.7) as

$$(p \vee h)(0) = \tau_p(h(0)). \tag{1.33}$$

1.7.5 Extension to Functions

We extend the definitions (1.23)–(1.27) to functions $f \in \mathcal{F}_1$. For such functions, we do not necessarily have $\mu_f(0; c) = 1$ and $\kappa_f(0) = 0$. However, the other properties we have shown, still hold, provided that the functions exist. For convenience, we also introduce

$$\mu_f(-1) = \sum_{x=1}^{\infty} \frac{f(x)}{x}. \qquad (f \in \mathcal{F}_{11})$$

1.8 Some Useful Operators

The operator Φ is applied to univariate functions and gives the function multiplied by its argument, that is, if f is a univariate function, then $\Phi f(x) = xf(x)$ for all x. We easily see that

$$\tau_{\Phi f} = \Phi(\tau_f'). \tag{1.34}$$

The operator Ψ is applied to functions in $\mathcal{F}_{11}$ and gives the function divided by its argument, that is, if $f \in \mathcal{F}_{11}$, then $\Psi f \in \mathcal{F}_{11}$ is given by

$$\Psi f(x) = \frac{f(x)}{x}. \qquad (x = 1, 2, \dots)$$

We obviously have that $\Phi\Psi f = \Psi\Phi f = f$.

We introduce the *cumulation operator* Γ and the *tail operator* Λ defined by $\Gamma f(x) = \sum_{y=-\infty}^{x} f(y)$ and $\Lambda f(x) = \sum_{y=x+1}^{\infty} f(y)$ respectively for all integers x and all functions $f \in \mathcal{F}_1$ for which the summations exist. When $f \in \mathcal{F}_{1\_}$, Γf always exists as then there will always exist an integer l such that $\Gamma f(x) = \sum_{y=l}^{x} f(y)$ for all integers $x \geq l$. We shall also need the *difference operator* Δ defined by $\Delta f(x) = f(x) - f(x-1)$ for all integers x and all functions $f \in \mathcal{F}_1$.

For all integers x, we have $\Gamma f(x) + \Lambda f(x) = \mu_f(0)$ when these functions exist. We have

$$f = \Delta\Gamma f = -\Delta\Lambda f \tag{1.35}$$

when the functions involved exist.

Note that even if $f \in \mathcal{P}_{10}$, that is not necessarily the case with Λf as then we have that for all negative integers x,

$$\Lambda f(x) = \Lambda f(-1) = \Lambda f(0) + f(0) = \sum_{x=0}^{\infty} f(x).$$

If $f \in \mathcal{P}_1$, then Γf is the corresponding cumulative distribution function and Λf is the tail.

Theorem 1.2 *For all* $f \in \mathcal{F}_{10}$, *we have*

$$\Gamma^t f(x) = \sum_{y=0}^{x} \binom{x-y+t-1}{t-1} f(y). \quad (x = 0, 1, 2, \dots; t = 1, 2, \dots) \tag{1.36}$$

Proof It is immediately seen that application of the operator Γ is equivalent with convolution with the function $\gamma \in \mathcal{F}_{10}$ given by

$$\gamma(x) = 1. \qquad (x = 0, 1, 2, \dots) \tag{1.37}$$

Thus, $\Gamma^t f = \gamma^{t*} * f$. For $|s| < 1$, we have

$$\tau_\gamma(s) = \sum_{x=0}^{\infty} s^x = (1-s)^{-1},$$

from which we obtain

$$\tau_{\gamma^{t*}}(s) = \tau_\gamma(s)^t = (1-s)^{-t} = \sum_{x=0}^{\infty} \binom{t+x-1}{x} s^x = \sum_{x=0}^{\infty} \binom{t+x-1}{t-1} s^x,$$

so that

$$\gamma^{t*}(x) = \binom{t+x-1}{t-1}. \qquad (x = 0, 1, 2, \dots)$$

This gives

$$\Gamma^t f(x) = (\gamma^{t*} * f)(x) = \sum_{y=0}^{x} \gamma^{t*}(x-y) f(y) = \sum_{y=0}^{x} \binom{t+x-y-1}{t-1} f(y),$$

that is, (1.36) holds. □

To derive a similar result for $\Lambda^t f$, we need the following lemma.

Lemma 1.1 *For all* $f \in \mathcal{F}_{10}$, *we have*

$$\nu_{\Lambda f}(j-1) = \frac{\nu_f(j)}{j} \tag{1.38}$$

for all positive integers j *for which* $\nu_f(j)$ *exists.*

Proof Let j be a positive integer for which $\nu_f(j)$ exists.

We shall first show by induction that

$$\sum_{x=0}^{y-1} x^{(j-1)} = \frac{y^{(j)}}{j}. \qquad (y = 1, 2, \dots) \tag{1.39}$$

This equality obviously holds for $y = 1$. Let us now assume that it holds for $y = z - 1$ for some integer $z > 1$. Then

$$\sum_{x=0}^{z-1} x^{(j-1)} = \sum_{x=0}^{z-2} x^{(j-1)} + (z-1)^{(j-1)} = \frac{(z-1)^{(j)}}{j} + (z-1)^{(j-1)}$$
$$= (z-1)^{(j-1)}\left(\frac{z-j}{j} + 1\right) = (z-1)^{(j-1)}\frac{z}{j} = \frac{z^{(j)}}{j}.$$

Thus, (1.39) holds also for $y = z$, and by induction it then holds for all positive integers y.

Application of (1.39) gives

$$\nu_{\Lambda f}(j-1) = \sum_{x=1}^{\infty} x^{(j-1)} \Lambda f(x) = \sum_{x=0}^{\infty} x^{(j-1)} \sum_{y=x+1}^{\infty} f(y)$$
$$= \sum_{y=1}^{\infty} f(y) \sum_{x=0}^{y-1} x^{(j-1)} = \sum_{y=1}^{\infty} f(y) \frac{y^{(j)}}{j} = \frac{\nu_f(j)}{j}.$$

□

The following theorem follows by repeated application of (1.38).

Theorem 1.3 *For all $f \in \mathcal{F}_{10}$, we have*

$$\Lambda^t f(-1) = \nu_{\Lambda^{t-1} f}(0) = \frac{\nu_f(t-1)}{(t-1)!} \tag{1.40}$$

for all positive integers t for which $\Lambda^t f$ exists.

We shall prove the following corollary by translation.

Corollary 1.1 *For all $f \in \mathcal{F}_1$ and all integers x, we have*

$$\Lambda^t f(x) = \sum_{y=t+x}^{\infty} \binom{y-x-1}{t-1} f(y) \tag{1.41}$$

for all positive integers t for which $\Lambda^t f$ exists.

Proof For $x = -1$, (1.41) follows easily from (1.40) if $f \in \mathcal{F}_{10}$. This result obviously holds even if we drop the restriction $f \in \mathcal{F}_{10}$ as $\Lambda^t f(-1)$ does not depend on $f(x)$ for negative integers x.

For all integers x, y, let

$$f_x(y) = f(y + x + 1).$$

Then

$$\begin{aligned}\Lambda^t f(x) = \Lambda^t f_x(-1) &= \sum_{y=t-1}^{\infty} \binom{y}{t-1} f_x(y) = \sum_{y=t-1}^{\infty} \binom{y}{t-1} f(y + x + 1) \\ &= \sum_{y=t+x}^{\infty} \binom{y-x-1}{t-1} f(y),\end{aligned}$$

that is, (1.41) holds for all integers x. □

1.9 Stop Loss Premiums

Let X be the aggregate claims of an insurance portfolio and F its distribution. As explained in Sect. 1.1, if this portfolio is covered by stop loss reinsurance with retention x, then the reinsurer pays $(X - x)_+$ and the cedant pays $\min(X, x)$. Let $\Pi_F(x) = \mathrm{E}(X - x)_+$ and $\Omega_F(x) = \mathrm{E}\min(X, x)$. We call the functions Π_F and Ω_F the *stop loss transform* and *retention transform* of F. We define these functions for all real numbers and all distributions on the real numbers although in reinsurance applications, the distribution will normally be restricted to the non-negative numbers. For simplicity, we assume that $\mu_F(1)$ exists and is finite.

We have

$$\begin{aligned}\Pi_F(x) = \mathrm{E}(X - x)_+ &= \int_x^{\infty} (y - x)\, \mathrm{d}F(y) = \int_x^{\infty} \int_x^{y} \mathrm{d}z\, \mathrm{d}F(y) \\ &= \int_x^{\infty} \int_{(z,\infty)} \mathrm{d}F(y)\, \mathrm{d}z,\end{aligned}$$

that is,

$$\Pi_F(x) = \int_x^{\infty} (1 - F(z))\, \mathrm{d}z. \tag{1.42}$$

Letting $x \downarrow -\infty$ gives $\mu_F(1) = \int_{-\infty}^{\infty} (1 - F(z))\, \mathrm{d}z$. Thus,

$$\begin{aligned}\Omega_F(x) = \mu_F(1) - \Pi_F(x) &= \int_{-\infty}^{\infty} (1 - F(z))\, \mathrm{d}z - \int_x^{\infty} (1 - F(z))\, \mathrm{d}z \\ &= \int_{-\infty}^{x} (1 - F(z))\, \mathrm{d}z.\end{aligned}$$

We shall now deduce some results on stop loss ordering. We start with mixed distributions.

Theorem 1.4 *For all values of θ in a set A, let F_θ and G_θ be univariate distributions such that $\Pi_{F_\theta} \leq \Pi_{G_\theta}$, and let U be a mixing distribution on A. Then the mixed distributions F and G given by $F = \int_A F_\theta \, \mathrm{d}U(\theta)$ and $G = \int_A G_\theta \, \mathrm{d}U(\theta)$ satisfy the inequality $\Pi_F \leq \Pi_G$.*

Proof For any real number x, we have

$$\begin{aligned}\Pi_F(x) &= \int_x^\infty (1 - F(z))\,\mathrm{d}z = \int_x^\infty \int_A (1 - F_\theta(z))\,\mathrm{d}U(\theta)\,\mathrm{d}z \\ &= \int_A \int_x^\infty (1 - F_\theta(z))\,\mathrm{d}z\,\mathrm{d}U(\theta) = \int_A \Pi_{F_\theta}(x)\,\mathrm{d}U(\theta) \leq \int_A \Pi_{G_\theta}(x)\,\mathrm{d}U(\theta) \\ &= \Pi_G(x),\end{aligned}$$

which proves the theorem. □

We now turn to convolutions.

Lemma 1.2 *Let F, G, and H be univariate distributions such that $\Pi_F \leq \Pi_G$. Then $\Pi_{F*H} \leq \Pi_{G*H}$.*

Proof For any real number x, we have

$$\begin{aligned}\Pi_{F*H}(x) &= \int_x^\infty (1 - (F*H)(z))\,\mathrm{d}z = \int_x^\infty \int_{(-\infty,\infty)} (1 - F(z-y))\,\mathrm{d}H(y)\,\mathrm{d}z \\ &= \int_{(-\infty,\infty)} \int_x^\infty (1 - F(z-y))\,\mathrm{d}z\,\mathrm{d}H(y) = \int_{(-\infty,\infty)} \Pi_F(x-y)\,\mathrm{d}H(y) \\ &\leq \int_{(-\infty,\infty)} \Pi_G(x-y)\,\mathrm{d}H(y) = \Pi_{G*H}(x),\end{aligned}$$

which proves the lemma. □

Theorem 1.5 *For $j = 1, 2, \ldots$, let F_j and G_j be univariate distributions such that $\Pi_{F_j} \leq \Pi_{G_j}$. Then*

$$\Pi_{*_{j=1}^M F_j} \leq \Pi_{*_{j=1}^M G_j}. \qquad (M = 0, 1, 2, \ldots) \tag{1.43}$$

Proof The inequality (1.43) obviously holds for $M = 0$. Let us now assume that it holds for $M = k - 1$ for some positive integer k. Then the induction hypothesis and two applications of Lemma 1.2 give

$$\Pi_{*_{j=1}^k F_j} = \Pi_{(*_{j=1}^{k-1} F_j)*F_k} \leq \Pi_{(*_{j=1}^{k-1} F_j)*G_k} \leq \Pi_{(*_{j=1}^{k-1} G_j)*G_k} = \Pi_{*_{j=1}^k G_j},$$

so that the induction hypothesis holds also for $M = k$. By induction, it then holds for all positive integers M.

This completes the proof of Theorem 1.5. □

The following corollary follows immediately from Theorem 1.5 by letting $F_j = F$ and $G_j = G$ for $j = 1, 2, \ldots$.

Corollary 1.2 *If F and G are univariate distributions such that $\Pi_F \leq \Pi_G$, then*

$$\Pi_{F^{M*}} \leq \Pi_{G^{M*}}. \qquad (M = 0, 1, 2, \ldots)$$

The following corollary follows by application of Theorem 1.4 and Corollary 1.2.

Corollary 1.3 *If $p \in \mathcal{P}_{10}$ and F and G are univariate distributions such that $\Pi_F \leq \Pi_G$, then $\Pi_{p \vee F} \leq \Pi_{p \vee G}$.*

Theorem 1.6 *If $F_1, F_2, \ldots$ are univariate distributions, then*

$$\Pi_{*_{j=1}^{M} F_j} \geq \sum_{j=1}^{M} \Pi_{F_j}. \qquad (M = 1, 2, \ldots) \tag{1.44}$$

Proof The inequality (1.44) obviously holds for $M = 1$. Let us now assume that it holds for $M = k - 1$ for some integer $k > 1$, and let Y and Z be independent random variables with distribution $*_{j=1}^{k-1} F_j$ and F_k respectively. For any numbers x, y, and z, we have

$$(y + z - x)_+ \geq (y - x)_+ + (z - x)_+.$$

This gives that for all numbers x, we have

$$\begin{aligned} \Pi_{*_{j=1}^{k} F_j}(x) &= \mathrm{E}(Y + Z - x)_+ \geq \mathrm{E}(Y - x)_+ + \mathrm{E}(Z - x)_+ \\ &= \Pi_{*_{j=1}^{k-1} F_j}(x) + \Pi_{F_k}(x) \geq \sum_{j=1}^{k} \Pi_{F_j}(x); \end{aligned}$$

the last inequality follows by the induction hypothesis (1.44). Hence, (1.44) holds also for $M = k$, and by induction, it then holds for all positive integers M. □

The following corollary follows immediately from Theorem 1.6 by letting $F_j = F$ for $j = 1, 2, \ldots$.

Corollary 1.4 *If F is a univariate distribution, then*

$$\Pi_{F^{M*}} \geq M \Pi_F. \qquad (M = 1, 2, \ldots)$$

Theorem 1.7 *If $p \in \mathcal{P}_{10}$ and H is a univariate distribution, then $\Pi_{p \vee H} \geq \mu_p(1)\Pi_H$.*

Proof Application of Corollary 1.4 gives

$$\Pi_{p \vee H} = \Pi_{\sum_{n=0}^{\infty} p(n) H^{n*}} = \sum_{n=0}^{\infty} p(n) \Pi_{H^{n*}} \geq \sum_{n=0}^{\infty} p(n) n \Pi_H = \mu_p(1)\Pi_H.$$ □

In the following theorem, we apply the stop loss transform of the same distribution to obtain an upper and a lower bound of the stop loss transform of another distribution.

Theorem 1.8 *Let F and G be univariate distributions with finite mean, and assume that $F \leq G$. Then*

$$\Pi_G \leq \Pi_F \leq \Pi_G + \varepsilon \tag{1.45}$$

$$\Pi_F - \varepsilon \leq \Pi_G \leq \Pi_F \tag{1.46}$$

with $\varepsilon = \mu_F(1) - \mu_G(1)$.

Proof As $F \leq G$, (1.42) immediately gives that $\Pi_G \leq \Pi_F$.

On the other hand, for any real number x, we have

$$\begin{aligned}\Pi_F(x) &= \mu_F(1) - \int_{-\infty}^{x} (1 - F(z))\,\mathrm{d}z \leq \mu_F(1) - \int_{-\infty}^{x} (1 - G(z))\,\mathrm{d}z \\ &= \mu_F(1) - \mu_G(1) + \Pi_G(x),\end{aligned}$$

so that also the second inequality in (1.45) holds.

We easily obtain (1.46) by rearranging (1.45).

This completes the proof of Theorem 1.8. □

Now let $f \in \mathcal{P}_1$ and assume that $\mu_f(1)$ exists and is finite. Then, for all integers x,

$$\Pi_f(x) = \sum_{y=x+1}^{\infty} (y - x) f(y) = \sum_{y=x}^{\infty} \Lambda f(y) = \Lambda^2 f(x-1) \tag{1.47}$$

$$\Omega_f(x) = \sum_{y=-\infty}^{x-1} \Lambda f(y).$$

Theorem 1.2 gives

$$\Gamma^2 f(x) = \sum_{y=0}^{x} \binom{x - y + 1}{1} f(y) = \sum_{y=0}^{x} (x - y + 1) f(y) = \mathrm{E}(x + 1 - X)_+.$$

Thus,

$$\Pi_f(x) = \mathrm{E}(X-x)_+ = \mathrm{E}(x-X)_+ + \mathrm{E}\,X - x = \Gamma^2 f(x-1) + \mu_f(1) - x$$

and

$$\begin{aligned}\Omega_f(x) &= \mathrm{E}\min(X,x) = \mu_f(1) - \Pi_f(x)\\ &= \mu_f(1) - (\Gamma^2 f(x-1) + \mu_f(1) - x) = x - \Gamma^2 f(x-1).\end{aligned}$$

In some respects, distributions in $\mathcal{P}_1$, and, by change of unit, more generally arithmetic distributions, are often easier to handle than more general distribution, for instance, with respect to recursions. Hence, it can be desirable to approximate a more general distribution by an arithmetic distribution. We shall now have a look at some such approximations obtained by rounding. Let X be a random variable with distribution F and h some positive number. We introduce the approximation X_{h+} (X_{h-}) with distribution F_{h+} (F_{h-}) obtained by rounding X upwards (downwards) to the nearest whole multiple of h. Then $X_{h-} \leq X \leq X_{h+}$, and, hence, $F_{h+} \leq F \leq F_{h-}$. Theorem 1.8 gives

$$\Pi_{F_{h+}} - \varepsilon \leq \Pi_F \leq \Pi_{F_{h+}}$$

with

$$\varepsilon = \mu_{F_{h+}}(1) - \mu_F(1) = \mathrm{E}(X_{h+} - X) \leq \mathrm{E}\,h = h,$$

so that

$$\Pi_{F_{h+}} - h \leq \Pi_F \leq \Pi_{F_{h+}}.$$

Analogously, we obtain

$$\Pi_{F_{h-}} \leq \Pi_F \leq \Pi_{F_{h-}} + h.$$

Hence, we can obtain any desired accuracy of the approximation to the stop loss transform by choosing h sufficiently small.

We shall prove the following corollary to Theorem 1.8.

Corollary 1.5 *Let $f, g \in \mathcal{P}_1$ with finite mean, and assume that $\Gamma f \leq \Gamma g$. Then*

$$\Lambda g \leq \Lambda f \leq \Lambda g + \varepsilon \tag{1.48}$$

$$\Lambda f - \varepsilon \leq \Lambda g \leq \Lambda f \tag{1.49}$$

$$\Gamma g - \varepsilon \leq \Gamma f \leq \Gamma g \tag{1.50}$$

$$\Gamma f \leq \Gamma g \leq \Gamma f + \varepsilon \tag{1.51}$$

$$|f - g| \leq \varepsilon \tag{1.52}$$

with $\varepsilon = \mu_f(1) - \mu_g(1)$.

Proof The first inequality in (1.48) follows immediately from the assumption that $\Gamma f \leq \Gamma g$ by using that

$$\Lambda f = 1 - \Gamma f; \qquad \Lambda g = 1 - \Gamma g. \tag{1.53}$$

By application of (1.47) and (1.45), we obtain that for any integer x,

$$\Lambda f(x) = \Pi f(x) - \Pi f(x+1) \leq \Pi g(x) + \varepsilon - \Pi g(x+1) = \Lambda g(x) + \varepsilon,$$

so that also the second inequality in (1.48) holds.

We easily obtain (1.49) by rearranging (1.48).

We easily obtain (1.50) and (1.51) by application of (1.53) in (1.48) and (1.49).

By application of (1.50), we obtain that for any integer x,

$$f(x) = \Gamma f(x) - \Gamma f(x-1) \leq \Gamma g(x) - \Gamma g(x-1) + \varepsilon = g(x) + \varepsilon$$
$$f(x) = \Gamma f(x) - \Gamma f(x-1) \geq \Gamma g(x) - \varepsilon - \Gamma g(x-1) = g(x) - \varepsilon,$$

and these two inequalities give (1.52).

This completes the proof of Corollary 1.5. □

It should be stressed that even if Theorem 1.8 gives reasonable bounds, that is not necessarily the case with the bounds of Corollary 1.5 as the magnitude of the quantities that we bound there, is much smaller whereas ε is the same.

1.10 Convergence of Infinite Series with Positive Terms

In this section, we shall deduce some criteria for convergence of infinite series with positive terms. As the partial sums of such a series constitutes an increasing sequence, it will either converge to a positive number or diverge to infinity.

Theorem 1.9 *If l is a non-negative integer and $\{x_n\}_{n=l}^{\infty}$ a sequence of positive numbers such that $\lim_{n\uparrow\infty} x_n/x_{n-1} = c$ exists, then the infinite series $\sum_{n=l}^{\infty} x_n$ is convergent if $c < 1$ and divergent if $c > 1$.*

Proof We first assume that $c < 1$. Then, for any $d \in (c, 1)$, there exists a positive integer $m > l$ such that $x_n/x_{n-1} < d$ for all integers $n \geq m$. Hence,

$$\sum_{n=m}^{\infty} x_n \leq x_m \sum_{j=0}^{\infty} d^j = \frac{x_m}{1-d} < \infty.$$

Hence, $\sum_{n=l}^{\infty} x_n$ is convergent.

Now, assume that $c > 1$. Then, for any $d \in (1, c)$, there exists a positive integer $m > l$ such that $x_n/x_{n-1} > d$ for all integers $n \geq m$. Hence,

$$\sum_{n=m}^{\infty} x_n \geq x_m \sum_{j=0}^{\infty} d^j = \infty.$$

Hence, $\sum_{n=l}^{\infty} x_n$ is divergent.

This completes the proof of Theorem 1.9. □

Unfortunately, Theorem 1.9 does not give any clue on convergence or divergence when $c = 1$. In that case, the following theorem can be useful. In the proof, we shall use that

$$\sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

This divergence follows by observing that

$$\sum_{n=1}^{\infty} \frac{1}{n} > \sum_{n=1}^{\infty} \frac{x^n}{n} = -\ln(1-x). \qquad (0 < x < 1)$$

Theorem 1.10 *If l is a non-negative integer and $\{x_n\}_{n=l}^{\infty}$ a sequence of positive numbers such that*

$$\lim_{n\uparrow\infty} n\left(1 - \frac{x_n}{x_{n-1}}\right) = c$$

exists, then the infinite series $\sum_{n=l}^{\infty} x_n$ is convergent if $c > 1$ and divergent if $c < 1$.

Proof We first assume that $c > 1$. Then, for any $d \in (1, c)$, there exists a positive integer $m > l$ such that for all integers $n \geq m$, we have

$$n\left(1 - \frac{x_n}{x_{n-1}}\right) \geq d,$$

that is,

$$(n-1)x_{n-1} - nx_n \geq (d-1)x_{n-1} > 0.$$

Hence, the sequence $\{nx_n\}_{n=m}^{\infty}$ is decreasing, and then the limit $\lim_{n\uparrow\infty} nx_n = v$ exists and is finite. Summation gives

$$(m-1)x_{m-1} - v \geq (d-1)\sum_{n=m}^{\infty} x_{n-1}.$$

Hence, $\sum_{n=l}^{\infty} x_n$ is convergent.

Now, assume that $c < 1$. Then, for any $d \in (c, 1)$, there exists a positive integer $m > l$ such that for all integers $n \geq m$, we have

$$n\left(1 - \frac{x_n}{x_{n-1}}\right) \leq d,$$

that is,

$$(n-1)x_{n-1} - nx_n \leq (d-1)x_{n-1} < 0.$$

Hence, the sequence $\{nx_n\}_{n=m}^{\infty}$ is increasing, so that for all integers $n > m$, we have $nx_n > mx_m$. Thus,

$$\sum_{n=m}^{\infty} x_n > mx_m \sum_{n=m}^{\infty} \frac{1}{n} = \infty.$$

Hence, $\sum_{n=l}^{\infty} x_n$ is divergent.

This completes the proof of Theorem 1.10. □

Whereas Theorems 1.9 and 1.10 are standard results from the theory of infinite series, the following theorem is more tailor-made for our purpose in Chap. 2, where we apply this theorem for characterisation of classes of distributions that satisfy some specific types of recursions.

Theorem 1.11 *Let*

$$x_n = \left(a + \frac{b}{n}\right)x_{n-1} \qquad (n = l+1, l+2, \ldots) \tag{1.54}$$

with $x_l, a > 0$ and $b \geq -(l+1)a$ for some non-negative integer l. Then the infinite series $\sum_{n=l}^{\infty} x_n$ is convergent if $a < 1$ and divergent if $a > 1$. If $a = 1$, then it is convergent if $b < -1$ and divergent if $b \geq -1$.

Proof From (1.54), we immediately see that $\lim_{n\uparrow\infty} x_n/x_{n-1} = a$. Hence, Theorem 1.9 gives that $\sum_{n=l}^{\infty} x_n$ is convergent if $a < 1$ and divergent if $a > 1$.

Now let $a = 1$. Then, for $n > l$,

$$n\left(1 - \frac{x_n}{x_{n-1}}\right) = -b,$$

and Theorem 1.10 gives that $\sum_{n=l}^{\infty} x_n$ is convergent if $b < -1$ and divergent if $b > -1$.

When $a = 1$ and $b = -1$, we have

$$\sum_{n=l}^{\infty} x_n = l \sum_{n=l}^{\infty} \frac{1}{n} = \infty,$$

that is, $\sum_{n=l}^{\infty} x_n$ is divergent.

This completes the proof of Theorem 1.11. □

Further Remarks and References

For more details on modelling of the aggregate claims of insurance portfolios, see e.g. Panjer and Willmot (1992), Sundt (1999b), Kaas et al. (2001), Klugman et al. (2004), or Dickson (2005). Schröter (1995) gives an overview of methods for evaluation of aggregate claims distributions. Denuit et al. (2005) specialise on modelling dependent risks.

An extensive presentation of reinsurance is given by Gerathewohl (1980, 1983). For stop loss reinsurance, see also Rytgaard (2004).

For more details on life assurance, see e.g. Bowers et al. (1997).

For more details on experience rating, see e.g. Neuhaus (2004), Lemaire (2004), and Norberg (2004). Interesting outlines of Bayes statistics are given by DeGroot (2004) and Jewell (2004).

Theorem 1.1 was proved by Dhaene et al. (1996). The recursion (1.17) was pointed out by Sundt et al. (1998).

Antzoulakos and Chadjiconstantinidis (2004) proved Theorems 1.2 and 1.3.

Applications of Theorem 1.8 and Corollary 1.5 have been discussed by Sundt (1986a, 1986b, 1991c, 1999b). Most of the other results in Sect. 1.9 are based on Bühlmann et al. (1977). There exists an extensive literature on inequalities for stop loss transforms; Kaas et al. (1994) give an extensive presentation.

Gerber and Jones (1976), Gerber (1982), Panjer and Lutek (1983), De Vylder and Goovaerts (1988), Walhin and Paris (1998), Sundt (1999b), and Grübel and Hermesmeier (1999, 2000) discussed various ways of approximating a distribution with an arithmetic distribution.

For convergence criteria of infinite series, see e.g. Knopp (1990).

In older literature, cumulants are sometimes called semi-invariants.

In non-actuarial literature, compound distributions are sometimes called generalised or stopped sum distributions.

Chapter 2
Counting Distributions with Recursion of Order One

Summary

In Chap. 1, we defined compound distributions, presented some of their properties, and mentioned their importance in modelling aggregate claims distributions in an insurance setting. The main topic of the present chapter is recursions for compound distributions, mainly with severity distribution in $\mathcal{P}_{11}$, but in Sect. 2.7, we extend the theory to severity distributions in $\mathcal{P}_{10}$; as a special case, we consider thinning in Sect. 2.7.2.

Section 2.3 is devoted to the Panjer class of distributions $p \in \mathcal{P}_{1\underline{0}}$ that satisfy a recursion in the form

$$p(n) = \left(a + \frac{b}{n}\right) p(n-1) \qquad (n = 1, 2, \dots)$$

for some constants a and b. This class is characterised in Sect. 2.3.2. The key result of the present chapter is the Panjer recursion for compound distributions with counting distribution in the Panjer class. This recursion is motivated and deduced in Sect. 2.3.1, where we also give a continuous version. Section 2.3.3 discusses an alternative recursion that for some severity distributions is more efficient than the Panjer recursion.

To motivate the Panjer recursion and the sort of deductions that we shall mainly apply in the present book, we first discuss two special cases, geometric counting distribution in Sect. 2.1 and Poisson counting distribution in Sect. 2.2. Within the Poisson case, in Sect. 2.2.2 we also discuss an alternative way of deduction based on generating functions as well as present an alternative recursion that for some severity distributions can be more efficient than the Panjer recursion.

Section 2.5 is devoted to an extension of the Panjer class, and that class is further extended in Sect. 2.6.

Although the main emphasis is on compound distributions in the present chapter, Sect. 2.4 is devoted to recursions for convolutions of a distribution on the integers with range bounded on at least one side as these recursions are closely related to the Panjer recursion.

2.1 Geometric Distribution

Let N be a random variable with distribution $p \in \mathcal{P}_{10}$ and $Y_1, Y_2, \dots$ independent and identically distributed random variables with distribution $h \in \mathcal{P}_{11}$. It is assumed

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_2, © Springer-Verlag Berlin Heidelberg 2009

that the Y_js are independent of N. We want to evaluate the distribution f of $X = Y_{\bullet N}$, that is, $f = p \vee h$. From (1.9), we obtain the initial value

$$f(0) = p(0). \tag{2.1}$$

The simplest case is when p is the *geometric distribution* geo(π) given by

$$p(n) = (1-\pi)\pi^n. \qquad (n = 0, 1, 2, \ldots; 0 < \pi < 1) \tag{2.2}$$

Theorem 2.1 *When p is the geometric distribution* geo(π) *and $h \in \mathcal{P}_{11}$, then $f = p \vee h$ satisfies the recursion*

$$f(x) = \pi \sum_{y=1}^{x} h(y) f(x-y) \qquad (x = 1, 2, \ldots) \tag{2.3}$$

$$f(0) = 1 - \pi. \tag{2.4}$$

Proof The initial condition (2.4) follows immediately from (2.1) and (2.2).

From (2.2), we see that

$$p(n) = \pi p(n-1). \qquad (n = 1, 2, \ldots) \tag{2.5}$$

Insertion in (1.6) gives that for $x = 1, 2, \ldots$, we have

$$f(x) = \sum_{n=1}^{\infty} p(n) h^{n*}(x) = \pi \sum_{n=1}^{\infty} p(n-1)(h * h^{(n-1)*})(x)$$
$$= \pi \left(h * \left(\sum_{n=1}^{\infty} p(n-1) h^{(n-1)*} \right) \right)(x) = \pi (h * f)(x) = \pi \sum_{y=1}^{x} h(y) f(x-y),$$

which proves (2.3).

This completes the proof of Theorem 2.1. □

2.2 Poisson Distribution

2.2.1 General Recursion

We now assume that the claim number distribution p is the *Poisson distribution* Po(λ) given by

$$p(n) = \frac{\lambda^n}{n!} e^{-\lambda}, \qquad (n = 0, 1, 2, \ldots; \lambda > 0) \tag{2.6}$$

but keep the other assumptions and notation of Sect. 2.1.

Theorem 2.2 *When p is the Poisson distribution* $\mathrm{Po}(\lambda)$ *and* $h \in \mathcal{P}_{11}$, *then* $f = p \vee h$ *satisfies the recursion*

$$f(x) = \frac{\lambda}{x} \sum_{y=1}^{x} y h(y) f(x-y) \qquad (x = 1, 2, \dots) \tag{2.7}$$

$$f(0) = \mathrm{e}^{-\lambda}. \tag{2.8}$$

Proof The initial condition (2.8) follows immediately from (2.1) and (2.6).

For the recursion for the compound geometric distribution, we utilised a recursion for the counting distribution, so let us try to do something similar in the Poisson case. From (2.6), we obtain

$$p(n) = \frac{\lambda}{n} p(n-1). \qquad (n = 1, 2, \dots) \tag{2.9}$$

Insertion in (1.6) gives that for $x = 1, 2, \dots,$ we have

$$f(x) = \sum_{n=1}^{\infty} \frac{\lambda}{n} p(n-1) h^{n*}(x). \tag{2.10}$$

This one looks more awkward, but let us see what we can do. This $h^{n*}(x)$, the probability that $Y_{\bullet n} = x$, might lead to something. If we condition on that event, then the conditional expectation of each Y_j must be x/n, that is,

$$\frac{1}{n} = \mathrm{E}\left[\frac{Y_1}{x}\middle| Y_{\bullet n} = x\right] = \sum_{y=1}^{x} \frac{y}{x} \frac{h(y) h^{(n-1)*}(x-y)}{h^{n*}(x)}. \tag{2.11}$$

Insertion in (2.10) gives

$$\begin{aligned} f(x) &= \lambda \sum_{n=1}^{\infty} p(n-1) \sum_{y=1}^{x} \frac{y}{x} \frac{h(y) h^{(n-1)*}(x-y)}{h^{n*}(x)} h^{n*}(x) \\ &= \frac{\lambda}{x} \sum_{y=1}^{x} y h(y) \sum_{n=1}^{\infty} p(n-1) h^{(n-1)*}(x-y) = \frac{\lambda}{x} \sum_{y=1}^{x} y h(y) f(x-y), \end{aligned}$$

which proves (2.7).

This completes the proof of Theorem 2.2. □

2.2.2 *Application of Generating Functions*

The proofs we have given for Theorems 2.1 and 2.2, introduce a technique we shall apply to deduce many recursions in this book. However, the results can often also be proved by using generating functions. Some authors do that with great elegance. However, in the opinion of the present authors, when working on the distributions themselves instead of through generating functions, you get a more direct feeling of what is going on. Using generating functions seems more like going from one place to another by an underground train; you get where you want, but you do not have any feeling of how the landscape gradually changes on the way.

To illustrate how generating functions can be used as an alternative to the technique that we shall normally apply, we shall now first give an alternative proof of Theorem 2.2 based on such functions. After that, we shall deduce an alternative recursion for f based on the form of τ_h.

Alternative Proof of Theorem 2.2 We have

$$\tau_p(s) = \sum_{n=0}^{\infty} s^n p(n) = \sum_{n=0}^{\infty} s^n \frac{\lambda^n}{n!} e^{-\lambda} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{(s\lambda)^n}{n!},$$

that is,

$$\tau_p(s) = e^{\lambda(s-1)}. \tag{2.12}$$

By application of (1.30), we obtain

$$\tau_f(s) = \tau_p(\tau_h(s)) = e^{\lambda(\tau_h(s)-1)}. \tag{2.13}$$

Differentiation with respect to s gives

$$\tau_f'(s) = \lambda \tau_h'(s) \tau_f(s), \tag{2.14}$$

that is,

$$\sum_{x=1}^{\infty} x s^{x-1} f(x) = \lambda \sum_{y=1}^{\infty} y s^{y-1} h(y) \sum_{x=0}^{\infty} s^x f(x), \tag{2.15}$$

from which we obtain

$$\sum_{x=1}^{\infty} s^x x f(x) = \lambda \sum_{y=1}^{\infty} \sum_{x=0}^{\infty} y s^{x+y} h(y) f(x) = \lambda \sum_{y=1}^{\infty} \sum_{x=y}^{\infty} y s^x h(y) f(x-y)$$
$$= \sum_{x=1}^{\infty} s^x \lambda \sum_{y=1}^{x} y h(y) f(x-y).$$

Comparison of coefficients gives (2.7). □

This proof still holds when $h \in \mathcal{P}_{10}$. From (2.12) and (1.33), we then get the initial value $f(0) = e^{-\lambda(1-h(0))}$.

With some experience, one would see (2.7) immediately from (2.15).

Let us now assume that $h \in \mathcal{P}_{10}$ satisfies the relation

$$\tau_h'(s) = \frac{\sum_{y=1}^{r} \eta(y) s^{y-1}}{1 - \sum_{y=1}^{r} \chi(y) s^y} \tag{2.16}$$

with r being a positive integer or infinity. Then

$$\begin{aligned}
\sum_{y=1}^{\infty} y h(y) s^{y-1} = \tau_h'(s) &= \sum_{y=1}^{r} \eta(y) s^{y-1} + \sum_{y=1}^{r} \chi(y) s^y \tau_h'(s) \\
&= \sum_{y=1}^{r} \eta(y) s^{y-1} + \sum_{z=1}^{r} \chi(z) s^z \sum_{u=1}^{\infty} u s^{u-1} h(u) \\
&= \sum_{y=1}^{\infty} \left(\eta(y) + \sum_{z=1}^{r} (y-z) \chi(z) h(y-z) \right) s^{y-1}.
\end{aligned}$$

Comparison of coefficients gives

$$h(y) = \frac{\eta(y)}{y} + \sum_{z=1}^{r} \left(1 - \frac{z}{y}\right) \chi(z) h(y-z). \qquad (y = 1, 2, \dots) \tag{2.17}$$

Theorem 2.3 *If p is the Poisson distribution* $\mathrm{Po}(\lambda)$ *and $h \in \mathcal{P}_{10}$ satisfies the recursion* (2.17) *for functions η and χ on $\{1, 2, \dots, r\}$ with r being a positive integer or infinity, then $f = p \vee h$ satisfies the recursion*

$$f(x) = \sum_{y=1}^{r} \left(\frac{\lambda}{x} \eta(y) + \left(1 - \frac{y}{x}\right) \chi(y) \right) f(x-y). \quad (x = 1, 2, \dots) \tag{2.18}$$

Proof Insertion of (2.17) in (2.7) gives that for $x = 1, 2, \dots$,

$$\begin{aligned}
f(x) &= \frac{\lambda}{x} \sum_{y=1}^{x} \left(\eta(y) + \sum_{z=1}^{r} (y-z) \chi(z) h(y-z) \right) f(x-y) \\
&= \frac{\lambda}{x} \left(\sum_{y=1}^{r} \eta(y) f(x-y) + \sum_{z=1}^{r} \chi(z) \sum_{y=z+1}^{x} (y-z) h(y-z) f(x-y) \right),
\end{aligned}$$

and by application of (2.7), we get (2.18). □

We see that the conditions of the theorem are always satisfied with $r = \infty$, $\eta = \Phi h$, and $\chi \equiv 0$. In that case, (2.18) reduces to (2.7).

Let us now look at three examples where Theorem 2.3 gives some simplification.

Example 2.1 Let h be the *logarithmic distribution* $\mathrm{Log}(\pi)$ given by

$$h(y) = \frac{1}{-\ln(1-\pi)} \frac{\pi^y}{y}. \qquad (y = 1, 2, \dots; 0 < \pi < 1) \tag{2.19}$$

Then

$$h(y) = \frac{\pi}{-\ln(1-\pi)} I(y = 1) + \left(1 - \frac{1}{y}\right)\pi h(y-1), \qquad (y = 1, 2, \dots)$$

that is, h satisfies the conditions of Theorem 2.3 with

$$r = 1; \qquad \eta(1) = \frac{\pi}{-\ln(1-\pi)}; \qquad \chi(1) = \pi.$$

Insertion in (2.18) gives that for $x = 1, 2, \dots$, we have

$$f(x) = \frac{\pi}{x}\left(\frac{\lambda}{-\ln(1-\pi)} + x - 1\right) f(x-1) = \frac{\alpha + x - 1}{x} \pi f(x-1)$$

with

$$\alpha = \frac{\lambda}{-\ln(1-\pi)}, \tag{2.20}$$

that is,

$$f(x) = \frac{(\alpha + x - 1)^{(x)}}{x!} \pi^x f(0) = \binom{\alpha + x - 1}{x} \pi^x f(0).$$

From (2.8) and (2.20), we obtain $f(0) = \mathrm{e}^{-\lambda} = (1-\pi)^{\alpha}$. Hence,

$$f(x) = \binom{\alpha + x - 1}{x} \pi^x (1-\pi)^{\alpha}.$$

This is the *negative binomial distribution* $\mathrm{NB}(\alpha, \pi)$. Hence, we have shown that a compound Poisson distribution with logarithmic severity distribution can be expressed as a negative binomial distribution. In Example 4.1, we shall show this in another way. □

Example 2.2 Let h be the shifted geometric distribution given by

$$h(y) = (1-\pi)\pi^{y-1}. \qquad (y = 1, 2, \dots; 0 < \pi < 1) \tag{2.21}$$

In this case, the compound distribution f is called a *Pólya–Aeppli distribution*. We have

$$\tau_h(s) = \sum_{y=1}^{\infty} (1-\pi)\pi^{y-1} s^y = \frac{(1-\pi)s}{1-\pi s}, \tag{2.22}$$

from which we obtain

$$\tau_h'(s) = \frac{1-\pi}{1-2\pi s+\pi^2 s^2},$$

that is, h satisfies the conditions of Theorem 2.3 with

$$r = 2; \quad \eta(1) = 1-\pi; \quad \eta(2) = 0; \quad \chi(1) = 2\pi; \quad \chi(2) = -\pi^2. \tag{2.23}$$

Insertion in (2.18) gives

$$f(x) = \frac{1}{x}((\lambda(1-\pi)+2(x-1)\pi)f(x-1)-(x-2)\pi^2 f(x-2)).$$
$$(x = 1, 2\ldots)$$

□

Example 2.3 Let h be the *uniform distribution* on the integers $0, 1, 2, \ldots, k$, that is,

$$h(y) = \frac{1}{k+1}. \qquad (y = 0, 1, 2, \ldots, k) \tag{2.24}$$

Then

$$\tau_h(s) = \sum_{y=0}^{k} \frac{1}{k+1} s^y = \frac{1}{k+1} \frac{1-s^{k+1}}{1-s}, \tag{2.25}$$

from which we obtain

$$\tau_h'(s) = \frac{\frac{1}{k+1} - s^k + \frac{k}{k+1} s^{k+1}}{1-2s+s^2}, \tag{2.26}$$

that is, h satisfies the conditions of Theorem 2.3 with

$$r = k+2$$
$$\eta(1) = \frac{1}{k+1}; \qquad \eta(k+1) = -1; \qquad \eta(k+2) = \frac{k}{k+1}$$
$$\chi(1) = 2; \qquad \chi(2) = -1$$

and $\eta(y)$ and $\chi(y)$ equal to zero for all other values of y. Insertion in (2.18) gives

$$f(x) = \frac{1}{x}\left(\left(\frac{\lambda}{k+1} + 2(x-1)\right)f(x-1) - (x-2)f(x-2)\right.$$
$$\left. - \lambda\left(f(x-k-1) - \frac{k}{k+1} f(x-k-2)\right)\right). \qquad (x = 1, 2, \ldots)$$

□

2.3 The Panjer Class

2.3.1 Panjer Recursions

Let us now compare the proof of Theorem 2.1 and the first proof of Theorem 2.2. In both cases, we utilised that the claim number distribution $p \in \mathcal{P}_{1\underline{0}}$ satisfied a recursion in the form

$$p(n) = v(n)p(n-1). \qquad (n = 1, 2, \dots) \tag{2.27}$$

In the Poisson case, we found a function t such that

$$\mathrm{E}[t(Y_1, x)|Y_{\bullet n} = x] = v(n) \qquad (x = 1, 2, \dots; n = 1, 2, \dots) \tag{2.28}$$

was independent of x; we had

$$t(y, x) = \lambda\frac{y}{x}; \qquad v(n) = \frac{\lambda}{n}.$$

In the geometric case, we actually did the same with $t(y, x) = v(n) = \pi$. In both cases, (2.28) was satisfied for any choice of $h \in \mathcal{P}_{11}$.

For any $p \in \mathcal{P}_{10}$, if (2.28) is satisfied, then by proceeding like in the first proof of Theorem 2.2, using that

$$\mathrm{E}[t(Y_1, x)|Y_{\bullet n} = x] = \sum_{y=1}^{x} t(y, x)\frac{h(y)h^{(n-1)*}(x-y)}{h^{n*}(x)} \tag{2.29}$$
$$(x = 1, 2, \dots; n = 1, 2, \dots)$$

like in (2.11), we obtain

$$\sum_{n=1}^{x} v(n)p(n-1)h^{n*}(x) = \sum_{y=1}^{x} t(y, x)h(y)f(x-y). \tag{2.30}$$
$$(x = 1, 2, \dots)$$

For $n = 1, 2, \dots$, we have $p(n) = q(n) + v(n)p(n-1)$ with

$$q(n) = p(n) - v(n)p(n-1), \tag{2.31}$$

so that

$$f(x) = (q \vee h)(x) + \sum_{n=1}^{x} v(n)p(n-1)h^{n*}(x). \qquad (x = 1, 2, \dots)$$

Insertion of (2.30) gives

$$f(x) = (q \vee h)(x) + \sum_{y=1}^{x} t(y,x)h(y)f(x-y)$$
$$= \sum_{n=1}^{x}(p(n) - v(n)p(n-1))h^{n*}(x)$$
$$+ \sum_{y=1}^{x} t(y,x)h(y)f(x-y). \qquad (x = 1, 2, \dots) \tag{2.32}$$

We have $t(x,x) = v(1)$ because in the conditional distribution of Y_1 given $\sum_{j=1}^{n} Y_j = x$ we have $Y_1 = x$ iff $n = 1$ as the Y_js are strictly positive. Insertion in (2.32) gives

$$f(x) = p(1)h(x) + \sum_{n=2}^{x}(p(n) - v(n)p(n-1))h^{n*}(x)$$
$$+ \sum_{y=1}^{x-1} t(y,x)h(y)f(x-y). \qquad (x = 1, 2, \dots) \tag{2.33}$$

If $p \in \mathcal{P}_{1\underline{l}}$ and $h \in \mathcal{P}_{1\underline{r}}$, then we have $f(x) = 0$ for all $x < lr$. Thus, $f \in \mathcal{P}_{1\underline{lr}}$, and the initial value of the recursion is

$$f(lr) = \begin{cases} p(l)h^{l*}(r) & (l = 1, 2, \dots) \\ p(0). & (l = 0) \end{cases}$$

If p satisfies (2.27), then $q \equiv 0$, so that (2.32) reduces to

$$f(x) = \sum_{y=1}^{x} t(y,x)h(y)f(x-y). \qquad (x = 1, 2, \dots) \tag{2.34}$$

If both (t_1, v_1) and (t_2, v_2) satisfy (2.28), then $(t, v) = (at_1 + bt_2, av_1 + bv_2)$ also satisfies (2.28) for any constants a and b. In particular, this gives that for all $h \in \mathcal{P}_{11}$, (2.28) is satisfied for all linear combinations of the ts of Theorems 2.1 and 2.2, that is, for

$$t(y,x) = a + b\frac{y}{x}; \qquad v(n) = a + \frac{b}{n}. \tag{2.35}$$

Insertion in (2.31)–(2.33) gives

$$q(n) = p(n) - \left(a + \frac{b}{n}\right)p(n-1) \qquad (n = 1, 2, \dots) \tag{2.36}$$

and

$$\begin{aligned} f(x) &= (q \vee h)(x) + \sum_{y=1}^{x}\left(a + b\frac{y}{x}\right)h(y)f(x-y) \\ &= \sum_{n=1}^{x}\left(p(n) - \left(a + \frac{b}{n}\right)p(n-1)\right)h^{n*}(x) \\ &\quad + \sum_{y=1}^{x}\left(a + b\frac{y}{x}\right)h(y)f(x-y) \\ &= p(1)h(x) + \sum_{n=2}^{x}\left(p(n) - \left(a + \frac{b}{n}\right)p(n-1)\right)h^{n*}(x) \\ &\quad + \sum_{y=1}^{x-1}\left(a + b\frac{y}{x}\right)h(y)f(x-y), \qquad (x = 1, 2, \dots) \end{aligned} \tag{2.37}$$

from which we immediately obtain the following theorem.

Theorem 2.4 *If* $p \in \mathcal{P}_{1\underline{0}}$ *satisfies the recursion*

$$p(n) = \left(a + \frac{b}{n}\right)p(n-1) \qquad (n = 1, 2, \dots) \tag{2.38}$$

for some constants a *and* b, *then*

$$f(x) = \sum_{y=1}^{x}\left(a + b\frac{y}{x}\right)h(y)f(x-y) \qquad (x = 1, 2, \dots) \tag{2.39}$$

for any $h \in \mathcal{P}_{11}$.

The following theorem is a continuous version of Theorem 2.4.

Theorem 2.5 *The compound distribution with continuous severity distribution on* $(0, \infty)$ *with density* h *and counting distribution* $p \in \mathcal{P}_{1\underline{0}}$ *that satisfies the recursion* (2.38), *has mass* $p(0)$ *at zero and for* $x > 0$ *density* f *that satisfies the integral equation*

$$f(x) = p(1)h(x) + \int_0^x \left(a + b\frac{y}{x}\right)h(y)f(x-y)\,dy. \tag{2.40}$$

Proof We immediately see that the compound distribution has mass $p(0)$ at zero.

For $x > 0$, we have

$$f(x) = \sum_{n=1}^{\infty} p(n)h^{n*}(x) = p(1)h(x) + \sum_{n=2}^{\infty}\left(a + \frac{b}{n}\right)p(n-1)h^{n*}(x)$$

$$= p(1)h(x) + \sum_{n=2}^{\infty} \int_0^x \left(a + b\frac{y}{x}\right) \frac{h(y)h^{(n-1)*}(x-y)}{h^{n*}(x)} \mathrm{d}y\, p(n-1)h^{n*}(x)$$

$$= p(1)h(x) + \int_0^x \left(a + b\frac{y}{x}\right) h(y) \sum_{n=2}^{\infty} p(n-1)h^{(n-1)*}(x-y)\,\mathrm{d}y$$

$$= p(1)h(x) + \int_0^x \left(a + b\frac{y}{x}\right) h(y) f(x-y)\,\mathrm{d}y.$$

This completes the proof of Theorem 2.5. □

The Volterra integral equation (2.40) can be solved by numerical methods. However, in practice it seems more natural to approximate the continuous severity distribution with a discrete one, perhaps using an approximation that gives an upper or lower bound for the exact distribution.

Analogously, other recursions presented in this book can be modified to integral equations when the severity distribution is continuous.

2.3.2 Subclasses

The class of counting distributions satisfying the recursion (2.38) is often called the *Panjer class*. We already know that this class contains the geometric distribution and the Poisson distribution. The following theorem gives a complete characterisation of the Panjer class.

Theorem 2.6 *If* $p \in \mathcal{P}_{1\underline{0}}$ *satisfies the recursion* (2.38), *then we must have one of the following four cases*:

1. *Degenerate distribution in zero*:

$$p(n) = I(n = 0). \tag{2.41}$$

2. *Poisson distribution* $\mathrm{Po}(\lambda)$.
3. *Negative binomial distribution* $\mathrm{NB}(\alpha, \pi)$:

$$p(n) = \binom{\alpha + n - 1}{n} \pi^n (1-\pi)^{\alpha}. \qquad (n = 0, 1, 2, \ldots; 0 < \pi < 1; \alpha > 0) \tag{2.42}$$

4. *Binomial distribution* $\mathrm{bin}(M, \pi)$:

$$p(n) = \binom{M}{n} \pi^n (1-\pi)^{M-n}. \quad (n = 0, 1, 2, \ldots, M; 0 < \pi < 1; M = 1, 2, \ldots) \tag{2.43}$$

Proof To avoid negative probabilities, we must have $a + b \geq 0$.

If $a + b = 0$, we obtain $p(n) = 0$ for all $n > 0$, so that we get the degenerate distribution given by (2.41).

For the rest of the proof, we assume that $a + b > 0$.

From (2.9), we see that if $a = 0$, then p satisfies (2.6) with $\lambda = b$.

We now assume that $a > 0$. Then Theorem 1.11 gives that $a < 1$. With $\alpha = (a + b)/a$ and $\pi = a$, we obtain

$$p(n) = p(0)\prod_{i=1}^{n}\left(a + \frac{b}{i}\right) = p(0)\pi^n \prod_{i=1}^{n}\left(1 + \frac{\alpha - 1}{i}\right) = p(0)\pi^n \prod_{i=1}^{n}\frac{\alpha + i - 1}{i}$$

$$= p(0)\pi^n \frac{(\alpha + n - 1)^{(n)}}{n!} = p(0)\pi^n \binom{\alpha + n - 1}{n}.$$

Comparison with (2.42) gives that p must now be the negative binomial distribution $\mathrm{NB}(\alpha, \pi)$.

Let us finally consider the case $a < 0$. To avoid negative probabilities, there must then exist an integer M such that

$$a + \frac{b}{M + 1} = 0,$$

that is,

$$M = \frac{a + b}{-a}; \qquad b = -a(M + 1).$$

In that case, we have $p(n) = 0$ for all $n > M$. For $n = 0, 1, 2, \ldots, M$, we obtain

$$p(n) = p(0)\prod_{i=1}^{n}\left(a + \frac{b}{i}\right) = p(0)a^n \prod_{i=1}^{n}\left(1 - \frac{M + 1}{i}\right)$$

$$= p(0)(-a)^n \prod_{i=1}^{n}\frac{M - i + 1}{i} = p(0)(-a)^n \frac{M^{(n)}}{n!} = p(0)(-a)^n \binom{M}{n},$$

which gives (2.43) when $-a = \pi/(1 - \pi)$, that is, $\pi = -a/(1 - a)$.

This completes the proof of Theorem 2.6. □

In Fig. 2.1, the four cases of Theorem 2.6 are illustrated in an (a, b) diagram. Table 2.1 presents the recursion of Theorem 2.4 for the three non-degenerate cases of Theorem 2.6.

As a special case of the negative binomial distribution, we obtain the geometric distribution of Theorem 2.1 with $\alpha = 1$.

It is well known that both the binomial class and the negative binomial class satisfy the property that the convolution of two distribution within the class with the same value of the parameter π is the distribution in the same class with the same value of π and the other parameter being the sum of that parameter from the two

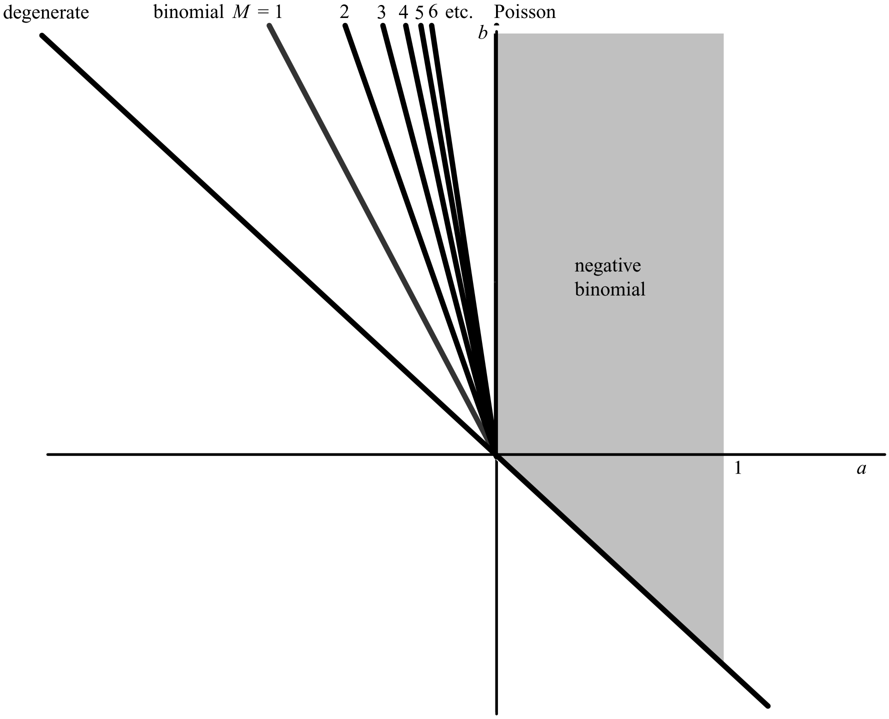

Fig. 2.1 (a, b) for the Panjer class

Table 2.1 Recursions for compound Panjer distributions

Distribution	a	b	$f(x)$	$f(0)$
$\mathrm{NB}(\alpha, \pi)$	π	$(\alpha - 1)\pi$	$\pi \sum_{y=1}^{x} (1 + (\alpha - 1)\frac{y}{x}) h(y) f(x - y)$	$(1 - \pi)^{\alpha}$
$\mathrm{Po}(\lambda)$	0	λ	$\frac{\lambda}{x} \sum_{y=1}^{x} y h(y) f(x - y)$	$\mathrm{e}^{-\lambda}$
$\mathrm{bin}(M, \pi)$	$-\frac{\pi}{1-\pi}$	$\frac{(M+1)\pi}{1-\pi}$	$\frac{\pi}{1-\pi} \sum_{y=1}^{x} ((M + 1)\frac{y}{x} - 1) h(y) f(x - y)$	$(1 - \pi)^{M}$

original distributions. When looking at the expressions for a and b for these two classes, we see that the two original distributions and their convolutions are in the Panjer class and have the same value of a, and the b of their convolution is a plus the sum of the bs of the two original distributions. As the convolution of two Poisson distributions is a Poisson distribution with parameter equal to the sum of the parameters of the two original distributions, this property also holds for the Poisson distributions, and, hence, for the whole Panjer class. We formulate this result as a theorem.

Theorem 2.7 *The convolution of two distributions that satisfy the recursion* (2.38) *with the same value of a, satisfies* (2.38) *with the same value of a and b equal to a plus the sum of the bs of the two original distributions.*

In Sect. 5.3.4, we shall prove a more general version of this theorem.

Let us now consider the moments of a distribution $p \in \mathcal{P}_{1\underline{0}}$ that satisfies the recursion (2.38). For $j = 1, 2, \ldots$, we have

$$\begin{aligned}\mu_p(j) &= \sum_{n=1}^{\infty} n^j p(n) = \sum_{n=1}^{\infty} n^j \left(a + \frac{b}{n}\right) p(n-1) = \sum_{n=1}^{\infty} n^{j-1}(an+b)p(n-1) \\ &= \sum_{n=0}^{\infty} (n+1)^{j-1}(a(n+1)+b)p(n),\end{aligned}$$

that is,

$$\mu_p(j) = a\mu_p(j;-1) + b\mu_p(j-1;-1). \tag{2.44}$$

By using that

$$\mu_p(k;-1) = \sum_{i=0}^{k} \binom{k}{i} \mu_p(i) \qquad (k = 0, 1, \ldots)$$

and solving (2.44) for $\mu_p(j)$, we obtain a recursion for $\mu_p(j)$; we shall return to that in Sect. 9.2.2. In particular, we get

$$\mu_p(1) = a\mu_p(1;-1) + b\mu_p(0;-1) = a\mu_p(1) + a + b,$$

which gives

$$\mu_p(1) = \frac{a+b}{1-a}.$$

Furthermore,

$$\begin{aligned}\mu_p(2) &= a\mu_p(2;-1) + b\mu_p(1;-1) \\ &= a(\mu_p(2) + 2\mu_p(1) + 1) + b(\mu_p(1) + 1) \\ &= a\mu_p(2) + (a+b)\mu_p(1) + a\mu_p(1) + a + b,\end{aligned}$$

from which we obtain

$$\begin{aligned}\mu_p(2) &= \frac{1}{1-a}((a+b)\mu_p(1) + a\mu_p(1) + a + b) \\ &= \mu_p(1)^2 + \frac{1}{1-a}\left(a\frac{a+b}{1-a} + a + b\right) = \mu_p(1)^2 + \frac{a+b}{(1-a)^2},\end{aligned}$$

which gives

$$\kappa_p(2) = \frac{a+b}{(1-a)^2}.$$

Table 2.2 Moments of distributions in the Panjer class

p	$\mu_p(1)$	$\kappa_p(2)$	$\kappa_p(2)/\mu_p(1)$
$\mathrm{NB}(\alpha, \pi)$	$\alpha\pi/(1-\pi)$	$\alpha\pi/(1-\pi)^2$	$1/(1-\pi)$
$\mathrm{Po}(\lambda)$	λ	λ	1
$\mathrm{bin}(M, \pi)$	$M\pi$	$M\pi(1-\pi)$	$1-\pi$

Hence,

$$\frac{\kappa_p(2)}{\mu_p(1)} = \frac{1}{1-a}.$$

From this we see that the variance is greater than the mean when $a > 0$, that is, negative binomial distribution, equal to the mean when $a = 0$, that is, Poisson distribution, and less than the mean when $a < 0$, that is, binomial distribution. This makes the Panjer class flexible for fitting counting distributions by matching of moments. In Table 2.2, we display the mean and the variance and their ratio for the three non-degenerate cases of Theorem 2.6.

2.3.3 An Alternative Recursion

For evaluation of $f = p \vee h$ with $h \in \mathcal{P}_{11}$ and $p \in \mathcal{P}_{1\underline{0}}$ satisfying the recursion (2.38), we shall deduce an alternative recursive procedure that can sometimes be more efficient than Theorem 2.4. We assume that h satisfies the relation

$$\tau_h(s) = \frac{\sum_{y=1}^{m} \alpha(y) s^y}{1 - \sum_{y=1}^{m} \beta(y) s^y} \tag{2.45}$$

with m being a positive integer or infinity. Rewriting this as $\tau_h = \tau_\alpha + \tau_\beta \tau_h$ and using (1.20), we obtain that

$$h = \alpha + \beta * h, \tag{2.46}$$

which gives the recursion

$$h(y) = \alpha(y) + \sum_{z=1}^{m} \beta(z) h(y-z). \qquad (y = 1, 2, \dots)$$

We shall need the following lemma.

Lemma 2.1 *If $w \in \mathcal{F}_{10}$ and $h \in \mathcal{P}_{11}$ satisfies the relation (2.45) with m being a positive integer or infinity, then $h * w$ satisfies the recursion*

$$(h * w)(x) = \sum_{y=1}^{m} (\alpha(y) w(x-y) + \beta(y)(h * w)(x-y)). \quad (x = 1, 2, \dots) \tag{2.47}$$

Proof Application of (2.46) gives $h * w = \alpha * w + \beta * h * w$, from which (2.47) follows. □

We see that (2.45) is always satisfied with $m = \infty$, $\alpha = h$, and $\beta \equiv 0$. In that case, (2.47) gives

$$(h * w)(x) = \sum_{y=1}^{x} h(y)w(x-y), \qquad (x = 1, 2, \dots)$$

which we already know.

We can express (2.39) in the form

$$f(x) = (a+b)(h * f)(x) - \frac{b}{x}(h * \Phi f)(x). \qquad (x = 1, 2, \dots) \tag{2.48}$$

For $x = 1, 2, \dots$, we can first evaluate $(h * f)(x)$ and $(h * \Phi f)(x)$ by (2.47) and then $f(x)$ by insertion in (2.48).

Example 2.4 Let p be the geometric distribution geo(π). Then $a = \pi$ and $b = 0$ so that (2.48) reduces to $f = \pi(h * f)$. Application of (2.47) gives that for $x = 1, 2, \dots,$

$$\begin{aligned} f(x) &= \pi \sum_{y=1}^{m} (\alpha(y) f(x-y) + \beta(y)(h * f)(x-y)) \\ &= \sum_{y=1}^{m} (\pi\alpha(y) + \beta(y)) f(x-y). \end{aligned}$$

This recursion can be considered as a parallel to the recursion (2.18). □

Example 2.5 Let h be the shifted geometric distribution given by (2.21). From (2.22), we see that (2.45) is satisfied with

$$m = 1; \qquad \alpha(1) = 1 - \pi; \qquad \beta(1) = \pi.$$

Insertion in (2.47) gives

$$(h * w)(x) = (1-\pi)w(x-1) + \pi(h * w)(x-1). \qquad (x = 1, 2, \dots)$$ □

By differentiating (2.45), we obtain

$$\tau_h'(s) = \frac{\sum_{y=1}^{m} y\alpha(y)s^{y-1}(1 - \sum_{z=1}^{m} \beta(z)s^z) + \sum_{y=1}^{m} \alpha(y)s^y \sum_{z=1}^{m} z\beta(z)s^{z-1}}{(1 - \sum_{z=1}^{m} \beta(z)s^z)^2},$$

which can be written in the form (2.16) with $r = 2m$. Hence, when p is the Poisson distribution Po(λ), we can also evaluate f by the recursion (2.18). In this case, $a = 0$

and $b = \lambda$ so that we can write (2.48) as

$$f(x) = \lambda\left((h * f)(x) - \frac{(h * \Phi f)(x)}{x}\right). \qquad (x = 1, 2, \ldots) \tag{2.49}$$

As $r = 2m$, the number of terms in the summation in (2.18) is twice the number of terms in the summation in (2.47). On the other hand, for each value of x in (2.49), we have to apply (2.47) twice, whereas in the recursion of Theorem 2.3, it suffices with one application of (2.18). As it seems to be an advantage to have the recursion expressed in one formula, we tend to go for the recursion of Theorem 2.3 in this case.

2.4 Convolutions of a Distribution

Let us now for a moment leave compound distributions and instead let $f = g^{M*}$ with $g \in \mathcal{P}_{1\underline{0}}$, that is, f is the distribution of $X = Y_{\bullet M}$, where $Y_1, Y_2, \ldots, Y_M$ are independent and identically distributed with distribution g. Then we have the following result.

Theorem 2.8 *The M-fold convolution $f = g^{M*}$ of $g \in \mathcal{P}_{1\underline{0}}$ satisfies the recursion*

$$f(x) = \frac{1}{g(0)} \sum_{y=1}^{x} \left((M+1)\frac{y}{x} - 1\right) g(y) f(x-y) \quad (x = 1, 2, \ldots) \tag{2.50}$$

$$f(0) = g(0)^M. \tag{2.51}$$

Proof Formula (2.51) follows immediately from (1.4).

Let us now prove (2.50). We introduce an auxiliary random variable Y_0, which is independent of X and has distribution g. Then, because of symmetry, we easily see that for $x = 1, 2, \ldots$

$$\mathrm{E}\left((M+1)\frac{Y_0}{x} - 1\right) I(Y_0 + X = x) = 0, \tag{2.52}$$

that is,

$$\sum_{y=0}^{x} \left((M+1)\frac{y}{x} - 1\right) g(y) f(x-y) = 0.$$

Solving for $f(x)$ gives (2.50).

This completes the proof of Theorem 2.8. □

Example 2.6 If g is the discrete uniform distribution given by (2.24), then the recursion (2.50) reduces to

$$f(x) = \sum_{y=1}^{k} \left((M+1)\frac{y}{x} - 1 \right) f(x-y), \qquad (x = 1, 2, \dots)$$

and (2.51) gives the initial condition $f(0) = (k+1)^{-M}$. In Example 5.2, we shall deduce a simpler recursion for f in the present situation. □

The simplest special case of a non-degenerate distribution g in Theorem 2.8 is the Bernoulli distribution $\mathrm{Bern}(\pi)$ given by

$$g(1) = \pi = 1 - g(0), \tag{2.53}$$

that is, the binomial distribution $\mathrm{bin}(1, \pi)$. By Theorem 2.7 and Table 2.1, we obtain that then f is $\mathrm{bin}(M, \pi)$. Insertion of (2.53) in (2.50) gives

$$f(x) = \frac{\pi}{1-\pi} \left(\frac{M+1}{x} - 1 \right) f(x-1), \qquad (x = 1, 2, \dots) \tag{2.54}$$

which is (2.38) with a and b given by Table 2.1 for the binomial distribution.

More generally, for any $g \in \mathcal{P}_{1\underline{0}}$, it follows from (1.8) that $g = q \vee h$ with q being $\mathrm{Bern}(\pi)$ with

$$\pi = 1 - g(0) \tag{2.55}$$

and $h \in \mathcal{P}_{11}$ given by

$$h(y) = \frac{g(y)}{\pi}. \qquad (y = 1, 2, \dots) \tag{2.56}$$

Then

$$f = g^{M*} = (q \vee h)^{M*} = q^{M*} \vee h = p \vee h$$

with $p = q^{M*}$, that is $\mathrm{bin}(M, \pi)$. Insertion of (2.55) and (2.56) in the recursion for the compound binomial distribution in Table 2.1 gives Theorem 2.8.

If

$$k = \max(x : g(x) > 0) < \infty, \tag{2.57}$$

then $f(x) = 0$ for all integers $x > Mk$, and we can turn the recursion (2.50) around and start it from $f(Mk)$. This can be convenient if we are primarily interested in $f(x)$ for high values of x. Furthermore, in this converted recursion, we can also allow for negative integers in the range of g.

Theorem 2.9 *If the distribution $g \in \mathcal{P}_1$ satisfies the condition* (2.57), *then $f = g^{M*}$ satisfies the recursion*

$$f(x) = \frac{1}{g(k)} \sum_{y=1}^{Mk-x} \left(\frac{(M+1)y}{Mk-x} - 1 \right) g(k-y) f(x+y)$$

$$(x = Mk-1, Mk-2, \ldots, 0)$$

$$f(Mk) = g(k)^M.$$

Proof Let $\tilde{Y}_j = k - Y_j$ $(j = 1, 2, \ldots, M)$ and

$$\tilde{X} = \sum_{j=1}^{M} \tilde{Y}_j = \sum_{j=1}^{M} (k - Y_j) = Mk - X,$$

and denote the distributions of $\tilde{Y}_j$ and $\tilde{X}$ by $\tilde{g}$ and $\tilde{f}$ respectively. Then $\tilde{g}, \tilde{f} \in \mathcal{P}_{1\underline{0}}$. Thus, they satisfy the recursion of Theorem 2.8, and we obtain

$$f(Mk) = \tilde{f}(0) = \tilde{g}(0)^M = g(k)^M.$$

For $x = Mk-1, Mk-2, \ldots, 0$, (2.50) gives

$$f(x) = \tilde{f}(Mk-x) = \frac{1}{\tilde{g}(0)} \sum_{y=1}^{Mk-x} \left(\frac{(M+1)y}{Mk-x} - 1 \right) \tilde{g}(y) \tilde{f}(Mk-x-y)$$

$$= \frac{1}{g(k)} \sum_{y=1}^{Mk-x} \left(\frac{(M+1)y}{Mk-x} - 1 \right) g(k-y) f(x+y).$$

This completes the proof of Theorem 2.9. □

If $g \in \mathcal{P}_{1\underline{l}}$ for some non-zero integer l, then we can also obtain a recursion for $g = f^{M*}$ from Theorem 2.8 by shifting g and f to $\mathcal{P}_{1\underline{0}}$.

Theorem 2.10 *If $g \in \mathcal{P}_{1\underline{l}}$ for some integer l, then $f = g^{M*}$ satisfies the recursion*

$$f(x) = \frac{1}{g(l)} \sum_{y=1}^{x-Ml} \left(\frac{(M+1)y}{x-Ml} - 1 \right) g(l+y) f(x-y)$$

$$(x = Ml+1, Ml+2, \ldots)$$

$$f(Ml) = g(l)^M.$$

Proof Let $\tilde{Y}_j = Y_j - l$ $(j = 1, 2, \dots, M)$ and

$$\tilde{X} = \sum_{j=1}^{M} \tilde{Y}_j = \sum_{j=1}^{M} (Y_j - l) = X - Ml,$$

and denote the distributions of $\tilde{Y}_j$ and $\tilde{X}$ by $\tilde{g}$ and $\tilde{f}$ respectively. Then $\tilde{g}$ and $\tilde{f}$ satisfy the recursion of Theorem 2.8, and we obtain

$$f(Ml) = \tilde{f}(0) = \tilde{g}(0)^M = g(l)^M.$$

For $x = Ml + 1, Ml + 2, \dots,$

$$f(x) = \tilde{f}(x - Ml) = \frac{1}{\tilde{g}(0)} \sum_{y=1}^{x-Ml} \left(\frac{(M+1)y}{x - Ml} - 1 \right) \tilde{g}(y) \tilde{f}(x - Ml - y)$$

$$= \frac{1}{g(l)} \sum_{y=1}^{x-Ml} \left(\frac{(M+1)y}{x - Ml} - 1 \right) g(l + y) f(x - y).$$

This completes the proof of Theorem 2.10. □

2.5 The Sundt–Jewell Class

2.5.1 Characterisation

Let us now return to the situation of Sect. 2.3.1. There we showed that if there exist functions t and v such that (2.28) holds, then (2.30) holds. If, in addition, the counting distribution p satisfies the recursion (2.27), then the compound distribution $f = p \vee h$ satisfies the recursion (2.34). Furthermore, we showed that for all severity distributions $h \in \mathcal{P}_{11}$, (2.28) is satisfied for t and v given by (2.35). A natural question is then for what other couples (t, v) (2.28) is satisfied for all h. The following theorem gives the answer.

Theorem 2.11 *There exists a function t that satisfies the relation (2.28) for all $h \in \mathcal{P}_{11}$ iff there exist constants a and b such that*

$$v(n) = a + \frac{b}{n}. \qquad (n = 2, 3, \dots) \tag{2.58}$$

Proof If the function v satisfies (2.58), then (2.28) is satisfied for all $h \in \mathcal{P}_{11}$ with

$$t(y, x) = \begin{cases} a + b\frac{y}{x} & (x \neq y) \\ v(1). & (x = y) \end{cases} \tag{2.59}$$

The reason that it works with a different value when $x = y$, is that in the conditional distribution of Y_1 given $Y_{\bullet n} = x$, we have $Y_1 = x$ iff $n = 1$ as the Y_js are strictly positive.

Let us now assume that there exists a function t that satisfies (2.28) for all $h \in \mathcal{P}_{11}$. We want to prove that then v must satisfy (2.58). It is then sufficient to show that for a particular choice of h (2.58) must be satisfied. We let

$$h(1) = h(2) = \frac{1}{2}.$$

By using that h^{n*} is a shifted binomial distribution, we obtain

$$h^{n*}(y) = \binom{n}{y-n} 2^{-n} \qquad (y = n, n+1, n+2, \dots, 2n; n = 1, 2, \dots)$$

from (2.43). Letting

$$h_n(y|x) = \Pr(Y_1 = y | Y_{\bullet n} = x)$$

for $n = 1, 2, \dots; x = n, n+1, n+2, \dots, 2n$, and $y = 1, 2$, we obtain

$$h_n(1|x) = \frac{h(1)h^{(n-1)*}(x-1)}{h^{n*}(x)} = \frac{\frac{1}{2}\binom{n-1}{x-n}2^{-(n-1)}}{\binom{n}{x-n}2^{-n}} = 2 - \frac{x}{n}$$

$$h_n(2|x) = 1 - h_n(1|x) = \frac{x}{n} - 1.$$

Insertion in (2.28) gives

$$\begin{aligned} v(n) &= \mathrm{E}[t(Y_1, x) | Y_{\bullet n} = x] = t(1, x)h_n(1|x) + t(2, x)h_n(2|x) \\ &= t(1, x)\left(2 - \frac{x}{n}\right) + t(2, x)\left(\frac{x}{n} - 1\right). \end{aligned}$$

With $x = n$ and $x = 2n$ respectively, we obtain

$$v(n) = t(1, n) = t(2, 2n).$$

Letting $x = 2z$ be an even number, we obtain

$$\begin{aligned} v(n) &= t(1, 2z)\left(2 - \frac{2z}{n}\right) + t(2, 2z)\left(\frac{2z}{n} - 1\right) \\ &= v(2z)\left(2 - \frac{2z}{n}\right) + v(z)\left(\frac{2z}{n} - 1\right), \end{aligned}$$

that is,

$$v(n) = A(z) + \frac{B(z)}{n} \qquad (1 \le n \le 2z \le 2n)$$

with

$$A(z) = 2v(2z) - v(z); \qquad B(z) = 2z(v(z) - v(2z)).$$

In particular, for $z \geq 2$, we obtain

$$v(z+1) = A(z+1) + \frac{B(z+1)}{z+1} = A(z) + \frac{B(z)}{z+1}$$
$$v(z+2) = A(z+1) + \frac{B(z+1)}{z+2} = A(z) + \frac{B(z)}{z+2},$$

which gives

$$A(z+1) = A(z); \qquad B(z+1) = B(z),$$

that is, (2.58) must be satisfied for some a and b.

This completes the proof of Theorem 2.11. □

2.5.2 Recursions

From (2.37) we immediately obtain that if $p \in \mathcal{P}_{1\underline{0}}$ satisfies the recursion

$$p(n) = \left(a + \frac{b}{n}\right) p(n-1) \qquad (n = l+1, l+2, \dots) \tag{2.60}$$

for some positive integer l and $h \in \mathcal{P}_{11}$, then the compound distribution $f = p \vee h$ satisfies the recursion

$$\begin{aligned} f(x) &= \sum_{n=1}^{l} \left(p(n) - \left(a + \frac{b}{n}\right) p(n-1) \right) h^{n*}(x) \\ &\quad + \sum_{y=1}^{x} \left(a + b\frac{y}{x}\right) h(y) f(x-y) \\ &= p(1)h(x) + \sum_{n=2}^{l} \left(p(n) - \left(a + \frac{b}{n}\right) p(n-1) \right) h^{n*}(x) \\ &\quad + \sum_{y=1}^{x-1} \left(a + b\frac{y}{x}\right) h(y) f(x-y). \qquad (x = 1, 2, \dots) \end{aligned} \tag{2.61}$$

In particular, if

$$p(n) = \left(a + \frac{b}{n}\right) p(n-1), \qquad (n = 2, 3, \dots) \tag{2.62}$$

then

$$f(x) = p(1)h(x) + \sum_{y=1}^{x-1}\left(a + b\frac{y}{x}\right)h(y)f(x-y). \quad (x = 1, 2, \dots) \tag{2.63}$$

The class of counting distributions given by (2.62) is sometimes called the *Sundt–Jewell class*.

2.5.3 Subclasses

The Sundt–Jewell class obviously contains the Panjer class. In the proof of Theorem 2.6, we pointed out that to avoid negative probabilities, we had to have $a + b \geq 0$; this is also illustrated in Fig. 2.1. In the Sundt–Jewell class, the recursion (2.62) starts at $n = 2$, so that we need only $2a + b \geq 0$ when $a > 0$.

Let us now look at what sort of distributions we have in the Sundt–Jewell class:

1. *The Panjer class.*
2. *Degenerate distribution concentrated in one*. Here we have $2a + b = 0$. In this case, $f = h$.
3. *Logarithmic distribution* $\mathrm{Log}(\pi)$. Here we have

$$a = \pi; \qquad b = -\pi, \tag{2.64}$$

so that

$$f(x) = \pi\left(\frac{h(x)}{-\ln(1-\pi)} + \sum_{y=1}^{x-1}\left(1 - \frac{y}{x}\right)h(y)f(x-y)\right). \qquad (x = 1, 2, \dots) \tag{2.65}$$

4. *Extended truncated negative binomial distribution* $\mathrm{ETNB}(\alpha, \pi)$. Let

$$p(n) = \frac{1}{\sum_{j=1}^{\infty}\binom{\alpha+j-1}{j}\pi^j}\binom{\alpha+n-1}{n}\pi^n. \tag{2.66}$$

$$(n = 1, 2, \dots; 0 < \pi \leq 1; -1 < \alpha < 0)$$

Then

$$a = \pi; \qquad b = (\alpha - 1)\pi \tag{2.67}$$

so that

$$f(x) = \pi\left(\frac{\alpha h(x)}{\sum_{j=1}^{\infty}\binom{\alpha+j-1}{j}\pi^j} + \sum_{y=1}^{x-1}\left(1 - (1-\alpha)\frac{y}{x}\right)h(y)f(x-y)\right).$$

$$(x = 1, 2, \dots) \tag{2.68}$$

When $0 < \pi < 1$, then

$$\sum_{j=1}^{\infty} \binom{\alpha + j - 1}{j} \pi^j = (1 - \pi)^{-\alpha} - 1.$$

5. *Truncated Panjer distributions*. Let $\tilde{p}$ be a distribution in the Panjer class and define the distribution $p \in \mathcal{P}_{11}$ by

$$p(n) = \frac{\tilde{p}(n)}{1 - \tilde{p}(0)}. \qquad (n = 1, 2, \dots)$$

6. *Zero-modification of distributions in the* Sundt–Jewell *class*. If $\tilde{p}$ is in the Sundt–Jewell class and $0 \le \rho \le 1$, then the mixed distribution p given by

$$p(n) = \rho I(n = 0) + (1 - \rho)\tilde{p}(n) \qquad (n = 0, 1, 2, \dots)$$

is also in the Sundt–Jewell class with the same a and b. We can also have $\rho \notin [0, 1]$ as long as $p(n) \in [0, 1]$ for all n. If $f = p \vee h$ and $\tilde{f} = \tilde{p} \vee h$, then we also have

$$f(x) = \rho I(x = 0) + (1 - \rho)\tilde{f}(x). \qquad (x = 0, 1, 2, \dots)$$

The following theorem shows that there are no other members in the Sundt–Jewell class than those in these classes.

Theorem 2.12 *The six classes described above contain all distributions in the Sundt–Jewell class.*

Proof Because of the mixtures in the class 6, there is an infinite number of counting distributions p with the same a and b. On the other hand, for each distribution in the Sundt–Jewell class, the corresponding distribution with lower truncation at one also belongs to the class, and any distribution in the Sundt–Jewell class can be obtained as a mixture between one of these truncated distributions and a degenerate distribution concentrated in zero. Hence, it suffices to study the distributions in the intersection between the Sundt–Jewell class and $\mathcal{P}_{11}$, and for each admissible pair (a, b) there exist only one such $p \in \mathcal{P}_{11}$.

We know that the class of Panjer distributions is contained in the Sundt–Jewell class, and, hence, that also goes for the truncated Panjer distributions. In Theorem 2.6 and Fig. 2.1, we have characterised the classes of (a, b) for these distributions.

Now, what other admissible values of (a, b) have we got with the extension from the Panjer class to the Sundt–Jewell class? We must now have $2a + b \ge 0$, and we have already considered the distributions with $a + b > 0$ in Theorem 2.6 and Fig. 2.1. From Theorem 1.11, we see that we still cannot have $a > 1$, and that we can have $a = 1$ only when $b < -1$. Hence, it suffices to check all possibilities within the closed triangle bounded by the lines $2a + b = 0$, $a + b = 0$, and $a = 1$, apart from the point with $a = 1$ and $b = -1$. This set is illustrated in Fig. 2.2.

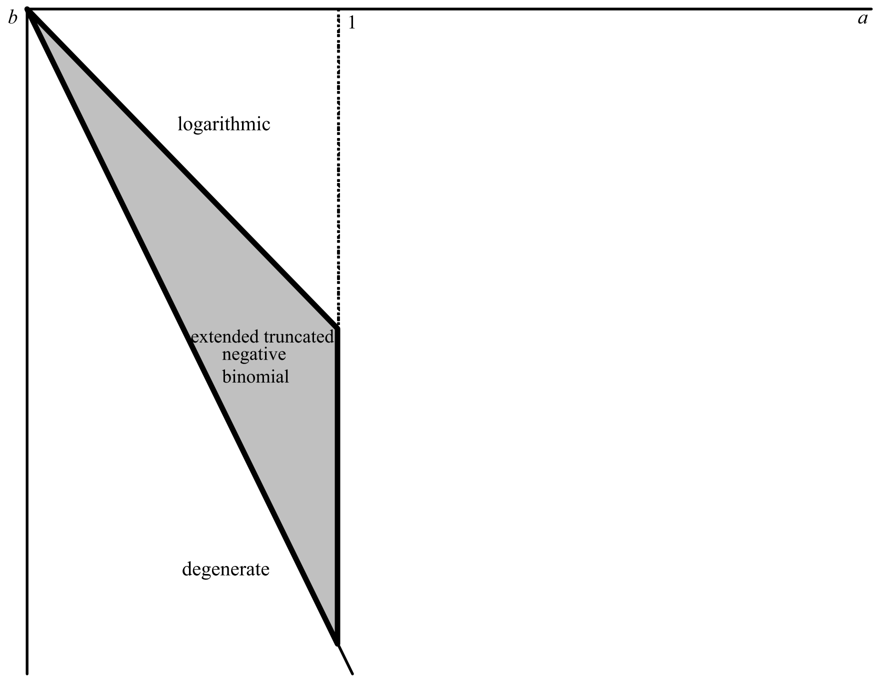

Fig. 2.2 (a, b) diagram for the complement of the Panjer class in the Sundt–Jewell class

When $2a+b=0$, we obtain the degenerate distribution concentrated in one, and, when $a+b=0$ with $a<1$, the logarithmic distribution $\mathrm{Log}(a)$.

For any (a, b) in the remaining area, we define π and α by (2.67), that is,

$$\pi = a; \qquad \alpha = \frac{a+b}{a}.$$

We then obviously have $0 < \pi \leq 1$. Furthermore, as $a+b<0$,

$$\alpha = \frac{a+b}{a} < 0,$$

and, as $2a+b>0$,

$$\alpha = \frac{2a+b-a}{a} > -1.$$

Hence, for each (a, b) in our remaining area, there exists an extended truncated negative binomial distribution $\mathrm{ETNB}(\alpha, \pi)$.

We have now allocated distributions in $\mathcal{P}_{11}$ from the classes 2–5 to all admissible pairs (a, b).

This completes the proof of Theorem 2.12. □

2.6 Higher Order Panjer Classes

2.6.1 Characterisation

For $l = 0, 1, 2, \dots$, let $\mathcal{S}_l$ denote the class of counting distributions $p \in \mathcal{P}_{1\underline{l}}$ that satisfy the recursion (2.60). We call this class the *Panjer class of order* l. Like we have done earlier, we often call the Panjer class of order zero simply the Panjer class.

The following theorem is proved analogous to Theorem 2.12.

Theorem 2.13 *If* $p \in \mathcal{S}_l$ *with* $l = 2, 3, \dots$, *then* p *belongs to one of the four classes*:

$$p(n) = \frac{\tilde{p}(n)}{1 - \tilde{p}(l-1)} \qquad (\tilde{p} \in \mathcal{S}_{l-1})$$

$$p(n) = I(n = l) \tag{2.69}$$

$$p(n) = \frac{1}{\sum_{j=l}^{\infty} \pi^j \binom{j}{l}^{-1}} \pi^n \binom{n}{l}^{-1} \qquad (0 < \pi \le 1) \tag{2.70}$$

$$p(n) = \frac{1}{\sum_{j=l}^{\infty} \binom{\alpha+j-1}{j} \pi^j} \binom{\alpha+n-1}{n} \pi^n \qquad (0 < \pi \le 1; -l < \alpha < -l+1) \tag{2.71}$$

for $n = l, l+1, l+2, \dots$.

With the distribution (2.70), we have $a = \pi$ and $b = -l\pi$, so that in the (a, b) plane, we cover the line $la + b = 0$ with $a \in (0, 1]$, and we obtain the distribution (2.69) when $(l+1)a + b = 0$. With the distribution (2.71), we have a and b given by (2.67), that is, in the (a, b) plane, we cover the triangle given by the restrictions $0 < a \le 1$, $la + b < 0$, and $a(l+1) + b \ge 0$.

If $p \in \mathcal{S}_l$ and $h \in \mathcal{P}_{11}$, then $f = p \vee h \in \mathcal{P}_{1l}$, and from (2.61), we obtain that f satisfies the recursion

$$f(x) = p(l)h^{l*}(x) + \sum_{y=1}^{x-l} \left(a + b\frac{y}{x}\right) h(y) f(x-y). \qquad (x = l, l+1, l+2, \dots) \tag{2.72}$$

We can evaluate h^{l*} recursively by Theorem 2.10.

2.6.2 Shifted Counting Distribution

Let us now assume that $p \in \mathcal{P}_{1\underline{0}}$ satisfies the recursion

$$p(n) = \left(a + \frac{b}{n+l}\right) p(n-1), \qquad (n = 1, 2, \dots) \tag{2.73}$$

and that we want to evaluate the compound distribution $f = p \vee h$ with $h \in \mathcal{P}_{11}$. Then the shifted distribution p_l given by $p_l(n) = p(n-l)$ for $n = l, l+1, l+2, \dots$ is the distribution in $\mathcal{S}_l$ given by (2.60). Thus, we can evaluate the compound distribution $f_l = p_l \vee h$ recursively by (2.72). Furthermore, we have $f_l = h^{l*} * f$, so that for $x = lr, lr+1, lr+2, \dots$

$$f_l(x) = \sum_{y=lr}^{x} h^{l*}(y) f(x-y)$$

if $h \in \mathcal{P}_{1\underline{r}}$. By solving this equation for $f(x-lr)$, using that $h^{l*}(lr) = h(r)^l$, we obtain

$$f(x-lr) = \frac{1}{h(r)^l}\left(f_l(x) - \sum_{y=lr+1}^{x} h^{l*}(y) f(x-y)\right).$$

Change of variable gives

$$f(x) = \frac{1}{h(r)^l}\left(f_l(x+lr) - \sum_{y=1}^{x} h^{l*}(y+lr) f(x-y)\right). \qquad (x = 0, 1, 2, \dots)$$

Let us look at shifting the opposite way. We want to evaluate $f = p \vee h$ with $h \in \mathcal{P}_{11}$ and $p \in \mathcal{P}_{1\underline{l}}$ satisfying the recursion

$$p(n) = \left(a + \frac{b}{n-l}\right) p(n-1). \qquad (n = l+1, l+2, \dots)$$

Then the shifted distribution p_{-l} given by $p_{-l}(n) = p(n+l)$ satisfies the recursion (2.38), and, thus, the compound distribution $f_{-l} = p_{-l} \vee h$ satisfies the recursion (2.39). As $f = h^{l*} * f_{-l}$, we can evaluate f by

$$f(x) = \sum_{y=l}^{x} h^{l*}(y) f_{-l}(x-y). \qquad (x = l, l+1, l+2, \dots)$$

2.6.3 Counting Distribution with Range Bounded from Above

Let us now consider a distribution p on a range of non-negative integers $\{l, l+1, l+2, \dots, r\}$ obtained from a distribution $\tilde{p} \in \mathcal{S}_l$ by

$$p(n) = \frac{\tilde{p}(n)}{\sum_{j=l}^{r} \tilde{p}(j)}. \qquad (n = l, l+1, \dots, r)$$

Then p satisfies a recursion in the form

$$p(n) = \left(a + \frac{b}{n}\right) p(n-1). \qquad (n = l+1, l+2, \dots, r) \qquad (2.74)$$

Application of (2.37) gives

$$f(x) = p(l)h^{l*}(x) - \left(a + \frac{b}{r+1}\right)p(r)h^{(r+1)*}(x) + \sum_{y=1}^{x}\left(a + b\frac{y}{x}\right)h(y)f(x-y).$$
$$(x = l, l+1, l+2, \dots) \qquad (2.75)$$

The recursions (2.74) and (2.75) are satisfied for more general pairs (a, b) than what follows from the construction from distributions in $\mathcal{S}_l$, as for $n > r$, $a + b/n$ does not need to be non-negative.

Example 2.7 Let $p \in \mathcal{P}_{\underline{1}0}$ be given by

$$p(n) = \frac{\binom{M}{n}(\frac{\pi}{1-\pi})^n}{\sum_{j=0}^{r}\binom{M}{j}(\frac{\pi}{1-\pi})^j}.$$
$$(n = 0, 1, 2 \dots, r; 0 < \pi < 1; r = 1, 2, \dots; M \geq r)$$

Then

$$a = -\frac{\pi}{1-\pi}; \qquad b = (M+1)\frac{\pi}{1-\pi},$$

and (2.75) gives

$$f(x) = \frac{\pi}{1-\pi}\left(\sum_{y=1}^{x}\left((M+1)\frac{y}{x} - 1\right)h(y)f(x-y) - \frac{M-r}{r+1}p(r)h^{(r+1)*}(x)\right).$$
$$(x = 1, 2, \dots) \qquad (2.76)$$

If $r = M$, then p is the binomial distribution $\text{bin}(M, \pi)$, and (2.76) reduces to the recursion for compound binomial distributions given in Table 2.1. □

2.7 Extension to Severity Distributions in $\mathcal{P}_{10}$

2.7.1 Recursions

Apart from Sect. 2.2.2, till now, we have always assumed that the severity distribution belongs to $\mathcal{P}_{11}$ when discussing recursions for compound distributions. We shall now relax this assumption by allowing the severities to be equal to zero, so let $h \in \mathcal{P}_{10}$. In (2.29) and (2.30), we must then sum from $y = 0$ instead of $y = 1$, so that (2.32) becomes

$$f(x) = (q \vee h)(x) + \sum_{y=0}^{x} t(y, x)h(y)f(x-y). \qquad (x = 1, 2, \dots)$$

As $f(x)$ appears with $y=0$ in summation, this does not yet give an explicit expression for $f(x)$, so we solve for $f(x)$ and obtain

$$\begin{aligned}
f(x) &= \frac{1}{1-t(0,x)h(0)}\left((q \vee h)(x) + \sum_{y=1}^{x} t(y,x)h(y)f(x-y)\right) \\
&= \frac{1}{1-t(0,x)h(0)}\left(\sum_{n=1}^{\infty}(p(n)-v(n)p(n-1))h^{n*}(x)\right. \\
&\quad \left. + \sum_{y=1}^{x} t(y,x)h(y)f(x-y)\right) \\
&= \frac{1}{1-t(0,x)h(0)}\left(p(1)h(x) + \sum_{n=2}^{\infty}(p(n)-v(n)p(n-1))h^{n*}(x)\right. \\
&\quad \left. + \sum_{y=1}^{x-1} t(y,x)h(y)f(x-y)\right). \qquad (x=1,2,\dots) \qquad (2.77)
\end{aligned}$$

From (1.33) we obtain that $f(0)=\tau_p(h(0))$. If h and/or p belong to $\mathcal{P}_{1\underline{0}}$, then we can use this as initial value for the recursion (2.77).

If $p \in \mathcal{P}_{10}$ satisfies (2.60), then (2.77) gives the recursion

$$\begin{aligned}
f(x) &= \frac{1}{1-ah(0)}\left(\sum_{n=1}^{l}\left(p(n)-\left(a+\frac{b}{n}\right)p(n-1)\right)h^{n*}(x)\right. \\
&\quad \left. + \sum_{y=1}^{x}\left(a+b\frac{y}{x}\right)h(y)f(x-y)\right) \\
&= \frac{1}{1-ah(0)}\left(p(1)h(x) + \sum_{n=2}^{l}\left(p(n)-\left(a+\frac{b}{n}\right)p(n-1)\right)h^{n*}(x)\right. \\
&\quad \left. + \sum_{y=1}^{x-1}\left(a+b\frac{y}{x}\right)h(y)f(x-y)\right). \qquad (x=1,2,\dots) \qquad (2.78)
\end{aligned}$$

In particular, if p is in the Panjer class, we obtain

$$f(x) = \frac{1}{1-ah(0)}\sum_{y=1}^{x}\left(a+b\frac{y}{x}\right)h(y)f(x-y). \qquad (x=1,2,\dots) \qquad (2.79)$$

Table 2.3 presents this recursion and its initial value $f(0)$ for the three subclasses of non-degenerate distributions in the Panjer class as given by Theorem 2.6. We have already encountered the Poisson case in Sect. 2.2.2.

Table 2.3 Recursions for compound Panjer distributions

Distribution	$f(x)$	$f(0)$
$\mathrm{NB}(\alpha,\pi)$	$\frac{\pi}{1-\pi h(0)}\sum_{y=1}^{x}(1+(\alpha-1)\frac{y}{x})h(y)f(x-y)$	$(\frac{1-\pi}{1-\pi h(0)})^{\alpha}$
$\mathrm{Po}(\lambda)$	$\frac{\lambda}{x}\sum_{y=1}^{x} yh(y)f(x-y)$	$e^{-\lambda(1-h(0))}$
$\mathrm{bin}(M,\pi)$	$\frac{\pi}{1-\pi+\pi h(0)}\sum_{y=1}^{x}((M+1)\frac{y}{x}-1)h(y)f(x-y)$	$(1-\pi(1-h(0)))^{M}$

2.7.2 Thinning

The recursions introduced in Sect. 2.7.1 can be used to study the effect of thinning. Let N be the number of observations and Y_j the size of the jth of these. We assume that the Y_js are mutually independent and identically distributed with distribution h and independent of N which has distribution p. We also introduce $X = Y_{\bullet N}$ and its distribution $f = p \vee h$. Let us assume that we are interested in the number of observations that satisfy a certain criterion. In insurance, this could e.g. be the number of claims that exceed some retention. In this context, we can let Y_j be an indicator variable equal to one if the observation satisfies the criterion, and zero otherwise. Thus, we let h be the Bernoulli distribution $\mathrm{Bern}(\pi)$ with $0 < \pi < 1$. Then, for $n = 1, 2, \ldots,$ h^{n*} is the binomial distribution $\mathrm{bin}(n,\pi)$, and insertion of (2.43) in (2.78) gives that for $x = 1, 2, \ldots,$

$$f(x) = \frac{1}{1-a(1-\pi)}\left(\sum_{n=x}^{l}\left(p(n) - \left(a + \frac{b}{n}\right)p(n-1)\right)\binom{n}{x}\pi^x(1-\pi)^{n-x} + \left(a + \frac{b}{x}\right)\pi f(x-1)\right)$$

$$= \frac{1}{1-a(1-\pi)}\sum_{n=x}^{l}\left(p(n) - \left(a + \frac{b}{n}\right)p(n-1)\right)\binom{n}{x}\pi^x(1-\pi)^{n-x} + \left(a_\pi + \frac{b_\pi}{x}\right)f(x-1)$$

with

$$a_\pi = \frac{a\pi}{1-a+a\pi}; \qquad b_\pi = \frac{b\pi}{1-a+a\pi}. \tag{2.80}$$

For $x > l$, the first term vanishes, so that

$$f(x) = \left(a_\pi + \frac{b_\pi}{x}\right)f(x-1). \qquad (x = l+1, l+2, \ldots)$$

It is interesting to note that (a_π, b_π) is on the line between $(0,0)$ and (a,b) in Fig. 2.1. As each of the classes of distributions given by (2.6), (2.42), (2.43), (2.70), and (2.71) satisfy the property that for any point in its area in the (a, b) diagram, all points on the line between that point and $(0, 0)$ also belong to the same class, we see that all these classes, in particular the Panjer class, are closed under thinning.

The thinned distribution f is called the π*-thinning* of p, that is, p is thinned with thinning probability π. Analogously, X is called a π-thinning of N.

2.7.3 Conversion to Severity Distributions in $\mathcal{P}_{11}$

Let $h \in \mathcal{P}_{1\underline{0}}$. When discussing the connection between the recursions for the M-fold convolutions and compound binomial distributions after the proof of Theorem 2.8, we showed how any distribution in $\mathcal{P}_{1\underline{0}}$ can be expressed as a compound Bernoulli distribution with severity distribution in $\mathcal{P}_{11}$. Let us now do this with a distribution $h \in \mathcal{P}_{1\underline{0}}$, denoting the counting distribution by q, its Bernoulli parameter by π, and the severity distribution by $\tilde{h}$, so that $h = q \vee \tilde{h}$. We want to evaluate the compound distribution $f = p \vee h$ with $p \in \mathcal{P}_{10}$ satisfying the recursion (2.60). Then

$$f = p \vee h = p \vee (q \vee \tilde{h}) = (p \vee q) \vee \tilde{h} = \tilde{p} \vee \tilde{h}$$

with $\tilde{p} = p \vee q$. Hence, we have now transformed a compound distribution with severity distribution in $\mathcal{P}_{1\underline{0}}$ to a compound distribution with severity distribution in $\mathcal{P}_{11}$. Furthermore, from the discussion above, we know that the counting distribution satisfies a recursion of the same type as the original counting distribution.

Further Remarks and References

With a different parameterisation, the Panjer class was studied by Katz (1945, 1965) and is sometimes referred to as the *Katz class*. In particular, Katz (1965) gave a characterisation of this class similar to Theorem 2.6 and visualised it in a diagram similar to Fig. 2.1. However, he seems to believe that when $a < 0$, we obtain a proper distribution even when b/a is not an integer; as we have indicated in the proof of Theorem 2.6, we then get negative probabilities.

Even earlier traces of the Panjer class are given by Carver (1919), Guldberg (1931), and Ottestad (1939); see Johnson et al. (2005, Sect. 2.3.1).

Luong and Garrido (1993) discussed parameter estimation within the Panjer class, and Katz (1965) and Fang (2003a, 2003b) discussed testing the hypothesis that a distribution within the Panjer class is Poisson.

In the actuarial literature, Theorem 2.4 is usually attributed to Panjer (1981). However, there are earlier references both within and outside the actuarial area. In the actuarial literature, the Poisson case was presented by Panjer (1980) and Williams (1980) and the Poisson and negative binomial cases by Stroh (1978). The Poisson, binomial, and negative binomial cases were deduced separately by Tilley in a discussion to Panjer (1980). Outside the actuarial literature, the Poisson case was treated by Neyman (1939), Beall and Rescia (1953), Katti and Gurland (1958), Shumway and Gurland (1960), Adelson (1966), Kemp (1967), and Plackett (1969); Khatri and Patel (1961) treat the Poisson, binomial, and negative binomial cases separately. Other proofs for Theorem 2.2 are given by Gerber (1982) and Hürlimann (1988).

Panjer (1981) also proved the continuous case given in Theorem 2.5. The negative binomial case was presented by Seal (1971). Ströter (1985) and den Iseger et al. (1997) discussed numerical solution of the integral equation (2.40). Another approach is to approximate the severity distribution by an arithmetic distribution; references for such approximations are given in Chap. 1.

Panjer (1981) was followed up by Sundt and Jewell (1981). They discussed various aspects of Panjer's framework. In particular, they proved Theorem 2.6 and visualised it in an (a, b) diagram like Fig. 2.1. They also introduced the framework with (2.27) and (2.28) and proved a slightly different version of Theorem 2.11. Furthermore, they presented the recursion (2.61) and its special case (2.72) as well as the recursion (2.75). They also extended the recursions to severity distributions in $\mathcal{P}_{1\underline{0}}$ like in Sect. 2.7.1.

Compound geometric distributions often appear in ruin theory and queuing theory. For applications of Theorem 2.1 in ruin theory, see e.g. Goovaerts and De Vylder (1984), Dickson (1995, 2005), Willmot (2000), and Cossette et al. (2004), and for applications in queuing theory, Hansen (2005) and Hansen and Pitts (2006). Reinhard and Snoussi (2004) applied Theorem 2.2 in ruin theory.

From Theorem 2.2, we obtain that if $f = p \vee h$ with $h \in \mathcal{P}_{11}$ and p being the Poisson distribution $\mathrm{Po}(\lambda)$, then

$$\lambda = -\ln f(0) \tag{2.81}$$

$$h(x) = \frac{1}{f(0)}\left(-\frac{xf(x)}{\ln f(0)} - \sum_{y=1}^{x-1} yh(y)f(x-y)\right). \quad (x = 1, 2, \dots) \tag{2.82}$$

Hence, λ and h are uniquely determined by f. Buchmann and Grübel (2003) proposed estimating λ and h by replacing f in (2.81) and (2.82) with the empirical distribution of a sample of independent observations from the distribution f. It should be emphasised that as a compound Poisson distribution with severity distribution in $\mathcal{P}_{11}$ always has infinite support, such an estimate of h based on the empirical distribution of a finite sample from f, can never be a distribution itself.

Such an estimation procedure can also be applied for other counting distributions p as long as p has only one unknown parameter. When the counting distribution has more parameters, we can still estimate h by replacing f in (2.82) with its empirical counterpart if we consider the parameters of p as given. Such estimation procedures were studied by Hansen and Pitts (2009).

Panjer and Willmot (1992) used generating functions extensively for deduction of recursions for aggregate claims distributions.

Theorem 2.3 was proved by De Pril (1986a), who also gave some examples. Chadjiconstantinidis and Pitselis (2008) present results based on that theorem.

In July 2006, Georgios Pitselis kindly gave us an early version of that paper and was positive to us to use material from it in our book. Since then, there has been some exchange of ideas between him and us, and this has influenced results both later in this book and in the paper.

The representation of a negative binomial distribution as a compound Poisson distribution with a logarithmic severity distribution was presented by Ammeter (1948, 1949) and Quenouille (1949).

The discussion on moments in Sect. 2.3.2 is based on Jewell (1984).

Lemma 2.1 and the related algorithm for recursive evaluation of compound distributions was presented by Hipp (2006) within the framework of phase distributions. He also discussed continuous and mixed severity distributions.

Willmot and Woo (2007) applied Panjer recursions in connection with evaluating discrete mixtures of Erlang distributions.

Reinsurance applications of Panjer recursions are discussed by Panjer and Willmot (1984), Sundt (1991a, 1991b), Mata (2000), Walhin (2001, 2002a), and Walhin et al. (2001).

McNeil et al. (2005) presented Theorem 2.4 within the framework of quantitative risk management.

Douligeris et al. (1997) applied Panjer recursions in connection with oil transportation systems.

For M-fold convolutions, De Pril (1985) deduced the recursions in Theorems 2.8 and 2.10. However, in pure mathematics, the recursion in Theorem 2.8 is well known for evaluation of the coefficients of powers of power series; Gould (1974) traces it back to Euler (1748). Sundt and Dickson (2000) compared the recursion of Theorem 2.8 with other methods for evaluation of M-fold convolutions of distributions in $\mathcal{P}_{1\underline{0}}$.

Willmot (1988) characterised the Sundt–Jewell class; see also Panjer and Willmot (1992, Sect. 7.2) and Johnson et al. (2005, Sect. 2.3.2). The higher order Panjer classes were characterised by Hess et al. (2002). Recursive evaluation of compound distributions with counting distribution satisfying (2.70) or (2.71) and severity distribution in $\mathcal{P}_{10}$ have been discussed by Gerhold et al. (2008). Sundt (2002) presented the procedure for recursive evaluation of a compound distribution with counting distribution given by (2.73).

Thinning in connection with the recursions has been discussed by Milidiu (1985), Willmot (1988), and Sundt (1991b). For more information, see also Willmot (2004) and Grandell (1991).

Panjer and Willmot (1982) and Hesselager (1994) discussed recursive evaluation of compound distributions with severity distribution in $\mathcal{P}_{10}$ and counting distribution $p \in \mathcal{P}_{1\underline{0}}$ that satisfies a recursion is the form

$$p(n) = \frac{\sum_{i=0}^{t} c(i)n^i}{\sum_{i=0}^{t} d(i)n^i} p(n-1). \qquad (n = 1, 2, \dots)$$

The Panjer class appears as a special case with $t = 1$ and $d(0) = 0$.

Ambagaspitiya (1995) discussed a class of distributions $p_{a,b} \in \mathcal{P}_{10}$ that satisfy a relation in the form

$$p_{a,b}(n) = \left(u(a,b) + \frac{v(a,b)}{n}\right) p_{a+b,b}(n-1). \qquad (n = l+1, l+2, \dots)$$

In particular, he discussed recursive evaluation of compound distributions with such a counting distribution and severity distribution in $\mathcal{P}_{11}$. The special case where $p_{a,b}$ satisfies the relation

$$p_{a,b}(n) = \frac{a}{a+b}\left(a + \frac{b}{n}\right) p_{a+b,b}(n-1), \qquad (n = 1, 2, \dots)$$

was treated by Ambagaspitiya and Balakrishnan (1994).

Hesselager (1997) deduced recursions for a compound Lagrange distribution and a compound shifted Lagrange distribution with kernel in the Panjer class. Recursions in connection with Lagrange distributions have also been studied by Sharif (1996) and Sharif and Panjer (1998). For more information on Lagrange distributions, see Johnson et al. (2005).

A special case of compound Lagrange distributions is the generalised Poisson distribution. Recursions in connection with this distribution have been studied by Goovaerts and Kaas (1991), Ambagaspitiya and Balakrishnan (1994), and Sharif and Panjer (1995).

By counting the number of dot operations (that is, multiplications and divisions), Bühlmann (1984) compared the recursive method of Theorem 2.2 with a method presented by Bertram (1981) (see also Feilmeier and Bertram 1987) based on the Fast Fourier Transform. Such comparison of methods presented in this book with each other or other methods have been performed by Kuon et al. (1987), Waldmann (1994), Dhaene and Vandebroek (1995), Sundt and Dickson (2000), Dickson and Sundt (2001), Dhaene et al. (2006), Sundt and Vernic (2006), and, in a bivariate setting, Walhin and Paris (2001c), some of them also counting bar operations (that is, additions and subtractions). Where both dot and bar operations are treated, these two classes are usually considered separately. The reason for distinguishing between these classes and sometimes dropping the bar operations, is that on computers, dot operations are usually more time-consuming than bar operations. Counting arithmetic operations is not a perfect criterion of comparing methods. There are also other aspects that should be taken into account. This is discussed by Sundt and Dickson (2000). It should be emphasised that when doing such counting, one should not just count the operations mechanically from the recursion formulae, but also consider how one could reduce the number of operations by introduction of auxiliary functions. For instance, in the recursion (2.7), one can save a lot of multiplications by first evaluating Φh instead of multiplying y by $h(y)$ at each occurrence, and if h has a finite range, we can further reduce the number of multiplications by instead evaluating $\lambda\Phi h$. How one sets up the calculations, can also affect the numerical accuracy of the evaluation. This aspect has been discussed by Waldmann (1995). We shall not pursue these issues further in this book.

Operators like $\vee$, Φ, and Ψ can be used to make our formulae more compact. However, it has sometimes been difficult to decide on how far to stretch this. As an example, let us look at the first part of (2.37), that is,

$$f(x) = (q \vee h)(x) + \sum_{y=1}^{x} \left(a + b\frac{y}{x}\right) h(y) f(x-y). \tag{2.83}$$

This formula can be made more compact as

$$f(x) = (q \vee h)(x) + a(h \vee f)(x) + b\Psi(\Phi h \vee f)(x). \tag{2.84}$$

On the other hand, it can be made less compact as

$$f(x) = \sum_{n=1}^{x} q(n) h^{n*}(x) + \sum_{y=1}^{x} \left(a + b\frac{y}{x}\right) h(y) f(x-y). \tag{2.85}$$

So why have we then used something between these two extremes? The reason that in (2.83) we have not written the last summation in the compact form we use in (2.84), is that with the compact form, the recursive nature of the formula becomes less clear; we do not immediately see how $f(x)$ depends on $f(0), f(1), \ldots, f(x-1)$ like in (2.83). On the other hand, in such a respect, we do not gain anything by writing the first term in (2.83) in the less compact form of (2.85), so it seems appropriate to use the compact form of that term. This reasoning may lead to apparent notational inconsistencies even within the same formula.

Chapter 3
Compound Mixed Poisson Distributions

Summary

In Sect. 1.5, we defined mixed distributions and described some insurance applications. In the present chapter, we shall concentrate on a special class of mixed distributions, the mixed Poisson distributions. Within insurance mathematics, this is the most common class of mixed distributions. Our main emphasis will be on recursions for mixed Poisson distributions and compound mixed Poisson distributions.

In Sect. 3.1, we introduce the mixed Poisson distribution and discuss some of its properties.

We then turn to compound mixed Poisson distributions. First, in Sect. 3.2, we discuss the Gamma mixing distribution as a simple example. Then, in Sect. 3.3, we turn to a rather general setting where we deduce a recursive procedure. In the special case of a finite mixture, we compare this procedure with an alternative procedure in Sect. 3.4.

The procedure of Sect. 3.3 can be simplified when the mixing distribution belongs to the Willmot class. This is the topic of Sect. 3.5, and the special case of evaluation of the counting distribution is considered in Sect. 3.6. In Sect. 3.7, we discuss some invariance properties of the Willmot class.

Finally, in Sect. 3.8, we look at some specific parametric classes of mixing distributions within the Willmot class.

3.1 Mixed Poisson Distributions

Let Θ be a positive random variable with distribution U. We assume that for all $\theta > 0$, the conditional distribution p_θ of the non-negative integer-valued random variable N given that $\Theta = \theta$ is the Poisson distribution $\mathrm{Po}(\theta)$, that is,

$$p_\theta(n) = \frac{\theta^n}{n!}\mathrm{e}^{-\theta}. \qquad (n = 0, 1, 2, \dots; \theta > 0) \tag{3.1}$$

Then the unconditional distribution p of N is a mixed Poisson distribution with mixing distribution U, that is,

$$p(n) = \int_{(0,\infty)} p_\theta(n)\,\mathrm{d}U(\theta) = \int_{(0,\infty)} \frac{\theta^n}{n!}\mathrm{e}^{-\theta}\,\mathrm{d}U(\theta) = \frac{(-1)^n}{n!}\gamma_U^{(n)}(1).$$
$$(n = 0, 1, 2, \dots) \tag{3.2}$$

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_3,

Application of (2.12) gives

$$\tau_p(s) = \mathrm{E}\, s^N = \mathrm{E}\,\mathrm{E}[s^N|\Theta] = \mathrm{E}\,\tau_{p_\Theta}(s) = \mathrm{E}\,\mathrm{e}^{\Theta(s-1)},$$

that is,

$$\tau_p(s) = \omega_U(s-1) = \gamma_U(1-s). \tag{3.3}$$

This gives that $\tau_p(s+1) = \omega_U(s)$, and by application of (1.15) and (1.13), we obtain that

$$\nu_p(j) = \mu_U(j). \qquad (j = 0, 1, 2, \dots)$$

In particular, when U is concentrated in a positive number λ, that is, p is the Poisson distribution $\mathrm{Po}(\lambda)$, then this gives

$$\nu_p(j) = \lambda^j. \qquad (j = 0, 1, 2, \dots) \tag{3.4}$$

In Sect. 2.3.2, we pointed out that for the Poisson distribution, the variance is equal to the mean. For our mixed Poisson distribution, we obtain

$$\begin{aligned}\mathrm{Var}\, N &= \mathrm{E}\,\mathrm{Var}[N|\Theta] + \mathrm{Var}\,\mathrm{E}[N|\Theta] = \mathrm{E}\,\mathrm{E}[N|\Theta] + \mathrm{Var}\,\Theta \\ &= \mathrm{E}\, N + \mathrm{Var}\,\Theta \geq \mathrm{E}\, N\end{aligned}$$

with strict inequality if the mixing distribution is non-degenerate.

In the following theorem, we relate a convolution of mixed Poisson distributions to their mixing distributions.

Theorem 3.1 *The convolution of a finite number of mixed Poisson distributions is a mixed Poisson distribution whose mixing distribution is the convolution of the mixing distributions of these distributions.*

Proof We shall first show that the theorem holds for the convolution of two mixed Poisson distributions. It is most convenient to prove this result in terms of random variables, so let N_1 and N_2 be independent random variables with mixed Poisson distributions, and let Θ_1 and Θ_2 be the corresponding mixing variables. Then we know that given Θ_1 and Θ_2, N_1 and N_2 are conditionally Poisson distributed with parameter Θ_1 and Θ_2 respectively, and from Theorem 2.7 then follows that $N_1 + N_2$ is conditionally Poisson distributed with parameter $\Theta_1 + \Theta_2$. Hence, unconditionally, $N_1 + N_2$ is mixed Poisson distributed, and the mixing distribution is the distribution of $\Theta_1 + \Theta_2$, that is, the convolution of the mixing distributions of N_1 and N_2. Hence, the theorem holds for the convolution of two mixed Poisson distributions, and by induction it follows that it holds for the convolution of an arbitrary finite number of mixed Poisson distributions. □

3.2 Gamma Mixing Distribution

In the following, we shall consider evaluation of a compound mixed Poisson distribution $f = p \vee h$ with mixing distribution U and severity distribution $h \in \mathcal{P}_{10}$. We first consider the special case when U is the *Gamma distribution* Gamma(α, β) with density

$$u(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}. \qquad (\theta > 0; \beta, \alpha > 0) \tag{3.5}$$

From (3.2), we obtain that for $n = 0, 1, 2, \ldots$

$$p(n) = \int_0^\infty \frac{\theta^n}{n!}e^{-\theta}\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}d\theta = \frac{\beta^\alpha}{n!\Gamma(\alpha)}\int_0^\infty \theta^{\alpha+n-1}e^{-(\beta+1)\theta}d\theta$$
$$= \frac{\beta^\alpha}{n!\Gamma(\alpha)}\frac{\Gamma(\alpha+n)}{(\beta+1)^{\alpha+n}}.$$

By partial integration, (1.1) gives that

$$\Gamma(x+1) = x\Gamma(x). \qquad (x > 0)$$

Hence,

$$p(n) = \binom{\alpha+n-1}{n}\left(\frac{1}{\beta+1}\right)^n\left(\frac{\beta}{\beta+1}\right)^\alpha. \tag{3.6}$$

By comparison with (2.42), we see that p is the negative binomial distribution NB$(\alpha, (\beta+1)^{-1})$. Then Table 2.3 gives the recursion

$$f(x) = \frac{1}{\beta+1-h(0)}\sum_{y=1}^{x}\left(1+(\alpha-1)\frac{y}{x}\right)h(y)f(x-y) \qquad (x = 1, 2, \ldots)$$
$$f(0) = \left(\frac{\beta}{\beta+1-h(0)}\right)^\alpha.$$

In particular, we obtain

$$p(n) = \frac{1}{\beta+1}\left(1+\frac{\alpha-1}{n}\right)p(n-1) \qquad (n = 1, 2, \ldots)$$
$$p(0) = \left(\frac{\beta}{\beta+1}\right)^\alpha.$$

In Sect. 3.1, we showed that for a mixed Poisson distribution with non-degenerate mixing distribution, the variance is greater than the mean. As a negative binomial distribution can be expressed as a mixed Poisson distribution with Gamma mixing distribution, it follows that for a negative binomial distribution, the variance is greater than the mean. In Sect. 2.3.2, we showed this by other means.

We shall return to mixed Poisson distributions with Gamma mixing distribution in Example 4.2.

Unfortunately, it is not in general that simple to evaluate f and p. In the following, we shall consider some more complicated procedures that can be applied more generally.

3.3 General Recursion

We now drop the assumption that the mixing distribution is a Gamma distribution.

Let

$$v_i(x) = \int_{(0,\infty)} \theta^i f_\theta(x)\,\mathrm{d}U(\theta) \qquad (x, i = 0, 1, 2, \dots) \tag{3.7}$$

with $f_\theta = p_\theta \vee h$. In particular, we have $v_0 = f$. By application of Table 2.3, we obtain that for $i = 0, 1, 2, \dots,$

$$\begin{aligned} v_i(0) &= \int_{(0,\infty)} \theta^i f_\theta(0)\,\mathrm{d}U(\theta) = \int_{(0,\infty)} \theta^i \mathrm{e}^{-\theta(1-h(0))}\,\mathrm{d}U(\theta) \\ &= (-1)^i \gamma_U^{(i)}(1 - h(0)); \end{aligned} \tag{3.8}$$

letting $i = 0$ gives

$$f(0) = \int_{(0,\infty)} \mathrm{e}^{-\theta(1-h(0))}\,\mathrm{d}U(\theta) = \gamma_U(1 - h(0)). \tag{3.9}$$

From (2.7), we obtain

$$f_\theta(x) = \frac{\theta}{x} \sum_{y=1}^{x} y h(y) f_\theta(x - y). \qquad (x = 1, 2, \dots) \tag{3.10}$$

Multiplication by $\theta^i\,\mathrm{d}U(\theta)$ and integration gives

$$v_i(x) = \frac{1}{x} \sum_{y=1}^{x} y h(y) v_{i+1}(x - y). \quad (x = 1, 2, \dots; i = 0, 1, 2, \dots) \tag{3.11}$$

We can now evaluate $f(0), f(1), f(2), \dots, f(x)$ by the following algorithm:

Evaluate $f(0)$ by (3.9).
For $y = 1, 2, \dots, x$:
  Evaluate $v_y(0)$ by (3.8).
  For $z = 1, 2, \dots, y$:
    Evaluate $v_{y-z}(z)$ by (3.11).
  Let $f(y) = v_0(y)$.

If h satisfies the conditions of Theorem 2.3, then we can replace (3.11) in this algorithm with

$$v_i(x) = \sum_{y=1}^{r} \left(\frac{\eta(y)}{x} v_{i+1}(x-y) + \left(1 - \frac{y}{x}\right) \chi(y) v_i(x-y) \right), \tag{3.12}$$

$$(x = 1, 2, \ldots; i = 0, 1, 2, \ldots)$$

which is found from (2.18) by proceeding like with the deduction of (3.11).

3.4 Finite Mixtures

Let us now consider the situation when the mixing distribution U is given by

$$\Pr(\Theta = \theta_k) = u_k \qquad (k = 1, 2, \ldots, t) \tag{3.13}$$

with $\sum_{k=1}^{t} u_k = 1$ for some positive integer t. We then have

$$f = \sum_{k=1}^{t} u_k f_{\theta_k}. \tag{3.14}$$

As an alternative to the algorithm of Sect. 3.3, we can evaluate each f_{θ_k} by Theorem 2.2 and then f by (3.14). We refer to the former method as Method A and the latter as Method B. The question is which of them to prefer for evaluation of $f(0), f(1), \ldots, f(x)$. To answer this question, we need an optimality criterion. A criterion that is sometimes applied for comparison of methods, is the number of elementary algebraic operations, that is, addition, subtraction, multiplication, and division. Under this criterion, one would intuitively expect Method A to be preferable when x is small and/or t is large. With Method B, we would have to apply the recursion of Theorem 2.2, and that would be inefficient when t is large so Method A seems better. On the other hand, with Method A, we would have to maintain x v_is, and that would be inefficient when x is large, so then Method B seems better.

3.5 Willmot Mixing Distribution

When x is large, the algorithm at the end of Sect. 3.3 can be rather time- and storage-consuming. We shall now show that if U is continuous with density u on an interval (γ, δ) with $0 \leq \gamma < \delta \leq \infty$ that satisfies the condition

$$\frac{d}{d\theta} \ln u(\theta) = \frac{u'(\theta)}{u(\theta)} = \frac{\sum_{i=0}^{k} \eta(i)\theta^i}{\sum_{i=0}^{k} \chi(i)\theta^i} \qquad (\gamma < \theta < \delta) \tag{3.15}$$

for some non-negative integer k, then we need to evaluate v_i only for $i = 0, 1, \ldots, k$. We call this class of mixing distributions the *Willmot class*. For the rest of the present chapter, we shall assume that (3.15) is satisfied.

We shall need the auxiliary functions

$$w_\theta(x) = f_\theta(x)u(\theta)\sum_{i=0}^{k}\chi(i)\theta^i \qquad (x = 0, 1, 2, \ldots; \gamma < \theta < \delta) \tag{3.16}$$

$$\rho(i) = (1 - h(0))\chi(i) - \eta(i) - (i+1)\chi(i+1) \tag{3.17}$$
$$(i = -1, 0, 1, \ldots, k)$$

with $\chi(-1) = \eta(-1) = \chi(k+1) = 0$.

Theorem 3.2 *If f is a compound mixed Poisson distribution with severity distribution $h \in \mathcal{P}_{10}$ and continuous mixing distribution on the interval (γ, δ) with differentiable density u that satisfies (3.15), and $w_{\gamma+}(x)$ and $w_{\delta-}(x)$ exist and are finite for all non-negative integers x, then*

$$\rho(k)v_k(x) = \sum_{y=1}^{x}h(y)\sum_{i=0}^{k}\chi(i)v_i(x-y) - \sum_{i=0}^{k-1}\rho(i)v_i(x)$$
$$+ w_{\gamma+}(x) - w_{\delta-}(x). \qquad (x = 1, 2, \ldots) \tag{3.18}$$

Proof Application of (3.7), (3.15), and partial integration gives that for $x = 1, 2, \ldots$

$$\sum_{i=0}^{k}\eta(i)v_i(x) = \sum_{i=0}^{k}\eta(i)\int_\gamma^\delta \theta^i f_\theta(x)u(\theta)\,\mathrm{d}\theta = \sum_{i=0}^{k}\chi(i)\int_\gamma^\delta \theta^i f_\theta(x)u'(\theta)\,\mathrm{d}\theta$$
$$= w_{\delta-}(x) - w_{\gamma+}(x) - \sum_{i=0}^{k}\chi(i)\int_\gamma^\delta\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\theta^i f_\theta(x)\right)u(\theta)\,\mathrm{d}\theta$$
$$= w_{\delta-}(x) - w_{\gamma+}(x)$$
$$- \sum_{i=0}^{k}\chi(i)\int_\gamma^\delta\left(i\theta^{i-1}f_\theta(x) + \theta^i\frac{\mathrm{d}}{\mathrm{d}\theta}f_\theta(x)\right)u(\theta)\,\mathrm{d}\theta,$$

that is,

$$\sum_{i=0}^{k}\eta(i)v_i(x) = w_{\delta-}(x) - w_{\gamma+}(x) - \sum_{i=1}^{k}i\chi(i)v_{i-1}(x)$$
$$- \sum_{i=0}^{k}\chi(i)\int_\gamma^\delta\theta^i\left(\frac{\mathrm{d}}{\mathrm{d}\theta}f_\theta(x)\right)u(\theta)\,\mathrm{d}\theta. \tag{3.19}$$

We have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\theta}f_\theta(x) &= \sum_{n=1}^{\infty} h^{n*}(x)\frac{\mathrm{d}}{\mathrm{d}\theta}\frac{\theta^n}{n!}\mathrm{e}^{-\theta} = \sum_{n=1}^{\infty} h^{n*}(x)\frac{1}{n!}(n\theta^{n-1}\mathrm{e}^{-\theta} - \theta^n\mathrm{e}^{-\theta})\\ &= \sum_{n=1}^{\infty} h^{n*}(x)(p_\theta(n-1) - p_\theta(n)) = (h * f_\theta)(x) - f_\theta(x)\\ &= \sum_{y=0}^{x} h(y)f_\theta(x-y) - f_\theta(x) = \sum_{y=1}^{x} h(y)f_\theta(x-y) - (1-h(0))f_\theta(x).\end{aligned}$$

Insertion in (3.19) gives

$$\begin{aligned}\sum_{i=0}^{k}\eta(i)v_i(x) &= w_{\delta-}(x) - w_{\gamma+}(x) - \sum_{i=0}^{k-1}(i+1)\chi(i+1)v_i(x)\\ &\quad - \sum_{i=0}^{k}\chi(i)\left(\sum_{y=1}^{x} h(y)v_i(x-y) - (1-h(0))v_i(x)\right).\end{aligned}$$

After some rearranging, we obtain

$$\sum_{i=0}^{k}\rho(i)v_i(x) = \sum_{i=0}^{k}\chi(i)\sum_{y=1}^{x} h(y)v_i(x-y) + w_{\gamma+}(x) - w_{\delta-}(x),$$

from which (3.18) follows. □

If $\rho(k) \neq 0$, then (3.18) gives

$$\begin{aligned}v_k(x) = \frac{1}{\rho(k)}\Bigg(&\sum_{y=1}^{x} h(y)\sum_{i=0}^{k}\chi(i)v_i(x-y)\\ &- \sum_{i=0}^{k-1}\rho(i)v_i(x) + w_{\gamma+}(x) - w_{\delta-}(x)\Bigg). \quad (x = 1, 2, \dots)\end{aligned} \tag{3.20}$$

In this case, we can evaluate $f(0), f(1), f(2), \dots, f(x)$ with x being an integer greater than k by the following algorithm:

Evaluate $f(0)$ by (3.9).
For $y = 1, 2, \dots, k$:
  Evaluate $v_y(0)$ by (3.8).
  For $z = 1, 2, \dots, y$:
    Evaluate $v_{y-z}(z)$ by (3.11).
  Let $f(y) = v_0(y)$.

For $y = k+1, k+2, \ldots, x$:
  Evaluate $v_k(y-k)$ by (3.20).
  For $z = 1, 2, \ldots, k$:
    Evaluate $v_{k-z}(y-k+z)$ by (3.11).
  Let $f(y) = v_0(y)$.

If h satisfies the conditions of Theorem 2.3, then we can replace (3.11) with (3.12) in this algorithm.

Let us now consider the condition that $w_{\gamma+}(x)$ and $w_{\delta-}(x)$ should exist and be finite for all positive integers x. For finite γ and δ, this condition holds when $u(\gamma+)$ and $u(\delta-)$ exist and are finite. In particular, we have

$$w_{0+}(x) = f_0(x)u(0+)\chi(0) = 0 \qquad (x = 1, 2, \ldots)$$

as the distribution f_0 is concentrated in zero.

From (3.16) and (3.10) we obtain the recursion

$$w_\theta(x) = \begin{cases} \frac{\theta}{x}\sum_{y=1}^{x} yh(y)w_\theta(x-y) & (x = 1, 2, \ldots) \\ \mathrm{e}^{-\theta(1-h(0))}u(\theta)\sum_{i=0}^{k}\chi(i)\theta^i. & (x = 0) \end{cases} \qquad (\gamma < \theta < \delta) \tag{3.21}$$

If $\gamma > 0$ and $u(\gamma+)$ exists and is finite, then we can evaluate $w_{\gamma+}$ recursively in the same way. Furthermore, we can make $w_{\gamma+}$ vanish by multiplying the numerator and denominator in (3.15) with $\theta - \gamma$ as then $\sum_{i=0}^{k}\chi(i)\theta^i$ is replaced with $(\theta-\gamma)\sum_{i=0}^{k}\chi(i)\theta^i$, which is zero when $\theta = \gamma$. However, then we increase k by one. Analogous for $w_{\delta-}$ when $\delta < \infty$ and $u(\delta-)$ exists and is finite.

3.6 The Counting Distribution

When h is concentrated in one, we get $f = p$. In that case, we can replace v_i, w_θ, and ρ with $\dot{v}_i$, $\dot{w}_\theta$, and $\dot{\rho}$ given by

$$\begin{aligned} \dot{v}_i(n) &= \int_{(0,\infty)} \theta^i p_\theta(n)\,\mathrm{d}U(\theta) = \frac{1}{n!}\int_{(0,\infty)} \theta^{n+i}\mathrm{e}^{-\theta}\,\mathrm{d}U(\theta) \\ &= (n+i)^{(i)}p(n+i) \qquad (n, i = 0, 1, 2, \ldots) \end{aligned} \tag{3.22}$$

$$\dot{w}_\theta(n) = p_\theta(n)u(\theta)\sum_{i=0}^{k}\chi(i)\theta^i = \frac{\theta^n}{n!}\mathrm{e}^{-\theta}u(\theta)\sum_{i=0}^{k}\chi(i)\theta^i$$
$$(n = 0, 1, 2, \ldots; \gamma < \theta < \delta)$$

$$\dot{\rho}(i) = \chi(i) - \eta(i) - (i+1)\chi(i+1). \qquad (i = -1, 0, 1, \ldots, k) \tag{3.23}$$

The recursion (3.21) now reduces to

$$\dot{w}_\theta(n) = \begin{cases} \frac{\theta}{n}\dot{w}_\theta(n-1) & (n = 1, 2, \dots) \\ \mathrm{e}^{-\theta} u(\theta) \sum_{i=0}^{k} \chi(i)\theta^i. & (n = 0) \end{cases} \qquad (\gamma < \theta < \delta) \tag{3.24}$$

From (3.22), we obtain that $p(n) = \dot{v}_n(0)/n!$ for $n = 0, 1, 2, \dots$.

Theorem 3.3 *If p is a mixed Poisson distribution with continuous mixing distribution on the interval (γ, δ) with differentiable density that satisfies (3.15), and $\dot{w}_{\gamma+}(n)$ and $\dot{w}_{\delta-}(n)$ exist and are finite for all non-negative integers n, then*

$$\dot{\rho}(k)p(n) = \sum_{i=1}^{k+1}((n-k)\chi(k-i+1) - \dot{\rho}(k-i))\frac{p(n-i)}{n^{(i)}} + \frac{\dot{w}_{\gamma+}(n-k) - \dot{w}_{\delta-}(n-k)}{n^{(k)}}. \quad (n = k+1, k+2, \dots) \tag{3.25}$$

Proof From (3.18), we obtain that for $x = 1, 2, \dots$,

$$\dot{\rho}(k)\dot{v}_k(x) = \sum_{i=0}^{k} \chi(i)\dot{v}_i(x-1) - \sum_{i=0}^{k-1}\dot{\rho}(i)\dot{v}_i(x) + \dot{w}_{\gamma+}(x) - \dot{w}_{\delta-}(x),$$

and insertion of (3.22) gives

$$\begin{aligned} &\dot{\rho}(k)(x+k)^{(k)}p(x+k) \\ &= \sum_{i=0}^{k}\chi(i)(x-1+i)^{(i)}p(x-1+i) - \sum_{i=0}^{k-1}\dot{\rho}(i)(x+i)^{(i)}p(x+i) \\ &\quad + \dot{w}_{\gamma+}(x) - \dot{w}_{\delta-}(x) \\ &= \sum_{i=-1}^{k-1}\chi(i+1)(x+i)^{(i+1)}p(x+i) - \sum_{i=0}^{k-1}\dot{\rho}(i)(x+i)^{(i)}p(x+i) \\ &\quad + \dot{w}_{\gamma+}(x) - \dot{w}_{\delta-}(x) \\ &= \sum_{i=-1}^{k-1}(x\chi(i+1) - \dot{\rho}(i))(x+i)^{(i)}p(x+i) + \dot{w}_{\gamma+}(x) - \dot{w}_{\delta-}(x). \end{aligned}$$

Letting $x = n - k$ for $n = k+1, k+2, \dots$ gives

$$\dot{\rho}(k)n^{(k)}p(n) = \sum_{i=-1}^{k-1}((n-k)\chi(i+1) - \dot{\rho}(i))(n-k+i)^{(i)}p(n-k+i) + \dot{w}_{\gamma+}(n-k) - \dot{w}_{\delta-}(n-k),$$

and by division by $n^{(k)}$ we obtain

$$\dot{\rho}(k)p(n) = \sum_{i=-1}^{k-1} ((n-k)\chi(i+1) - \dot{\rho}(i))\frac{p(n-k+i)}{n^{(k-i)}} + \frac{\dot{w}_{\gamma+}(n-k) - \dot{w}_{\delta-}(n-k)}{n^{(k)}}.$$

Finally, by changing the summation variable, we obtain (3.25). □

If $\dot{\rho}(k) \neq 0$, then (3.25) gives the recursion

$$p(n) = \frac{1}{\dot{\rho}(k)}\left(\sum_{i=1}^{k+1}((n-k)\chi(k-i+1) - \dot{\rho}(k-i))\frac{p(n-i)}{n^{(i)}} + \frac{\dot{w}_{\gamma+}(n-k) - \dot{w}_{\delta-}(n-k)}{n^{(k)}}\right). \quad (n = k+1, k+2, \dots) \tag{3.26}$$

Let us consider the special case $k = 1$. Then (3.26) gives

$$p(n) = \frac{1}{\dot{\rho}(1)}\left(b_{\gamma+}(n) - b_{\delta-}(n) + \left(\chi(1) - \frac{\chi(1) + \dot{\rho}(0)}{n}\right)p(n-1) + \frac{\chi(0)}{n}p(n-2)\right) \qquad (n = 2, 3, \dots) \tag{3.27}$$

with

$$b_\theta(n) = \frac{\dot{w}_\theta(n)}{\theta}. \qquad (n = 0, 1, 2, \dots; \gamma < \theta < \delta) \tag{3.28}$$

Thus,

$$p(n) = \frac{1}{\dot{\rho}(1)}\Bigg(I(n=1)(\dot{\rho}(1)p(1) + \dot{\rho}(0)p(0) + b_{\delta-}(1) - b_{\gamma+}(1)) + b_{\gamma+}(n) - b_{\delta-}(n) + \left(\chi(1) - \frac{\chi(1) + \dot{\rho}(0)}{n}\right)p(n-1) + \frac{\chi(0)}{n}p(n-2)\Bigg), \qquad (n = 1, 2, \dots)$$

which is in the form (5.6). Insertion in (5.8) gives

$$f(x) = \frac{1}{\dot{\rho}(1) - \chi(1)h(0)}\Bigg((\dot{\rho}(1)p(1) + \dot{\rho}(0)p(0) + b_{\delta-}(1) - b_{\gamma+}(1))h(x) + (b_{\gamma+} \vee h)(x) - (b_{\delta-} \vee h)(x)$$

$$+\sum_{y=1}^{x}\left(\left(\chi(1)-(\chi(1)+\dot{\rho}(0))\frac{y}{x}\right)h(y)+\frac{\chi(0)}{2}\frac{y}{x}h^{2*}(y)\right)f(x-y)\bigg).$$
$$(x=1,2,\dots) \qquad (3.29)$$

By (3.24), for $\gamma < \theta < \delta$, $\dot{w}_\theta$ and, hence, b_θ are proportional to the Poisson distribution Po(θ) as given by (2.6), so by Table 2.3, we can evaluate $b_\theta \vee h$ recursively by

$$(b_\theta \vee h)(x) = \frac{\theta}{x}\sum_{y=1}^{x} yh(y)(b_\theta \vee h)(x-y), \qquad (x=1,2,\dots) \qquad (3.30)$$

and by (3.28) and (3.24), we obtain the initial value

$$(b_\theta \vee h)(0) = \left(\frac{\chi(0)}{\theta}+\chi(1)\right)u(\theta)\mathrm{e}^{-\theta(1-h(0))}. \qquad (3.31)$$

When $\gamma > 0$, we can evaluate $b_{\gamma+} \vee h$ by such a recursion, and, when $\delta < \infty$, we can do it with $b_{\delta-} \vee h$.

When $k = 1$, the recursion (3.29) is more efficient than the algorithm given after formula (3.20).

3.7 Invariance Properties in the Willmot Class

3.7.1 Introduction

In Sect. 3.7, we shall consider how the condition (3.15) is affected by some types of transforms of the mixing distribution. In this connection, it will be convenient to introduce a mixing variable Θ having distribution U with continuous density satisfying (3.15). We distinguish the notation for the transformed situation by adding a tilde to the notation for the transformed situation.

3.7.2 Scaling

We let $\tilde{\Theta} = c\Theta$ for some positive constant c. Then we have

$$\tilde{U}(\theta) = \Pr(\tilde{\Theta} \le \theta) = \Pr(c\Theta \le \theta) = \Pr(\Theta \le \theta/c) = U(\theta/c),$$

so that for $\theta \in (\tilde{\gamma}, \tilde{\delta})$ with $\tilde{\gamma} = c\gamma$ and $\tilde{\delta} = c\delta$

$$\tilde{u}(\theta) = \frac{\mathrm{d}}{\mathrm{d}\theta}\tilde{U}(\theta) = \frac{\mathrm{d}}{\mathrm{d}\theta}U(\theta/c) = \frac{u(\theta/c)}{c},$$

which gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\theta}\ln\tilde{u}(\theta) &= \frac{\mathrm{d}}{\mathrm{d}\theta}\ln\frac{u(\theta/c)}{c} = \frac{\mathrm{d}}{\mathrm{d}\theta}\ln u(\theta/c) = \frac{1}{u(\theta/c)}\frac{\mathrm{d}}{\mathrm{d}\theta}u(\theta/c)\\ &= \frac{1}{c}\frac{u'(\theta/c)}{u(\theta/c)} = \frac{1}{c}\frac{\sum_{i=0}^{k}\eta(i)(\theta/c)^i}{\sum_{i=0}^{k}\chi(i)(\theta/c)^i} = \frac{\sum_{i=0}^{k}\tilde{\eta}(i)\theta^i}{\sum_{i=0}^{k}\tilde{\chi}(i)\theta^i}\end{aligned}$$

with $\tilde{\eta}(i) = \eta(i)c^{-i}$ and $\tilde{\chi}(i) = \chi(i)c^{1-i}$ for $i = 0, 1, 2, \ldots, k$.

This invariance of (3.15) under scale transforms can be convenient e.g. in connection with experience rating. In some insurance classes where experience rating is applied, there could be an objective risk measure that could vary between policies and for each policy between years. For instance, in worker's compensation insurance, the risk volume could be the number of workers insured under the policy. Then it seems reasonable to assume that the random Poisson parameter is proportional to the number of workers, so that this parameter is $c\Theta$ with c denoting the number of workers. In this situation, it would be unreasonable if (3.15) should hold only for one value of c.

3.7.3 Shifting

Let $\tilde{\Theta} = \Theta + c$ for some positive constant c. Then, for $\theta \geq c$, we have

$$\tilde{U}(\theta) = \Pr(\tilde{\Theta} \leq \theta) = \Pr(\Theta + c \leq \theta) = \Pr(\Theta \leq \theta - c) = U(\theta - c),$$

so that for $\theta \in (\tilde{\gamma}, \tilde{\delta})$ with $\tilde{\gamma} = \gamma + c$ and $\tilde{\delta} = \delta + c$, we have $\tilde{u}(\theta) = u(\theta - c)$. This gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\theta}\ln\tilde{u}(\theta) &= \frac{\mathrm{d}}{\mathrm{d}\theta}\ln u(\theta - c) = \frac{u'(\theta - c)}{u(\theta - c)} = \frac{\sum_{i=0}^{k}\eta(i)(\theta - c)^i}{\sum_{i=0}^{k}\chi(i)(\theta - c)^i}\\ &= \frac{\sum_{i=0}^{k}\eta(i)\sum_{j=0}^{i}\binom{i}{j}\theta^j(-c)^{i-j}}{\sum_{i=0}^{k}\chi(i)\sum_{j=0}^{i}\binom{i}{j}\theta^j(-c)^{i-j}} = \frac{\sum_{j=0}^{k}\theta^j\sum_{i=j}^{k}\eta(i)\binom{i}{j}(-c)^{i-j}}{\sum_{j=0}^{k}\theta^j\sum_{i=j}^{k}\chi(i)\binom{i}{j}(-c)^{i-j}}\\ &= \frac{\sum_{j=0}^{k}\tilde{\eta}(j)\theta^j}{\sum_{j=0}^{k}\tilde{\chi}(j)\theta^j}\end{aligned}$$

with

$$\tilde{\eta}(j) = \sum_{i=j}^{k}\eta(i)\binom{i}{j}(-c)^{i-j}; \qquad \tilde{\chi}(j) = \sum_{i=j}^{k}\chi(i)\binom{i}{j}(-c)^{i-j}.$$
$$(j = 0, 1, 2, \ldots, k)$$

In experience rating, shifting can be interpreted as if there are two sorts of claims, one sort where the number of claims depends on the individual unknown risk properties of the policy, and another sort where this is not the case. In automobile insurance, these two sorts could be accident and theft. Given the unknown random risk parameter Θ of the policy, the number of the first sort of claims would be Poisson distributed with parameter Θ, whereas the number of the second sort of claims would be Poisson distributed with parameter c independent of Θ. An example of such a shifted distribution is the Delaporte distribution, which will be studied in Example 5.3.

3.7.4 Truncating

Now let the transformed distribution $\tilde{U}$ be the conditional distribution Θ given that $\tilde{\gamma} < \Theta < \tilde{\delta}$ with $\gamma \le \tilde{\gamma} < \tilde{\delta} \le \delta$. Then, for $\theta \in (\tilde{\gamma}, \tilde{\delta})$, we have

$$\tilde{u}(\theta) = \frac{u(\theta)}{U(\tilde{\delta}) - U(\tilde{\gamma})},$$

so that

$$\frac{\mathrm{d}}{\mathrm{d}\theta} \ln \tilde{u}(\theta) = \frac{\mathrm{d}}{\mathrm{d}\theta} \ln \frac{u(\theta)}{U(\tilde{\delta}) - U(\tilde{\gamma})} = \frac{\mathrm{d}}{\mathrm{d}\theta} \ln u(\theta) = \frac{\sum_{i=0}^{k} \eta(i)\theta^i}{\sum_{i=0}^{k} \chi(i)\theta^i},$$

that is $\tilde{\eta} = \eta$ and $\tilde{\chi} = \chi$.

3.7.5 Power Transform

Let $\tilde{\Theta} = \Theta^{1/c}$ for some non-zero constant c. Then

$$\begin{aligned} \tilde{U}(\theta) &= \Pr(\tilde{\Theta} \le \theta) = \Pr(\Theta^{1/c} \le \theta) \\ &= \begin{cases} \Pr(\Theta \le \theta^c) = U(\theta^c) & (c > 0) \\ \Pr(\Theta \ge \theta^c) = 1 - U(\theta^c) & (c < 0) \end{cases} \end{aligned}$$

so that

$$\tilde{u}(\theta) = \frac{\mathrm{d}}{\mathrm{d}\theta} \tilde{U}(\theta) = |c|\theta^{c-1} u(\theta^c)$$

for $\theta \in (\tilde{\gamma}, \tilde{\delta})$ with

$$(\tilde{\gamma}, \tilde{\delta}) = \begin{cases} (\gamma^{1/c}, \delta^{1/c}) & (c > 0) \\ (\delta^{1/c}, \gamma^{1/c}). & (c < 0) \end{cases}$$

This gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\theta}\ln\tilde{u}(\theta) &= \frac{\mathrm{d}}{\mathrm{d}\theta}\ln|c|\theta^{c-1}u(\theta^c) = \frac{\mathrm{d}}{\mathrm{d}\theta}((c-1)\ln\theta + \ln u(\theta^c))\\ &= \frac{c-1}{\theta} + c\theta^{c-1}\frac{u'(\theta^c)}{u(\theta^c)} = \frac{c-1}{\theta} + c\theta^{c-1}\frac{\sum_{i=0}^k \eta(i)\theta^{ci}}{\sum_{i=0}^k \chi(i)\theta^{ci}}\\ &= \frac{\sum_{i=0}^k (c-1)\chi(i)\theta^{ci} + \sum_{i=0}^k c\eta(i)\theta^{c(i+1)}}{\sum_{i=0}^k \chi(i)\theta^{ci+1}},\end{aligned}$$

that is,

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln\tilde{u}(\theta) = \frac{\sum_{i=0}^{k+1}((c-1)\chi(i) + c\eta(i-1))\theta^{ci}}{\sum_{i=0}^k \chi(i)\theta^{ci+1}}. \tag{3.32}$$

If c is a positive integer greater than one, then this gives

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln\tilde{u}(\theta) = \frac{\sum_{j=0}^{\tilde{k}} \tilde{\eta}(j)\theta^j}{\sum_{j=0}^{\tilde{k}} \tilde{\chi}(j)\theta^j}$$

with $\tilde{k} = c(k+1)$,

$$\tilde{\eta}(ci) = (c-1)\chi(i) + c\eta(i-1); \qquad \tilde{\chi}(ci+1) = \chi(i),$$
$$(i = 0, 1, 2, \ldots, k+1)$$

and $\tilde{\chi}(j)$ and $\tilde{\eta}(j)$ equal to zero for all other values of j.

If c is a negative integer, then multiplication by $\theta^{-c(k+1)}$ in the numerator and denominator of (3.32) gives

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln\tilde{u}(\theta) = \frac{\sum_{i=0}^{k+1}((c-1)\chi(i) + c\eta(i-1))\theta^{-c(k+1-i)}}{\sum_{i=0}^k \chi(i)\theta^{-c(k+1-i)+1}} = \frac{\sum_{j=0}^{\tilde{k}} \tilde{\eta}(j)\theta^j}{\sum_{j=0}^{\tilde{k}} \tilde{\chi}(j)\theta^j}$$

with $\tilde{k} = -c(k+1)$,

$$\begin{aligned}&\tilde{\eta}(-c(k+1-i)) = (c-1)\chi(i) + c\eta(i-1)\\ &\tilde{\chi}(-c(k+1-i)+1) = \chi(i),\end{aligned} \qquad (i = 0, 1, 2, \ldots, k+1)$$

and $\tilde{\chi}(j)$ and $\tilde{\eta}(j)$ equal to zero for all other values of j.

3.8 Special Classes of Mixing Distributions

3.8.1 Shifted Pareto Distribution

Let U be the *shifted Pareto distribution* $\mathrm{SPar}(\alpha, \gamma)$ with density

$$u(\theta) = \frac{\alpha\gamma^{\alpha}}{\theta^{\alpha+1}}. \qquad (\theta > \gamma; \alpha, \gamma > 0) \tag{3.33}$$

Insertion in (3.2) gives

$$p(n) = \frac{\alpha\gamma^{\alpha}}{n!}\int_{\gamma}^{\infty} \theta^{n-\alpha-1}\mathrm{e}^{-\theta}\mathrm{d}\theta, \qquad (n = 0, 1, 2, \ldots)$$

and by partial integration, we obtain the recursion

$$p(n) = \frac{\alpha\gamma^{n-1}}{n!}\mathrm{e}^{-\gamma} + \left(1 - \frac{\alpha+1}{n}\right)p(n-1) \qquad (n = 1, 2, \ldots)$$

with initial value

$$p(0) = \alpha\gamma^{\alpha}\int_{\gamma}^{\infty} \theta^{-\alpha-1}\mathrm{e}^{-\theta}\mathrm{d}\theta.$$

From (2.78), we get

$$f(x) = \frac{1}{1-h(0)}\left((b_{\gamma+} \vee h)(x) + \sum_{y=1}^{x}\left(1 - (\alpha+1)\frac{y}{x}\right)h(y)f(x-y)\right)$$
$$(x = 1, 2, \ldots)$$

with

$$b_{\gamma+}(n) = \frac{\alpha\gamma^{n-1}}{n!}\mathrm{e}^{-\gamma}, \qquad (n = 0, 1, 2, \ldots)$$

and insertion of (3.33) in (3.9) gives the initial value

$$f(0) = \alpha\gamma^{\alpha}\int_{\gamma}^{\infty} \frac{\mathrm{e}^{-\theta(1-h(0))}}{\theta^{\alpha+1}}\mathrm{d}\theta.$$

We could also have obtained these recursions from (3.27) and (3.29).

3.8.2 Pareto Distribution

Let U be the *Pareto distribution* $\mathrm{Par}(\alpha, \beta)$ with density

$$u(\theta) = \frac{\alpha\beta^{\alpha}}{(\beta+\theta)^{\alpha+1}}. \qquad (\theta > 0; \alpha, \beta > 0) \tag{3.34}$$

Then

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln u(\theta)=\frac{-\alpha-1}{\beta+\theta}. \qquad (\theta>0)$$

Thus, (3.15) is satisfied with

$$k=1; \qquad \gamma=0; \qquad \delta=\infty \tag{3.35}$$

$$\eta(0)=-\alpha-1; \qquad \eta(1)=0 \tag{3.36}$$

$$\chi(0)=\beta; \qquad \chi(1)=1, \tag{3.37}$$

from which we obtain

$$\dot{\rho}(0)=\alpha+\beta; \qquad \dot{\rho}(1)=1.$$

Insertion in (3.27) gives

$$p(n)=\left(1-\frac{\alpha+\beta+1}{n}\right)p(n-1)+\frac{\beta}{n}p(n-2) \qquad (n=2,3,\dots)$$

and from (3.2) we obtain the initial values

$$p(0)=\alpha\beta^{\alpha}\int_0^{\infty}\frac{\mathrm{e}^{-\theta}}{(\beta+\theta)^{\alpha+1}}\mathrm{d}\theta; \qquad p(1)=\alpha\beta^{\alpha}\int_0^{\infty}\frac{\theta\mathrm{e}^{-\theta}}{(\beta+\theta)^{\alpha+1}}\mathrm{d}\theta.$$

From (3.29), we get the recursion

$$f(x)=\frac{1}{1-h(0)}\Bigg((p(1)+(\alpha+\beta)p(0))h(x)$$
$$+\sum_{y=1}^{x}\left(\left(1-(\alpha+\beta+1)\frac{y}{x}\right)h(y)+\frac{\beta}{2}\frac{y}{x}h^{2*}(y)\right)f(x-y)\Bigg),$$
$$(x=1,2,\dots)$$

and insertion of (3.34) in (3.9) gives the initial value

$$f(0)=\alpha\beta^{\alpha}\int_0^{\infty}\frac{\mathrm{e}^{-\theta(1-h(0))}}{(\beta+\theta)^{\alpha+1}}\mathrm{d}\theta.$$

3.8.3 Truncated Normal Distribution

Let U be the *truncated normal distribution* $\mathrm{TN}(\xi,\sigma)$ with density

$$u(\theta)=\frac{\mathrm{e}^{-\frac{1}{2\sigma^2}(\theta-\xi)^2}}{\int_0^{\infty}\mathrm{e}^{-\frac{1}{2\sigma^2}(y-\xi)^2}\mathrm{d}y}. \qquad (\theta>0;\ \xi\in\mathbb{R};\ \sigma>0) \tag{3.38}$$

Then

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln u(\theta) = \frac{\xi - \theta}{\sigma^2},$$

so that (3.15) is satisfied with

$$k = 1; \qquad \gamma = 0; \qquad \delta = \infty \tag{3.39}$$

$$\eta(0) = \xi; \qquad \eta(1) = -1 \tag{3.40}$$

$$\chi(0) = \sigma^2; \qquad \chi(1) = 0, \tag{3.41}$$

from which we obtain

$$\dot{\rho}(0) = \sigma^2 - \xi; \qquad \dot{\rho}(1) = 1.$$

Insertion in (3.27) gives

$$p(n) = \frac{\xi - \sigma^2}{n} p(n-1) + \frac{\sigma^2}{n} p(n-2), \qquad (n = 2, 3, \dots)$$

and by insertion of (3.38) in (3.2), we obtain

$$p(n) = \frac{\int_0^\infty \theta^n \mathrm{e}^{-\theta - \frac{1}{2\sigma^2}(\theta-\xi)^2} \mathrm{d}\theta}{n! \int_0^\infty \mathrm{e}^{-\frac{1}{2\sigma^2}(\theta-\xi)^2} \mathrm{d}\theta}, \qquad (n = 0, 1, 2, \dots)$$

which gives in particular the initial values $p(0)$ and $p(1)$.

From (3.29), we get the recursion

$$f(x) = (p(1) + (\sigma^2 - \xi)p(0))h(x) + \frac{1}{x}\sum_{y=1}^{x} y\left((\xi - \sigma^2)h(y) + \frac{\sigma^2}{2}h^{2*}(y)\right) f(x-y), \qquad (x = 1, 2, \dots)$$

and insertion of (3.38) in (3.9) gives the initial value

$$f(0) = \frac{\int_0^\infty \mathrm{e}^{-\theta(1-h(0)) - \frac{1}{2\sigma^2}(\theta-\xi)^2} \mathrm{d}\theta}{\int_0^\infty \mathrm{e}^{-\frac{1}{2\sigma^2}(\theta-\xi)^2} \mathrm{d}\theta}.$$

3.8.4 Inverse Gauss Distribution

Let U be the *inverse Gauss distribution* IGauss(ξ, β) with density

$$u(\theta) = \frac{\xi}{\sqrt{2\pi\beta}} \theta^{-\frac{3}{2}} \mathrm{e}^{-\frac{(\theta-\xi)^2}{2\beta\theta}} \qquad (\theta > 0; \xi, \beta > 0) \tag{3.42}$$

where π denotes the Ludolph number. Then

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln u(\theta) = -\frac{3}{2\theta} - \frac{1}{2\beta} + \frac{\xi^2}{2\beta\theta^2} = \frac{\xi^2 - 3\beta\theta - \theta^2}{2\beta\theta^2},$$

so that (3.15) is satisfied with

$$k = 2; \qquad \gamma = 0; \qquad \delta = \infty \tag{3.43}$$

$$\eta(0) = \xi^2; \qquad \eta(1) = -3\beta; \qquad \eta(2) = -1 \tag{3.44}$$

$$\chi(0) = \chi(1) = 0; \qquad \chi(2) = 2\beta. \tag{3.45}$$

Insertion in (3.17) gives

$$\rho(0) = -\xi^2; \qquad \rho(1) = -\beta; \qquad \rho(2) = 2(1 - h(0))\beta + 1,$$

and by letting $h(0) = 0$, we obtain

$$\dot{\rho}(0) = -\xi^2; \qquad \dot{\rho}(1) = -\beta; \qquad \dot{\rho}(2) = 2\beta + 1.$$

Insertion in (3.26) gives

$$p(n) = \frac{1}{2\beta + 1}\left(\beta\left(2 - \frac{3}{n}\right)p(n-1) + \frac{\xi^2}{n(n-1)}p(n-2)\right),$$
$$(n = 3, 4, \dots)$$

and by insertion of (3.42) in (3.2), we obtain

$$p(n) = \frac{\xi}{\sqrt{2\pi\beta}n!}\int_0^\infty \theta^{n-\frac{3}{2}}\mathrm{e}^{-\theta-\frac{(\theta-\xi)^2}{2\beta\theta}}\,\mathrm{d}\theta, \qquad (n = 0, 1, 2, \dots)$$

which gives in particular the initial values $p(0)$, $p(1)$, and $p(2)$.

From (3.20), we obtain

$$v_2(x) = \frac{1}{2(1 - h(0))\beta + 1}\left(2\beta\sum_{y=1}^{x} h(y)v_2(x - y) + \xi^2 f(x) + \beta v_1(x)\right).$$
$$(x = 1, 2, \dots)$$

In Example 4.3, we shall present another procedure for recursive evaluation of f.

3.8.5 Transformed Gamma Distribution

Let U be the transformed Gamma distribution TGamma(α, β, k) with density

$$u(\theta) = \frac{k\beta^\alpha}{\Gamma(\alpha)}\theta^{k\alpha-1}\mathrm{e}^{-\beta\theta^k}. \qquad (\theta > 0; \alpha, \beta, k > 0) \tag{3.46}$$

If Θ has the distribution Gamma(α, β) given by (3.5), then $\Theta^{1/k}$ has the distribution TGamma(α, β, k). From (3.46), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\theta} \ln u(\theta) = \frac{k\alpha - 1}{\theta} - \beta k \theta^{k-1} = \frac{k\alpha - 1 - k\beta\theta^k}{\theta},$$

so that (3.15) is satisfied when k is a positive integer. We have already discussed the case $k = 1$ in Sect. 3.2, so we now assume that k is a positive integer greater than one. In that case, we have

$$\gamma = 0; \qquad \delta = \infty \tag{3.47}$$

$$\eta(0) = k\alpha - 1; \qquad \eta(k) = -k\beta \tag{3.48}$$

$$\chi(1) = 1 \tag{3.49}$$

and $\eta(i)$ and $\chi(i)$ equal to zero for all other values of i. This gives

$$\rho(0) = -k\alpha; \qquad \rho(1) = 1 - h(0); \qquad \rho(k) = k\beta$$

$$\dot{\rho}(0) = -k\alpha; \qquad \dot{\rho}(1) = 1; \qquad \dot{\rho}(k) = k\beta$$

and $\rho(i) = \dot{\rho}(i) = 0$ for all other values of i.

Insertion in (3.26) gives

$$p(n) = \frac{1}{kn^{(k-1)}\beta}\left(\frac{n-k+k\alpha}{n-k+1} p(n-k) - p(n-k+1)\right),$$
$$(n = k+1, k+2, \dots)$$

and by insertion of (3.46) in (3.2) we obtain

$$p(n) = \frac{k\beta^\alpha}{n!\Gamma(\alpha)} \int_0^\infty \theta^{k\alpha+n-1} \mathrm{e}^{-\theta-\beta\theta^k} \mathrm{d}\theta, \qquad (n = 0, 1, 2, \dots)$$

which gives in particular the initial values $p(0), p(1), \dots, p(k)$.

From (3.20), we obtain

$$v_k(x) = \frac{1}{k\beta}\left(\sum_{y=1}^{x} h(y) v_1(x-y) + k\alpha f(x) - (1 - h(0)) v_1(x)\right).$$
$$(x = 1, 2, \dots)$$

Further Remarks and References

Willmot (1993) proved Theorem 3.3 by using generating functions and discussed the invariance properties of the condition (3.15). Hesselager (1996a) proved Theorem 3.2 and described how it can be applied for evaluation of compound mixed

Poisson distributions. He also indicated how Theorem 3.3 can be deduced from Theorem 3.2. Grandell (1997) deduced the recursion (3.29) from Theorem 3.3 and pointed out that it is more efficient than the recursion of Hesselager (1996a) when $k = 1$. Other recursions for compound mixed Poisson distributions have been discussed by Willmot (1986a, 1986b) and Gerhold et al. (2008).

Section 3.4 is based on Sect. 4B in Sundt and Vernic (2004).

The examples in Sect. 3.8 are based on examples in Willmot (1993) and Hesselager (1996a), who also discuss other parametric classes of mixing distributions. Mixed Poisson distributions with inverse Gauss mixing distribution are discussed by Willmot (1989). For more information on the parametric classes of distributions that we have applied for the mixing distributions, see e.g. Johnson et al. (1994, 1995).

For overviews of the theory of mixed Poisson distributions, see Grandell (1997) and Karlis and Xekalaki (2005).

In Sect. 4.3, we shall deduce some other recursions for mixed Poisson distributions and compound mixed Poisson distributions in the case when the mixing distribution is infinitely divisible.

Chapter 4
Infinite Divisibility

Summary

In the present chapter, the emphasis is on infinitely divisible distributions in $\mathcal{P}_{1\underline{0}}$.

In Sect. 4.1, we define infinite divisibility and prove some properties of infinitely divisible distributions.

The kernel of Sect. 4.2 is the result that a distribution in $\mathcal{P}_{1\underline{0}}$ is infinitely divisible iff it can be represented as a compound Poisson distribution with severity distribution in $\mathcal{P}_{11}$. To prove this result, we shall apply the recursion of Theorem 2.8 for convolutions of a distribution and the recursion of Theorem 2.2 for a compound Poisson distribution. We also discuss how to find the Poisson parameter and the severity distribution of this representation of an infinitely divisible distribution.

In Sect. 4.3, we study mixed Poisson distributions with infinitely divisible mixing distribution. Such mixed distributions are themselves infinitely divisible, and, hence, by the characterisation mentioned above, it follows that they can be represented as compound Poisson distributions. We discuss how to find the Poisson parameter and the severity distribution.

The De Pril transform is a tool for recursive evaluation of convolutions of distributions in $\mathcal{P}_{1\underline{0}}$, in insurance applications in particular within individual models. In Chap. 6, this transform will be discussed in full generality within $\mathcal{P}_{1\underline{0}}$. However, already in Sect. 4.4, we shall introduce it within the context of infinitely divisible distributions in $\mathcal{P}_{1\underline{0}}$. In this context, it is easy to motivate the De Pril transform and its properties in relation to compound Poisson distributions. In particular, we shall show that a distribution in $\mathcal{P}_{1\underline{0}}$ is infinitely divisible iff its De Pril transform is non-negative.

4.1 Definitions and Properties

A distribution is said to be *infinitely divisible* if for each positive integer M there exists a distribution of which the original distribution is the M-fold convolution. Correspondingly, we say that a random variable is infinitely divisible if its distribution is infinitely divisible.

We say that a distribution is infinitely divisible within a set S of distributions if for each positive integer M there exists in S a distribution of which the original distribution is the M-fold convolution. For convenience, we shall normally mean infinitely divisible in S when saying that a distribution in S is infinitely divisible.

From Theorem 2.7 and Fig. 2.1, we see that in the Panjer class, the Poisson and negative binomial distributions are infinitely divisible, but the binomial distribution is not.

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_4, © Springer-Verlag Berlin Heidelberg 2009

The following theorem summarises some properties of infinitely divisible distributions. As most of these results hold more generally than for distributions on integers, we shall identify the distributions with their cumulative distribution functions and therefore denote them by capitals.

Theorem 4.1 i) *A distribution concentrated in one point, is infinitely divisible.*

ii) *The convolution of two infinitely divisible distributions is infinitely divisible.*

iii) *A shifted infinitely divisible distribution is infinitely divisible.*

iv) *A scaled infinitely divisible distribution is infinitely divisible.*

v) *A mixed Poisson distribution is infinitely divisible if its mixing distribution is infinitely divisible.*

vi) *A compound distribution is infinitely divisible if its counting distribution is infinitely divisible in distributions in* $\mathcal{P}_{10}$.

vii) *An infinitely divisible distribution in* $\mathcal{P}_{10}$ *has a positive probability in zero.*

Proof i) For any positive integer M, the distribution concentrated in a point x is the M-fold convolution of a distribution concentrated in x/M.

ii) Let $F = F_1 * F_2$ where F_1 and F_2 are infinitely divisible distributions. Then, for each positive integer M, there exist distributions F_{1M} and F_{2M} such that $F_1 = F_{1M}^{M*}$ and $F_2 = F_{2M}^{M*}$. From this follows that

$$F = F_1 * F_2 = F_{1M}^{M*} * F_{2M}^{M*} = (F_{1M} * F_{2M})^{M*},$$

that is, for each positive integer M, $F = F_M^{M*}$ with $F_M = F_{1M} * F_{2M}$. Hence, F is infinitely divisible.

iii) This result follows from i) and ii).

iv) Let $X = aY$ where a is a constant and Y is an infinitely divisible random variable. Then, for each positive integer M, there exist independent and identically distributed random variables $Y_{M1}, Y_{M2}, \ldots, Y_{MM}$ such that $Y = \sum_{i=1}^{M} Y_{Mi}$. Then $X = \sum_{i=1}^{M} X_{Mi}$ with $X_{Mi} = aY_{Mi}$ for $i = 1, 2, \ldots, M$. Hence, for each positive integer M, X can be expressed as the sum of M independent and identically distributed random variables, and X is therefore infinitely divisible.

v) Let F be a mixed Poisson distribution with infinitely divisible mixing distribution U. Then, for each positive integer M, there exists a distribution U_M such that $U = U_M^{M*}$. Let F_M be the mixed Poisson distribution with mixing distribution U_M. From Theorem 3.1 follows that $F_M^{M*} = F$. Thus, for each positive integer M, there exists a distribution of which F is the M-fold convolution, and F is therefore infinitely divisible.

vi) Let $F = p \vee H$ be a compound distribution with severity distribution H and infinitely divisible counting distribution $p \in \mathcal{P}_{10}$, which is infinitely divisible in distributions in $\mathcal{P}_{10}$. Then, for each positive integer M, there exists a distribution $p_M \in \mathcal{P}_{10}$ such that $p = p_M^{M*}$. Thus,

$$F = p \vee H = p_M^{M*} \vee H = (p_M \vee H)^{M*},$$

that is, for each positive integer M, $F = F_M^{M*}$ with $F_M = p_M \vee H$. Hence, F is infinitely divisible.

vii) If $f \in \mathcal{P}_{10}$ is infinitely divisible, then, for each positive integer M, there exists a distribution $f_M \in \mathcal{P}_{10}$ such that $f = f_M^{M*}$. In particular, this gives

$$f(0) = f_M^{M*}(0) = f_M(0)^M,$$

so that $f_M(0) = f(0)^{1/M}$. Thus, if $f(0) = 0$, then we must have $f_M(0) = 0$ for all M. However, if $f_M(0) = 0$, then $f(x) = f_M^{M*}(x) = 0$ for all $x < M$, and, as this should hold for all positive integers M, we must have $f(x) = 0$ for all non-negative integers x, which is impossible for a distribution in $\mathcal{P}_{10}$. Hence, f must have a positive probability in zero.

This completes the proof of Theorem 4.1. □

4.2 Characterisation

We are now ready to prove the following characterisation of infinite divisible distributions in $\mathcal{P}_{10}$.

Theorem 4.2 *A non-degenerate distribution in $\mathcal{P}_{\underline{1}0}$ is infinitely divisible iff it can be expressed as a compound Poisson distribution with severity distribution in $\mathcal{P}_{11}$.*

Proof As Poisson distributions and compound distributions with infinitely divisible counting distribution are infinitely divisible, we have that all compound Poisson distributions with severity distribution in $\mathcal{P}_{11}$ are infinitely divisible.

Let us now assume that $f \in \mathcal{P}_{\underline{1}0}$ is non-degenerate and infinitely divisible. Then, for each positive integer M, there exists a distribution $f_M \in \mathcal{P}_{\underline{1}0}$ such that $f = f_M^{M*}$. From Theorem 2.8, we obtain that for $x = 1, 2, \dots,$

$$f(x) = \frac{1}{f_M(0)} \sum_{y=1}^{x} \left((M+1)\frac{y}{x} - 1 \right) f_M(y) f(x-y),$$

that is,

$$f(x) = \frac{1}{f(0)^{1/M}} \sum_{y=1}^{x} \left(\left(\frac{1}{M} + 1 \right) \frac{y}{x} - \frac{1}{M} \right) g_M(y) f(x-y)$$

with $g_M = M f_M$. Solving for $g_M(x)$ gives

$$g_M(x) = \frac{1}{f(0)} \left(f(0)^{1/M} f(x) - \sum_{y=1}^{x-1} \left(\left(\frac{1}{M} + 1 \right) \frac{y}{x} - \frac{1}{M} \right) g_M(y) f(x-y) \right). \tag{4.1}$$

We want to show by induction that $g(x) = \lim_{M\uparrow\infty} g_M(x)$ exists and is finite for all positive integers x. From (4.1), we obtain that

$$\lim_{M\uparrow\infty} g_M(1) = \lim_{M\uparrow\infty} \frac{1}{f(0)} f(0)^{1/M} f(1) = \frac{f(1)}{f(0)}.$$

Thus, the hypothesis holds for $x = 1$. Let us now assume that it holds for $x = 1, 2, \ldots, z-1$. We want to show that it also holds for $x = z$. Letting M go to infinity in (4.1) with $x = z$ gives

$$g(z) = \frac{1}{f(0)}\left(f(z) - \frac{1}{z}\sum_{y=1}^{z-1} yg(y)f(z-y)\right), \tag{4.2}$$

that is, the hypothesis holds also for $x = z$. Hence, it holds for all non-negative integers x.

Solving (4.2) for $f(z)$ gives

$$f(z) = \frac{1}{z}\sum_{y=1}^{z} yg(y)f(z-y).$$

Letting $h = g/\lambda$ and

$$\lambda = -\ln f(0), \tag{4.3}$$

we see that f satisfies the recursion of Theorem 2.2 for a compound Poisson distribution with Poisson parameter λ and severity distribution $h \in \mathcal{P}_{11}$. We know that h is non-negative. Hence, it is a distribution if it sums to one. If it had not summed to one, then f would not have summed to one either. However, as f is a distribution, it does sum to one, and then h must also sum to one.

We have now shown that a non-degenerate infinitely divisible distribution in $\mathcal{P}_{1\underline{0}}$ can be expressed as a compound Poisson distribution with severity distribution in $\mathcal{P}_{11}$.

This completes the proof of Theorem 4.2. □

By interpreting a distribution concentrated in zero as a Poisson distribution with parameter zero, we can drop the assumption in Theorem 4.2 that the distribution should be non-degenerate.

If a non-generate distribution $f \in \mathcal{P}_{1\underline{0}}$ is infinitely divisible, then Theorem 4.2 gives that it can be expressed as a compound Poisson distribution with severity distribution in $\mathcal{P}_{11}$, but how do we find the Poisson parameter λ and the severity distribution h? We have already seen that λ is given by (4.3). If we know f, then we can evaluate h recursively by solving the recursion (2.7) for $h(x)$, that is,

$$h(x) = \frac{1}{f(0)}\left(\frac{f(x)}{\lambda} - \frac{1}{x}\sum_{y=1}^{x-1} yh(y)f(x-y)\right). \qquad (x = 1, 2, \ldots) \tag{4.4}$$

Alternatively, we can use generating functions and solve (2.13) for $\tau_h(s)$, that is,

$$\tau_h(s) = 1 + \frac{\ln \tau_f(s)}{\lambda} = 1 - \frac{\ln \tau_f(s)}{\ln f(0)} = \frac{\ln \frac{\tau_f(0)}{\tau_f(s)}}{\ln \tau_f(0)}. \tag{4.5}$$

Example 4.1 We have already pointed out that the negative binomial distribution $\mathrm{NB}(\alpha, \pi)$ is infinitely divisible. By insertion of (2.42) in (4.3), we obtain

$$\lambda = -\alpha \ln(1-\pi), \tag{4.6}$$

as well as

$$\tau_f(s) = \sum_{x=0}^{\infty} s^x \binom{\alpha + x - 1}{x} \pi^x (1-\pi)^\alpha = (1-\pi)^\alpha \sum_{x=0}^{\infty} \binom{\alpha + x - 1}{x} (\pi s)^x,$$

that is,

$$\tau_f(s) = \left(\frac{1-\pi}{1-\pi s}\right)^\alpha. \tag{4.7}$$

Insertion in the last expression in (4.5) gives

$$\tau_h(s) = \frac{\ln(1-\pi s)^\alpha}{\ln(1-\pi)^\alpha} = \frac{\ln(1-\pi s)}{\ln(1-\pi)} = \frac{1}{-\ln(1-\pi)} \ln \frac{1}{1-\pi s}$$

$$= \frac{1}{-\ln(1-\pi)} \sum_{x=1}^{\infty} \frac{\pi^x}{x} s^x,$$

from which follows that h is the logarithmic distribution $\mathrm{Log}(\pi)$. Hence, we have shown that a negative binomial distribution can be expressed as a compound Poisson distribution with logarithmic severity distribution. In Example 2.1, we showed that in another way. □

4.3 Mixed Poisson Distributions

4.3.1 Infinitely Divisible Mixing Distribution

By combining Theorems 4.1v) and 4.2, we obtain the following corollary.

Corollary 4.1 *A mixed Poisson distribution with infinitely divisible mixing distribution can be expressed as a compound Poisson distribution with severity distribution in* $\mathcal{P}_{11}$.

If we know that a mixed Poisson distribution f has an infinitely divisible mixing distribution U, then Corollary 4.1 gives that it can be expressed as a compound

Poisson distribution with severity distribution in $\mathcal{P}_{11}$, but how do we find the Poisson parameter λ and the severity distribution h? This is closely related to what we did above for Theorem 4.2.

Application of (4.3), (1.28), and (3.3) gives

$$\lambda = -\ln f(0) = -\ln \tau_f(0) = -\ln \gamma_U(1), \tag{4.8}$$

and by insertion of (3.3) and (4.8) in (4.5), we obtain

$$\tau_h(s) = 1 + \frac{\ln \gamma_U(1-s)}{\lambda} = 1 - \frac{\ln \gamma_U(1-s)}{\ln \gamma_U(1)} = \frac{\ln \frac{\gamma_U(1)}{\gamma_U(1-s)}}{\ln \gamma_U(1)}. \tag{4.9}$$

We can also evaluate h recursively by (4.4) if f is known.

Example 4.2 From Sect. 3.2, we know that a mixed Poisson distribution with Gamma mixing distribution Gamma(α, β) is the negative binomial distribution NB$(\alpha, (\beta+1)^{-1})$ given by (3.6). From Example 4.1, we obtain that $\lambda = -\alpha \ln(\beta/(\beta+1))$ and h is the logarithmic distribution Log$((\beta+1)^{-1})$. □

Example 4.3 Let U be the inverse Gauss distribution IGauss(ξ, β) with density u given by (3.42). Then

$$\begin{aligned}
\gamma_U(s) &= \frac{\xi}{\sqrt{2\pi\beta}} \int_0^\infty \mathrm{e}^{-s\theta} \theta^{-\frac{3}{2}} \mathrm{e}^{-\frac{(\theta-\xi)^2}{2\beta\theta}} \,\mathrm{d}\theta \\
&= \frac{\xi}{\sqrt{2\pi\beta}} \int_0^\infty \theta^{-\frac{3}{2}} \exp\left(-s\theta - \frac{\theta^2 - 2\xi\theta + \xi^2}{2\beta\theta}\right) \mathrm{d}\theta \\
&= \frac{\xi}{\sqrt{2\pi\beta}} \int_0^\infty \theta^{-\frac{3}{2}} \exp\left(-\frac{2\beta s+1}{2\beta\theta}\left(\theta^2 - \frac{2\xi\theta}{2\beta s+1} + \frac{\xi^2}{2\beta s+1}\right)\right) \mathrm{d}\theta \\
&= \frac{\xi}{\sqrt{2\pi\beta}} \int_0^\infty \theta^{-\frac{3}{2}} \exp\left(-\frac{2\beta s+1}{2\beta\theta}\left(\left(\theta - \frac{\xi}{\sqrt{2\beta s+1}}\right)^2\right.\right. \\
&\qquad \left.\left. - \frac{2\xi\theta}{2\beta s+1} + \frac{2\xi\theta}{\sqrt{2\beta s+1}}\right)\right) \mathrm{d}\theta \\
&= \frac{\xi}{\sqrt{2\pi\beta}} \exp\left(\frac{\xi}{\beta}\left(1 - \sqrt{2\beta s+1}\right)\right) \\
&\qquad \times \int_0^\infty \theta^{-\frac{3}{2}} \exp\left(-\frac{2\beta s+1}{2\beta\theta}\left(\theta - \frac{\xi}{\sqrt{2\beta s+1}}\right)^2\right) \mathrm{d}\theta \\
&= \frac{\xi}{\sqrt{2\pi\beta}} \exp\left(\frac{\xi}{\beta}\left(1 - \sqrt{2\beta s+1}\right)\right) \frac{\sqrt{2\beta s+1}}{\xi} \sqrt{2\pi \frac{\beta}{2\beta s+1}} \\
&= \exp\left(-\frac{\xi}{\beta}\left(\sqrt{2\beta s+1} - 1\right)\right).
\end{aligned}$$

From this, we see that for any positive integer M, IGauss(ξ,β) is the M-fold convolution of IGauss$(\xi/M,\beta)$. Hence, U is infinitely divisible. Insertion in (4.8) and (4.9) gives

$$\lambda = \frac{\xi}{\beta}\big(\sqrt{2\beta+1}-1\big)$$

and

$$\begin{aligned}\tau_h(s) &= 1-\frac{\sqrt{2\beta(1-s)+1}-1}{\sqrt{2\beta+1}-1} = \frac{\sqrt{2\beta+1-2\beta s}-\sqrt{2\beta+1}}{1-\sqrt{2\beta+1}}\\ &= \frac{1}{(2\beta+1)^{-1/2}-1}\left(\left(1-\frac{2\beta}{2\beta+1}s\right)^{1/2}-1\right)\\ &= \frac{1}{(2\beta+1)^{-1/2}-1}\sum_{x=1}^{\infty}\binom{-\frac{1}{2}+x-1}{x}\left(\frac{2\beta}{2\beta+1}\right)^{x}s^{x}.\end{aligned}$$

By comparison with (2.66), we see that h is the extended truncated negative binomial distribution ETNB$(-1/2, 2\beta/(2\beta+1))$. If we want to evaluate the compound distribution $f\vee k$ for some $k\in\mathcal{P}_{11}$, then, as an alternative to the procedure of Sect. 3.8.4, we can first evaluate $h\vee k$ by the recursion (2.68) and then

$$f\vee k = (p\vee h)\vee k = p\vee(h\vee k)$$

by the recursion (2.7) with p being the Poisson distribution Po(λ). □

4.3.2 Mixing Distribution in $\mathcal{P}_{10}$

If the mixing distribution is in $\mathcal{P}_{10}$, then we can sometimes apply the following theorem for recursive evaluation of f.

Theorem 4.3 *A mixed Poisson distribution with mixing distribution in $\mathcal{P}_{10}$ can be expressed as a compound distribution with the mixing distribution as counting distribution and severity distribution* Po(1).

Proof Let f be a mixed Poisson distribution with mixing distribution $u\in\mathcal{P}_{10}$. Then $f=\sum_{\theta=1}^{\infty}u(\theta)h_\theta$ with h_θ denoting the Poisson distribution Po(θ). When θ is a non-negative integer, Po(θ) is the θ-fold convolution of Po(1) so that $h_\theta = h_1^{\theta *}$. Hence, by (1.6),

$$f=\sum_{\theta=1}^{\infty}u(\theta)h_\theta=\sum_{\theta=1}^{\infty}u(\theta)h_1^{\theta *}=u\vee h_1,$$

that is, a compound distribution with the mixing distribution as counting distribution and severity distribution Po(1). □

In particular, by combining Theorem 4.3 with (2.78), we obtain that if f is a mixed Poisson distribution with infinitely divisible mixing distribution $u \in \mathcal{P}_{10}$ that satisfies the recursion

$$u(\theta) = \left(a + \frac{b}{\theta}\right) u(\theta - 1) \qquad (\theta = l+1, l+2, \dots)$$

for some non-negative integer l and constants a and b, then f can be evaluated recursively by

$$f(x) = \frac{1}{\mathrm{e} - a}\left(\mathrm{e}^{1-\theta} \sum_{\theta=1}^{l}\left(u(\theta) - \left(a + \frac{b}{\theta}\right) u(\theta - 1)\right)\frac{\theta^x}{x!} + \sum_{y=1}^{x}\left(a + b\frac{y}{x}\right)\frac{f(x-y)}{y!}\right) \qquad (x = 1, 2, \dots) \tag{4.10}$$

with initial value

$$f(0) = \sum_{\theta=1}^{\infty} \mathrm{e}^{-\theta} u(\theta) = \gamma_u(1).$$

When $l = 0$, that is, u is in the Panjer class, then (4.10) reduces to

$$f(x) = \frac{1}{\mathrm{e} - a}\sum_{y=1}^{x}\left(a + b\frac{y}{x}\right)\frac{f(x-y)}{y!}. \qquad (x = 1, 2, \dots)$$

Other classes of recursions for u can be handled analogously.

4.4 De Pril Transforms of Infinitely Divisible Distributions in $\mathcal{P}_{1\underline{0}}$

4.4.1 Definition and Properties

We shall need the following theorem.

Theorem 4.4 *For $j = 1, 2, \dots, M$, let F_j be a compound Poisson distribution with Poisson parameter λ_j and severity distribution H_j. Then $F = *_{j=1}^{M} F_j$ is a compound Poisson distribution with Poisson parameter*

$$\lambda = \lambda_{\bullet M} \tag{4.11}$$

and severity distribution

$$H = \frac{1}{\lambda}\sum_{j=1}^{M} \lambda_j H_j.$$

Proof We have

$$\zeta_H(s) = \int_{(-\infty,\infty)} \mathrm{e}^{\mathrm{i}sx}\,\mathrm{d}H(x) = \frac{1}{\lambda}\sum_{j=1}^{M}\lambda_j \int_{(-\infty,\infty)} \mathrm{e}^{\mathrm{i}sx}\,\mathrm{d}H_j(x) = \frac{1}{\lambda}\sum_{j=1}^{M}\lambda_j\zeta_{Hj}(s),$$

and by using (1.18), (1.29), and (2.12), we obtain

$$\begin{aligned}\zeta_F(s) &= \prod_{j=1}^{M}\zeta_{F_j}(s) = \prod_{j=1}^{M}\exp(\lambda_j(\zeta_{H_j}(s)-1)) = \exp\left(\sum_{j=1}^{M}\lambda_j(\zeta_{H_j}(s)-1)\right)\\ &= \exp\left(\lambda\left(\frac{1}{\lambda}\sum_{j=1}^{M}\lambda_j\zeta_{H_j}(s)-1\right)\right) = \exp(\lambda(\zeta_H(s)-1)).\end{aligned}$$

This is the characteristic function of a compound Poisson distribution with Poisson parameter λ and severity distribution H, and, hence, F is such a compound distribution. □

Let $f_1, f_2, \dots, f_M \in \mathcal{P}_{1\underline{0}}$ be infinitely divisible. Then Theorem 4.2 gives that for $j = 1, 2, \dots, M$, f_j can be expressed as a compound Poisson distribution with severity distribution in $\mathcal{P}_{11}$; let λ_j denote the Poisson parameter and h_j the severity distribution. Theorem 4.4 gives that $f = *_{j=1}^{M} f_j$ is a compound Poisson distribution with Poisson parameter λ given by (4.11) and severity distribution

$$h = \frac{1}{\lambda}\sum_{j=1}^{M}\lambda_j h_j. \tag{4.12}$$

Thus, if we know the f_js, then we can evaluate the convolution f by for each j first finding λ_j by (4.3) and then h_j recursively by (4.4). After that, we find λ by (4.11) and h by (4.12) and evaluate f recursively by Theorem 2.2.

If the purpose of expressing the f_js as compound Poisson distributions is to evaluate their convolution, is it then necessary to evaluate the λ_js and h_js, or can we do some shortcuts? We see that for each j, λ_j and h_j appear in (2.7) only through the function

$$\varphi_{f_j} = \lambda_j \Phi h_j, \tag{4.13}$$

so why not rather work with that? From (4.4), we obtain that φ_{f_j} can be evaluated recursively by

$$\varphi_{f_j}(x) = \frac{1}{f_j(0)}\left(x f_j(x) - \sum_{y=1}^{x-1}\varphi_{f_j}(y) f_j(x-y)\right). \quad (x = 1, 2, \dots) \tag{4.14}$$

We call the function φ_{f_j} the *De Pril transform* of f_j and use the recursion (4.14) as definition of the De Pril transform of any infinitely divisible distribution $f_j \in \mathcal{P}_{1\underline{0}}$.

Solving (4.14) for $f_j(x)$ gives the recursion

$$f_j(x) = \frac{1}{x}\sum_{y=1}^{x} \varphi_{f_j}(y) f_j(x-y), \qquad (x = 1, 2, \dots) \tag{4.15}$$

for an infinitely divisible distribution in $\mathcal{P}_{1\underline{0}}$ in terms of its De Pril transform.

In terms of De Pril transforms, we can rewrite (4.12) as

$$\varphi_f = \sum_{j=1}^{M} \varphi_{f_j}. \tag{4.16}$$

We can now evaluate f by evaluating each φ_{f_j} recursively by (4.14), finding φ_f by (4.16), and finally evaluating f recursively by (4.15) with initial value $f(0) = \prod_{j=1}^{M} f_j(0)$. This is *De Pril's first method* for evaluating the convolution of M distributions.

As a Poisson distribution is a compound Poisson distribution with severity distribution concentrated in one, (4.13) gives that the De Pril transform of a Poisson distribution f with parameter λ is given by

$$\varphi_f(y) = \lambda I(y = 1). \qquad (y = 1, 2, \dots) \tag{4.17}$$

Now let f be an infinitely divisible distribution in $\mathcal{P}_{1\underline{0}}$ and $k \in \mathcal{P}_{11}$. Then there exists a $\lambda > 0$ and a distribution $h \in \mathcal{P}_{11}$ such that f is a compound Poisson distribution with Poisson parameter λ and severity distribution h. Hence, $f \vee k$ is a compound Poisson distribution with Poisson parameter λ and severity distribution $h \vee k$. Application of (4.13) and (1.8) gives that for $x = 1, 2, \dots,$

$$\varphi_{f\vee k}(x) = \lambda x (h \vee k)(x) = \lambda x \sum_{y=1}^{x} h(y) k^{y*}(x),$$

and by one more application of (4.13), we obtain

$$\varphi_{f\vee k}(x) = x \sum_{y=1}^{x} \frac{\varphi_f(y)}{y} k^{y*}(x), \qquad (x = 1, 2, \dots) \tag{4.18}$$

that is,

$$\varphi_{f\vee k} = \Phi(\Psi\varphi_f \vee k). \tag{4.19}$$

Now let f be the negative binomial distribution $\mathrm{NB}(\alpha, \pi)$. In Examples 2.1 and 4.1, we showed that this distribution is a compound Poisson distribution with Poisson parameter $-\alpha \ln(1-\pi)$ and logarithmic $\mathrm{Log}(\pi)$ severity distribution. Insertion of that Poisson parameter and (2.19) in (4.13) gives that

$$\varphi_f(y) = \alpha \pi^y. \qquad (y = 1, 2, \dots) \tag{4.20}$$

By insertion in (4.18), we obtain that if $k \in \mathcal{P}_{11}$, then

$$\varphi_{f\vee k}(x) = \alpha x \sum_{y=1}^{x} \frac{\pi^y}{y} k^{y*}(x). \qquad (x = 1, 2, \dots) \tag{4.21}$$

By introducing $\xi = \lambda\eta$ in (2.17) and using (4.13) and Theorem 2.3, we obtain that if an infinitely divisible distribution $f \in \mathcal{P}_{1\underline{0}}$ satisfies the recursion

$$\varphi_f(x) = \xi(x) + \sum_{z=1}^{r} \chi(z)\varphi_f(x - z) \qquad (x = 1, 2, \dots) \tag{4.22}$$

for functions ξ and χ on $\{1, 2, \dots, r\}$ with r being a positive integer or infinity, then it also satisfies the recursion

$$f(x) = \sum_{y=1}^{r} \left(\frac{\xi(y)}{x} + \left(1 - \frac{y}{x}\right)\chi(y)\right) f(x - y). \qquad (x = 1, 2, \dots) \tag{4.23}$$

Example 4.4 For $j = 0, 1, 2, \dots, M$, let $f_j = p_j \vee h_j$ with $h_j \in \mathcal{P}_{11}$. We let p_j be the Poisson distribution $\mathrm{Po}(\lambda)$ for $j = 0$ and the negative binomial distribution $\mathrm{NB}(\alpha_j, \pi_j)$ for $j = 1, 2, \dots, M$. We want to evaluate $f = *_{j=0}^{M} f_j$ and $p = *_{j=0}^{M} p_j$.

Insertion of (4.13) and (4.21) in (4.16) gives

$$\varphi_f(x) = x \left(\lambda h_0(x) + \sum_{y=1}^{x} \frac{1}{y} \sum_{j=1}^{M} \alpha_j \pi_j^y h_j^{y*}(x) \right), \qquad (x = 1, 2, \dots) \tag{4.24}$$

and by insertion in (4.15), we obtain

$$f(x) = \frac{1}{x} \sum_{y=1}^{x} y \left(\lambda h_0(y) + \sum_{z=1}^{y} \frac{1}{z} \sum_{j=1}^{M} \alpha_j \pi_j^z h_j^{z*}(y) \right) f(x - y) \qquad (x = 1, 2, \dots) \tag{4.25}$$

with initial value

$$f(0) = \prod_{j=0}^{M} p_j(0) = \mathrm{e}^{-\lambda} \prod_{j=1}^{M} (1 - \pi_j)^{\alpha_j}.$$

The expression (4.24) looks rather awkward, so let us have a look at how it can be evaluated.

For $j = 1, 2, \dots, M$, Example 4.1 gives that $p_j = q_j \vee k_j$ with q_j being the Poisson distribution $\mathrm{Po}(-\alpha_j \ln(1 - \pi_j))$ and k_j being the logarithmic distribution $\mathrm{Log}(\pi_j)$. Then $f_j = q_j \vee (k_j \vee h_j)$, and (4.13) gives that

$$\varphi_{f_j}(x) = -\alpha_j(\ln(1 - \pi_j))x(k_j \vee h_j)(x). \qquad (x = 1, 2, \dots) \tag{4.26}$$

The compound distribution $k_j \vee h_j$ can be evaluated recursively by (2.65). Combining this with (4.26) gives that for $x = 1, 2, \ldots$

$$\varphi_{f_j}(x) = -\alpha_j(\ln(1-\pi_j)) \times x\pi_j\left(\frac{h_j(x)}{-\ln(1-\pi_j)} + \sum_{y=1}^{x-1}\left(1-\frac{y}{x}\right)h_j(y)(k_j \vee h_j)(x-y)\right)$$
$$= \pi_j\left(\alpha_j x h_j(x) - \sum_{y=1}^{x-1} h_j(y)\alpha_j(\ln(1-\pi_j))(x-y)(k_j \vee h_j)(x-y)\right),$$

and by one more application of (4.26), we obtain the recursion

$$\varphi_{f_j}(x) = \pi_j\left(\alpha_j x h_j(x) + \sum_{y=1}^{x-1} h_j(y)\varphi_{f_j}(x-y)\right). \quad (x = 1, 2, \ldots) \tag{4.27}$$

From (4.13), we obtain that $\varphi_{f_0} = \lambda\Phi h_0$.

We have now explained how to evaluate φ_{f_j} for $j = 0, 1, 2, \ldots, M$. Then we can evaluate φ_f by (4.16), and finally we evaluate f by (4.15).

By letting h_j be concentrated in one for $j = 0, 1, 2, \ldots, M$, we obtain that $f_j = p_j$ for $j = 0, 1, 2, \ldots, M$ and $f = p$. Then (4.24) and (4.25) reduce to

$$\varphi_p(x) = \lambda I(x = 1) + \sum_{j=1}^{M} \alpha_j \pi_j^x \qquad (x = 1, 2, \ldots)$$

$$p(x) = \frac{1}{x}\left(\lambda p(x-1) + \sum_{y=1}^{x} p(x-y)\sum_{j=1}^{M}\alpha_j\pi_j^y\right). \qquad (x = 1, 2, \ldots)$$

□

Let us now assume that we have an insurance portfolio with independent policies of M different types. Policies of type j have infinitely divisible aggregate claims distribution $f_j \in \mathcal{P}_{1\underline{0}}$, and there are M_j such policies in the portfolio. The aggregate claims distribution of the portfolio is then

$$f = *_{j=1}^{M} f_j^{M_j *},$$

and its De Pril transform is

$$\varphi_f = \sum_{j=1}^{M} M_j \varphi_{f_j}.$$

Hence, if we have evaluated the φ_{f_j}s, then it is easy to change the M_js, that is, study the effect of changes in portfolio composition on the aggregate claims distribution of the portfolio.

Even when a distribution $f \in \mathcal{P}_{1\underline{0}}$ is not infinitely divisible, we can still define its De Pril transform by (4.14), and then (4.15) holds trivially. In Chap. 6, we shall show that (4.16) and (4.18) still hold. Hence, the De Pril transform is a general tool for evaluation of convolutions of distributions in $\mathcal{P}_{1\underline{0}}$ and compound distributions with counting distribution in $\mathcal{P}_{1\underline{0}}$ and severity distribution in $\mathcal{P}_{11}$.

4.4.2 Characterisation of Infinitely Divisible Distributions in Terms of De Pril Transforms

From the way we defined the De Pril transform of an infinitely divisible distribution in $\mathcal{P}_{1\underline{0}}$, it is clear that such distributions have non-negative De Pril transforms. From the following theorem, we see that this property holds only for such distributions. Hence, it can be applied for characterising them.

Theorem 4.5 *A distribution in $\mathcal{P}_{1\underline{0}}$ is infinitely divisible iff its De Pril transform is non-negative.*

To prove this theorem, we shall need the following lemma.

Lemma 4.1 *If a distribution $f \in \mathcal{P}_{1\underline{0}}$ has non-negative De Pril transform, then $\mu_{\varphi_f}(-1) < \infty$.*

Proof Application of (4.14) gives

$$\mu_{\varphi_f}(-1) = \sum_{x=1}^{\infty} \frac{\varphi_f(x)}{x} = \sum_{x=1}^{\infty} \frac{1}{f(0)} \left(f(x) - \frac{1}{x} \sum_{y=1}^{x-1} \varphi_f(y) f(x-y) \right)$$

$$\leq \sum_{x=1}^{\infty} \frac{f(x)}{f(0)} = \frac{1-f(0)}{f(0)} < \infty.$$

□

Proof of Theorem 4.5 We have already shown that infinitely divisible distributions in $\mathcal{P}_{1\underline{0}}$ have non-negative De Pril transforms.

Now, let $f \in \mathcal{P}_{1\underline{0}}$ with non-negative De Pril transform. From Lemma 4.1, we get that

$$\lambda = \mu_{\varphi_f}(-1) < \infty. \tag{4.28}$$

Let $h(x) = \varphi_f(x)/(\lambda x)$ for $x = 1, 2, \ldots$. Then h is non-negative and sums to one and is therefore a distribution in $\mathcal{P}_{11}$. Insertion of $\varphi_f(x) = \lambda x h(x)$ in (4.15) gives the recursion (2.7) for $p \vee h$ where p denotes the Poisson distribution Po(λ). Hence, f and $p \vee h$ must be proportional, and, as both of them are distributions and therefore sum to one, they must be equal. Thus, any distribution in $\mathcal{P}_{1\underline{0}}$ with non-negative De Pril transform can be expressed as a compound Poisson distribution with severity distribution in $\mathcal{P}_{11}$, and, by Theorem 4.2, it is therefore infinitely divisible.

This completes the proof of Theorem 4.5. □

From (4.28), we have that

$$f(0) = p(0) = \mathrm{e}^{-\lambda} = \mathrm{e}^{-\mu_{\varphi_f}(-1)}.$$

Hence,

$$\mu_{\varphi_f}(-1) = -\ln f(0). \tag{4.29}$$

Furthermore,

$$\mu_{\varphi_f}(0) = \sum_{x=1}^{\infty} \varphi_f(x) = \sum_{x=1}^{\infty} \lambda x h(x) = \lambda \mu_h(1) = \mu_p(1)\mu_h(1) = \mu_{p \vee h}(1),$$

that is,

$$\mu_{\varphi_f}(0) = \mu_f(1). \tag{4.30}$$

In Sect. 6.1, we shall show that (4.29) and (4.30) hold for a wider class of distributions in $\mathcal{P}_{1\underline{0}}$.

Further Remarks and References

In Theorem 4.1, we showed that a mixed Poisson distribution is infinitely divisible if the mixing distribution is infinitely divisible. Maceda (1948) showed the converse implication, that is, a mixed Poisson distribution is infinitely divisible *iff* the mixing distribution is infinitely divisible, see also Godambe and Patil (1975) and Bühlmann and Buzzi (1971).

Our proof of Theorem 4.2 was given by Ospina and Gerber (1987). Feller (1968, Sect. XII.2) proved it by using generating functions.

If we have a sample of independent observations from an infinitely divisible distribution in $\mathcal{P}_{1\underline{0}}$, then we can use the procedure outlined in connection with (2.81) and (2.82) for estimating the Poisson parameter and severity distribution in its representation as a compound Poisson distribution with severity distribution in $\mathcal{P}_{11}$.

Example 4.3 is based on Willmot (1987).

Kestemont and Paris (1985) deduced a recursion for the Hofmann distribution by expressing this distribution in $\mathcal{P}_{1\underline{0}}$ as a compound Poisson distribution and using Theorem 2.2. This recursion has also been studied by Walhin and Paris (1999). The class of Hofmann distributions contains the class of mixed Poisson distributions with Gamma and inverse Gauss mixing distributions.

The material on De Pril transforms in Sect. 4.4 is to a large extent based on Sundt (1995), who introduced the term De Pril transform. De Pril's first method was introduced by De Pril (1989). Formula (4.15) was given by Chan (1982a, 1982b). Theorem 4.5 was proved by Katti (1967) in a slightly different parameterisation by using probability generating functions; see also Theorem 4.4 in Chap. II of Steutel and van Harn (2004). The present proof and formulation in terms of De Pril transforms was presented in Sundt (1995) after having been suggested to Sundt by Gordon Willmot in a letter of December 3, 1992.

Example 4.4 is based on Willmot and Sundt (1989). Ammeter (1949) suggested to evaluate the convolution of compound negative binomial distributions by expressing it as a compound Poisson distribution, using Theorem 4.4 and the representation of a negative binomial distribution as a compound Poisson distribution with logarithmic severity distribution. A distribution in the form of f plays an important role in the credit risk management framework CreditRisk+. Recursive evaluation of that distribution is discussed in the CreditRisk+ manual (Credit Suisse Financial Products 1997) as well as by Giese (2003) and Haaf et al. (2004). Further developments are presented by Bürgisser et al. (2001), Kurth and Tasche (2003), and Gerhold et al. (2008).

For more information on infinite divisibility, see e.g. Lukacs (1970) and Steutel and van Harn (2004).

Chapter 5
Counting Distributions with Recursion of Higher Order

Summary

Chapter 2 circled around various forms of the recursion

$$p(n) = \left(a + \frac{b}{n}\right) p(n-1)$$

for a distribution $p \in \mathcal{P}_{10}$. In the present chapter, we extend this recursion to

$$p(n) = \sum_{i=1}^{k} \left(a(i) + \frac{b(i)}{n}\right) p(n-i). \tag{5.1}$$

In Sect. 5.1, we study compound distributions where the counting distribution satisfies various forms of this recursion. These recursions require evaluation of convolutions of the severity distribution, and such evaluation is the topic of Sect. 5.2.

Section 5.3 is devoted to study of the properties of the classes of distributions that satisfy the recursion (5.1) for $n = 1, 2, \dots$.

In Sect. 5.4, we consider a recursion that can be applied for compound distributions when the generating function of the counting distribution is rational.

5.1 Compound Distributions

In Chap. 2, we studied classes of distributions in $\mathcal{P}_{10}$ for which there existed a recursion of order one, that is, if p is such a counting distribution, then $p(n)$ should be a function of $p(n-1)$, at least for all n greater than a certain limit. We gave characterisations of some of these classes, and from the recursions for these distributions we deduced recursions for compound distributions with these distributions as counting distribution. We shall now extend our setup to classes of distributions in $\mathcal{P}_{10}$ of order k, that is, $p(n)$ shall be a function of $p(n-1)$, $p(n-2)$, $\dots$, $p(n-k)$, at least for n greater than a certain limit.

We return to the setting of Sect. 2.7.1: Let N be the number of claims occurring in an insurance portfolio within a given period and Y_i the amount of the ith of these claims. We assume that the amounts of the individual claims are independent of the number of claims and mutually independent and identically distributed on the non-negative integers with distribution h. Let p denote the distribution of N and $f = p \vee h$ the distribution of the aggregate claims $X = Y_{\bullet N}$.

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_5, © Springer-Verlag Berlin Heidelberg 2009

Our faithful workhorse in Chap. 2 was the relation

$$\mathrm{E}\left[a+b\frac{Y_1}{x}\middle|Y_{\bullet n}=x\right]=a+\frac{b}{n} \qquad (x=1,2,\dots;n=1,2,\dots) \tag{5.2}$$

for constants a and b. From this, we deduced that

$$\sum_{n=1}^{\infty}\left(a+\frac{b}{n}\right)p(n-1)h^{n*}(x)=\sum_{y=0}^{x}\left(a+b\frac{y}{x}\right)h(y)f(x-y),$$
$$(x=1,2,\dots)$$

from which we obtained

$$f(x)=\frac{1}{1-ah(0)}\left(\sum_{n=1}^{x}\left(p(n)-\left(a+\frac{b}{n}\right)p(n-1)\right)h^{n*}(x)\right.$$
$$\left.+\sum_{y=1}^{x}\left(a+b\frac{y}{x}\right)h(y)f(x-y)\right), \qquad (x=1,2,\dots) \tag{5.3}$$

and the terms in the first summation gracefully vanished when $p(n)=(a+b/n)p(n-1)$ for $n=1,2,\dots,x$.

As this was such a nice experience, it would be great if we could do something similar in our extended setup, but what do we then need? In the first summation in (5.3), we had the difference between $p(n)$ and the product of $p(n-1)$ with something, and we were delighted when that difference was equal to zero. In our extended setting, we want that difference replaced with the difference between $p(n)$ and a function of $p(n-1), p(n-2), \dots, p(n-k)$, but what should that function be? As multiplying $p(n-1)$ by something was so fruitful in Chap. 2, it is tempting to multiply $p(n-i)$ by something for $i=1,2,\dots,k$, and then perhaps take the sum of these products? But how should we choose the something? Well, as we were so successful with $a+b/n$ in Chap. 2, let us see what we can achieve with that. Thus, we need to extend (5.2) from one to i for $i=1,2,\dots,k$. The only place one appears in that equation, is through Y_1. Obviously, just replacing Y_1 with Y_i would not bring anything, but what about $Y_{\bullet i}$? Let us try. By the same reasoning we used for setting up (2.11), we obtain

$$\mathrm{E}\left[a+\frac{b}{i}\frac{Y_{\bullet i}}{x}\middle|Y_{\bullet n}=x\right]=a+\frac{b}{n}; \tag{5.4}$$
$$(x=1,2,\dots;n=i,i+1,i+2,\dots;i=1,2,\dots)$$

the reason that we have to divide b by i, is that the conditional expectation of the sum of i observations is ix/n. This gives that for $x=1,2,\dots,$

$$\sum_{n=1}^{\infty}\left(a+\frac{b}{n}\right)p(n-i)h^{n*}(x)$$
$$=\sum_{n=i}^{\infty}\left(a+\frac{b}{n}\right)p(n-i)h^{n*}(x)=\sum_{n=i}^{\infty}\mathrm{E}\left[a+\frac{b}{i}\frac{Y_{\bullet i}}{x}\middle|Y_{\bullet n}=x\right]p(n-i)h^{n*}(x)$$
$$=\sum_{n=i}^{\infty}\sum_{y=0}^{x}\left(a+\frac{b}{i}\frac{y}{x}\right)\frac{h^{i*}(y)h^{(n-i)*}(x-y)}{h^{n*}(x)}p(n-i)h^{n*}(x)$$
$$=\sum_{y=0}^{x}\left(a+\frac{b}{i}\frac{y}{x}\right)h^{i*}(y)\sum_{n=i}^{\infty}p(n-i)h^{(n-i)*}(x-y),$$

that is,

$$\sum_{n=1}^{\infty}\left(a+\frac{b}{n}\right)p(n-i)h^{n*}(x)=\sum_{y=0}^{x}\left(a+\frac{b}{i}\frac{y}{x}\right)h^{i*}(y)f(x-y). \quad (x=1,2,\dots) \tag{5.5}$$

With a slight abuse of notation, now letting a and b be functions, we have

$$p(n)=q(n)+\sum_{i=1}^{k}\left(a(i)+\frac{b(i)}{n}\right)p(n-i) \qquad (n=1,2,\dots) \tag{5.6}$$

with

$$q(n)=p(n)-\sum_{i=1}^{k}\left(a(i)+\frac{b(i)}{n}\right)p(n-i). \qquad (n=1,2,\dots)$$

Then

$$f(x)=(q\vee h)(x)+\sum_{i=1}^{k}\sum_{n=1}^{\infty}\left(a(i)+\frac{b(i)}{n}\right)p(n-i)h^{n*}(x), \quad (x=1,2,\dots) \tag{5.7}$$

and insertion of (5.5) gives

$$f(x)=(q\vee h)(x)+\sum_{y=0}^{x}f(x-y)\sum_{i=1}^{k}\left(a(i)+\frac{b(i)}{i}\frac{y}{x}\right)h^{i*}(y)$$
$$=(q\vee h)(x)+\sum_{y=0}^{x}\left((a\vee h)(y)+\frac{\Phi(\Psi b\vee h)(y)}{x}\right)f(x-y),$$
$$(x=1,2,\dots)$$

that is,

$$f(x)=\frac{1}{1-\tau_a(h(0))}\left((q\vee h)(x)+\sum_{y=1}^{x}\left((a\vee h)(y)+\frac{\Phi(\Psi b\vee h)(y)}{x}\right)f(x-y)\right)$$

$$=\frac{1}{1-\tau_a(h(0))}\left(\sum_{n=1}^{\infty}\left(p(n)-\sum_{i=1}^{k}\left(a(i)+\frac{b(i)}{n}\right)p(n-i)\right)h^{n*}(x)\right.$$

$$\left.+\sum_{y=1}^{x}\left((a\vee h)(y)+\frac{\Phi(\Psi b\vee h)(y)}{x}\right)f(x-y)\right)$$

$$=\tilde{q}(x)+\sum_{y=1}^{x}\left(\tilde{a}(y)+\frac{\tilde{b}(y)}{x}\right)f(x-y)\qquad(x=1,2,\dots)\tag{5.8}$$

with

$$\tilde{q}(y)=\frac{(q\vee h)(y)}{1-\tau_a(h(0))}\tag{5.9}$$

$$\tilde{a}(y)=\frac{(a\vee h)(y)}{1-\tau_a(h(0))};\qquad \tilde{b}(y)=\frac{\Phi(\Psi b\vee h)(y)}{1-\tau_a(h(0))}\tag{5.10}$$

for $y=1,2,\dots$, that is, the recursion (5.8) for f is in the same form as the recursion (5.6) for p with k, q, a, and b replaced with ∞, $\tilde{q}$, $\tilde{a}$, and $\tilde{b}$ respectively. This implies that many results for distributions represented in the shape (5.6) also hold for compound distributions with counting distribution represented in the shape (5.6).

We have already used (5.8) to deduce the recursion (3.29).

If

$$p(n)=\sum_{i=1}^{k}\left(a(i)+\frac{b(i)}{n}\right)p(n-i)\qquad(n=l+1,l+2,\dots)\tag{5.11}$$

for some positive integer l, then

$$f(x)=\frac{1}{1-\tau_a(h(0))}\left(\sum_{n=1}^{l}\left(p(n)-\sum_{i=1}^{k}\left(a(i)+\frac{b(i)}{n}\right)p(n-i)\right)h^{n*}(x)\right.$$

$$\left.+\sum_{y=1}^{x}\left((a\vee h)(y)+\frac{\Phi(\Psi b\vee h)(y)}{x}\right)f(x-y)\right),\qquad(x=1,2,\dots)$$

and when $l=0$, that is,

$$p(n)=\sum_{i=1}^{k}\left(a(i)+\frac{b(i)}{n}\right)p(n-i),\qquad(n=1,2,\dots)\tag{5.12}$$

this reduces to

$$f(x) = \frac{1}{1-\tau_a(h(0))}\sum_{y=1}^{x}\left((a \vee h)(y) + \frac{\Phi(\Psi b \vee h)(y)}{x}\right) f(x-y)$$
$$= \sum_{y=1}^{x}\left(\tilde{a}(y) + \frac{\tilde{b}(y)}{x}\right) f(x-y). \qquad (x = 1, 2, \dots) \tag{5.13}$$

When further assuming that $h \in \mathcal{P}_{11}$, we obtain

$$f(x) = \sum_{i=1}^{k}\sum_{y=i}^{x}\left(a(i) + \frac{b(i)}{i}\frac{y}{x}\right) h^{i*}(y) f(x-y). \qquad (x = 1, 2, \dots)$$

Before closing this section, let us deduce a more general recursion.

Theorem 5.1 *The compound distribution $f = p \vee h$ with $h, p \in \mathcal{P}_{10}$ and p satisfying the recursion*

$$p(n) = \sum_{i=1}^{k}\left(a(i) + \sum_{j=0}^{\max(l,i-1)} \frac{b(i,j)}{n-j}\right) p(n-i) \qquad (n = l+1, l+2, \dots)$$

satisfies the recursion

$$f(x) = \frac{1}{1-\tau_a(h(0))}\left(\sum_{n=1}^{l}\left(p(n) - \sum_{i=1}^{k} a(i)p(n-i)\right) h^{n*}(x) - \sum_{i=1}^{k}\sum_{j=0}^{\max(l,i-1)}\sum_{n=j+1}^{l}\frac{b(i,j)}{n-j}p(n-i)h^{n*}(x) + \sum_{i=1}^{k}\sum_{y=1}^{x} a(i)h^{i*}(y)f(x-y) + \sum_{i=1}^{k}\sum_{j=0}^{\max(l,i-1)} b(i,j)\left(c(i-j)h^{j*}(x) + \frac{1}{i-j}\sum_{y=0}^{x-1} h^{j*}(y)\sum_{z=1}^{x-y}\frac{z}{x-y}h^{(i-j)*}(z)f(x-y-z)\right)\right) \tag{5.14}$$

$(x = 1, 2, \dots)$

with

$$c(i) = \sum_{n=1}^{\infty}\frac{1}{n}p(n-i)h(0)^n. \qquad (i = 1, 2, \dots)$$

Proof Application of (5.5) gives that for $x, i = 1, 2, \ldots$ and $j = 0, 1, 2, \ldots, i-1$, we have

$$\sum_{n=j+1}^{\infty} \frac{1}{n-j} p(n-i) h^{n*}(x)$$
$$= \sum_{n=1}^{\infty} \frac{1}{n} p(n-i+j) h^{(n+j)*}(x) = \sum_{y=0}^{x} h^{j*}(y) \sum_{n=1}^{\infty} \frac{1}{n} p(n-i+j) h^{n*}(x-y)$$
$$= c(i-j) h^{j*}(x) + \sum_{y=0}^{x-1} h^{j*}(y) \sum_{z=1}^{x-y} \frac{1}{i-j} \frac{z}{x-y} h^{(i-j)*}(z) f(x-y-z).$$

Together with (5.5), this gives that for $x = 1, 2, \ldots,$

$$f(x) = \sum_{n=1}^{\infty} p(n) h^{n*}(x)$$
$$= \sum_{n=1}^{l} p(n) h^{n*}(x) + \sum_{n=l+1}^{\infty} \sum_{i=1}^{k} \left(a(i) + \sum_{j=0}^{\max(l,i-1)} \frac{b(i,j)}{n-j} \right) p(n-i) h^{n*}(x)$$
$$= \sum_{n=1}^{l} \left(p(n) - \sum_{i=1}^{k} a(i) p(n-i) \right) h^{n*}(x)$$
$$- \sum_{i=1}^{k} \sum_{j=0}^{\max(l,i-1)} \sum_{n=j+1}^{l} \frac{b(i,j)}{n-j} p(n-i) h^{n*}(x)$$
$$+ \sum_{i=1}^{k} \sum_{n=1}^{\infty} a(i) p(n-i) h^{n*}(x) + \sum_{i=1}^{k} \sum_{j=0}^{\max(l,i-1)} \sum_{n=j+1}^{\infty} \frac{b(i,j)}{n-j} p(n-i) h^{n*}(x)$$
$$= \sum_{n=1}^{l} \left(p(n) - \sum_{i=1}^{k} a(i) p(n-i) \right) h^{n*}(x)$$
$$- \sum_{i=1}^{k} \sum_{j=0}^{\max(l,i-1)} \sum_{n=j+1}^{l} \frac{b(i,j)}{n-j} p(n-i) h^{n*}(x) + \sum_{i=1}^{k} \sum_{y=0}^{x} a(i) h^{i*}(y) f(x-y)$$
$$+ \sum_{i=1}^{k} \sum_{j=0}^{\max(l,i-1)} b(i,j) \Bigg(c(i-j) h^{j*}(x)$$
$$+ \frac{1}{i-j} \sum_{y=0}^{x-1} h^{j*}(y) \sum_{z=1}^{x-y} \frac{z}{x-y} h^{(i-j)*}(z) f(x-y-z) \Bigg).$$

Solving for $f(x)$ gives (5.14). □

Unfortunately, evaluation of the function c involves a summation over an infinite range. However, this summation vanishes when $h \in \mathcal{P}_{11}$. In that case, (5.14) reduces to

$$\begin{aligned} f(x) &= \sum_{n=1}^{l} \left(p(n) - \sum_{i=1}^{k} a(i) p(n-i) \right) h^{n*}(x) \\ &\quad - \sum_{i=1}^{k} \sum_{j=0}^{\max(l,i-1)} \sum_{n=j+1}^{l} \frac{b(i,j)}{n-j} p(n-i) h^{n*}(x) + \sum_{i=1}^{k} \sum_{y=i}^{x} a(i) h^{i*}(y) f(x-y) \\ &\quad + \sum_{i=1}^{k} \sum_{j=0}^{\max(l,i-1)} \frac{b(i,j)}{i-j} \sum_{y=j}^{x-1} h^{j*}(y) \sum_{z=i-j}^{x-y} \frac{z}{x-y} h^{(i-j)*}(z) f(x-y-z). \\ &\quad (x = 1, 2, \dots) \end{aligned}$$

5.2 Convolutions of the Severity Distribution

We see that in recursions like (5.13) we might need to evaluate h^{i*} for $i = 1, 2, \dots, k$. For these evaluations, the recursions of Sect. 2.4 do not seem to be efficient unless many of the h^{i*}s vanish because $a(i) = b(i) = 0$. Normally, it would be more efficient to use that the convolution of two distributions $h_1, h_2 \in \mathcal{P}_{10}$ can be evaluated by

$$(h_1 * h_2)(y) = \sum_{z=0}^{y} h_1(z) h_2(y-z). \qquad (y = 0, 1, 2, \dots)$$

At first glance, it might then seem natural to evaluate h^{i*} by applying this procedure with $h_1 = h$ and $h_2 = h^{(i-1)*}$. However, when i is even, it seems more efficient to use $h_1 = h_2 = h^{i/2*}$ as then

$$h^{i*}(y) = 2 \sum_{z=0}^{[(y-1)/2]} h^{i/2*}(z) h^{i/2*}(y-z) + I(y \text{ even}) h^{i/2*}(y/2)^2. \qquad (5.15)$$
$$(y = 0, 1, 2, \dots)$$

For some classes of severity distributions, there are other ways of evaluating these convolutions. For instance, if h belongs to $\mathcal{P}_{1\underline{0}}$ and satisfies the recursion (2.38), then Theorem 2.7 gives that

$$h^{i*}(y) = \left(a + \frac{(a+b)i - a}{y} \right) h^{i*}(y-1). \qquad (i, y = 1, 2, \dots)$$

5.3 The $\mathcal{R}_k$ Classes

5.3.1 Definitions and Characterisation

If the extension (5.13) of the Panjer recursion (2.79) should be of interest, then there have to exist interesting counting distributions p satisfying (5.12). To find out about that, we shall in the following look into the properties of these distributions, and to do that, we need some notation.

We denote the distribution p given by (5.12) by $R_k[a, b]$. For simplicity, we often drop the argument in the functions a and b when $k = 1$. We denote by $\mathcal{R}_k$ the class of all distributions in $\mathcal{P}_{1\underline{0}}$ that can be represented in the form $R_k[a, b]$ with the same k. As a distribution in $\mathcal{R}_{k-1}$ can be expressed as a distribution in $\mathcal{R}_k$ by letting $a(k) = b(k) = 0$, we have that $\mathcal{R}_{k-1} \subset \mathcal{R}_k$. Hence, it is of interest to study the sets $\mathcal{R}_k^0 = \mathcal{R}_k \sim \mathcal{R}_{k-1}$ for $k = 1, 2, \dots$, that is, the set of distributions that satisfy (5.12), but not such a recursion of lower order. The class $\mathcal{R}_0$ consists of only the degenerate distribution concentrated in zero. We also introduce the limiting classes $\mathcal{R}_\infty = \mathcal{P}_{1\underline{0}}$ and $\mathcal{R}_\infty^0 = \mathcal{R}_\infty \sim \bigcup_{k=0}^\infty \mathcal{R}_k$, that is, the set of distributions in $\mathcal{P}_{1\underline{0}}$ that cannot be represented in the form $R_k[a, b]$ with a finite k.

For studying the properties of the $\mathcal{R}_k$ classes, it will be convenient to introduce the functions

$$\rho_p(s) = \frac{\mathrm{d}}{\mathrm{d}s} \ln \tau_p(s) = \frac{\tau_p'(s)}{\tau_p(s)}. \qquad (p \in \mathcal{P}_{1\underline{0}}) \tag{5.16}$$

The distribution p is uniquely determined by ρ_p.

Theorem 5.2 *A distribution $p \in \mathcal{P}_{1\underline{0}}$ can be expressed as $R_k[a, b]$ iff*

$$\rho_p(s) = \frac{\sum_{i=1}^k (ia(i) + b(i))s^{i-1}}{1 - \sum_{i=1}^k a(i)s^i} = \frac{\tau_a'(s) + \tau_b(s)/s}{1 - \tau_a(s)}. \tag{5.17}$$

Proof If p can be expressed as $R_k[a, b]$, then

$$\begin{aligned}
\tau_p'(s) &= \sum_{n=1}^\infty ns^{n-1}p(n) = \sum_{n=1}^\infty ns^{n-1} \sum_{i=1}^k \left(a(i) + \frac{b(i)}{n}\right) p(n-i) \\
&= \sum_{i=1}^k \sum_{n=1}^\infty s^{n-1}(na(i) + b(i))p(n-i) \\
&= \sum_{i=1}^k \sum_{n=0}^\infty s^{n+i-1}((n+i)a(i) + b(i))p(n) \\
&= \sum_{i=1}^k \left((ia(i) + b(i))s^{i-1} \sum_{n=0}^\infty s^n p(n) + a(i)s^i \sum_{n=0}^\infty ns^{n-1} p(n) \right)
\end{aligned}$$

$$= \sum_{i=1}^{k} ((ia(i) + b(i))s^{i-1}\tau_p(s) + a(i)s^i \tau_p'(s)),$$

that is,

$$\tau_p'(s)\left(1 - \sum_{i=1}^{k} a(i)s^i\right) = \tau_p(s) \sum_{i=1}^{k} (ia(i) + b(i))s^{i-1},$$

from which we obtain (5.17). Hence, (5.17) is a necessary condition for p to be representable in the form $R_k[a, b]$. As ρ_p determines p uniquely, it is also a sufficient condition, and, thus, Theorem 5.2 is proved. □

From this theorem, we immediately get the following characterisation of $\mathcal{R}_k$.

Corollary 5.1 *A distribution in $\mathcal{P}_{1\underline{0}}$ belongs to $\mathcal{R}_k$ iff the derivative of the natural logarithm of its generating function can be expressed as the ratio between a polynomial of degree at most $k-1$ and a polynomial of degree at most k with a non-zero constant term.*

Combining Theorem 5.2 with (3.3) gives the following corollary.

Corollary 5.2 *A mixed Poisson distribution with mixing distribution U can be expressed as $R_k[a, b]$ iff*

$$\frac{\mathrm{d}}{\mathrm{d}s} \ln \gamma_U(1-s) = -\frac{\gamma_U'(1-s)}{\gamma_U(1-s)} = \frac{\sum_{i=1}^{k}(ia(i) + b(i))s^{i-1}}{1 - \sum_{i=1}^{k} a(i)s^i}.$$

By multiplying numerator and denominator in the first fraction of (5.17) by $1 - qs$ for some constant q and expressing these products as polynomials in s, we can express ρ_p in the form

$$\rho_p(s) = \frac{\sum_{i=1}^{k+1}(ia_q(i) + b_q(i))s^{i-1}}{1 - \sum_{i=1}^{k+1} a_q(i)s^i}$$

with the functions a_q and b_q depending on q. Then we have that p is $R_{k+1}[a_q, b_q]$. As this holds for all values of q, we see that for a distribution $R_{k+1}[c, d] \in \mathcal{R}_{k+1}$, the functions c and d are not unique.

Example 5.1 Let us look at the simple case where p is the Poisson distribution $\mathrm{Po}(\lambda)$, that is, p is $R_1[0, \lambda]$. Then, for any number q,

$$\rho_p(s) = \lambda = \frac{\lambda(1 - qs)}{1 - qs},$$

from which we see that p is $R_2[a_q, b_q]$ with $a_q(1) = q$, $b_q(1) = \lambda - q$, $a_q(2) = 0$, and $b_q(2) = -q\lambda$, Hence,

$$p(n) = \left(q + \frac{\lambda - q}{n}\right) p(n-1) - \frac{q\lambda}{n} p(n-2), \qquad (n = 1, 2, \dots)$$

and this holds for any value of q. □

We have now seen that the functions a and b are not unique for a distribution $R_k[a, b] \in \mathcal{R}_k$, but what if $R_k[a, b] \in \mathcal{R}_k^0$? The following theorem gives the answer.

Theorem 5.3 *If $R_k[a, b] \in \mathcal{R}_k^0$ with $k < \infty$, then the functions a and b are unique.*

Proof If p is $R_k[a, b] \in \mathcal{R}_k^0$, then there exists no $k' < k$ and no functions a' and b' such that p is $R_{k'}[a', b']$, that is,

$$\rho_p(s) = \frac{\sum_{i=1}^{k'} (ia'(i) + b'(i)) s^{i-1}}{1 - \sum_{i=1}^{k'} a'(i) s^i}.$$

This means that the polynomials in the numerator and denominator in the first fraction of (5.17) cannot have any common factor, and hence, their coefficients must be unique, that is, a and b are unique. □

In Theorem 5.3, we had to make the assumption that $k < \infty$ as the uniqueness property does not hold for $\mathcal{R}_\infty^0$. If p is $R_\infty[a, b]$, then (5.12) can be written as

$$p(n) = \sum_{i=1}^{n} \left(a(i) + \frac{b(i)}{n}\right) p(n-i). \qquad (n = 1, 2, \dots) \tag{5.18}$$

From this, we easily see that the uniqueness property cannot hold for any distribution $R_\infty[a, b] \in \mathcal{R}_\infty$; for instance, for any function a, we can solve (5.18) for $b(n)$. Then we obtain

$$b(n) = \frac{1}{p(0)} \left(np(n) - \sum_{i=1}^{n-1} (na(i) + b(i)) p(n-i)\right) - na(n), \qquad (n = 1, 2, \dots) \tag{5.19}$$

which gives a recursion for b. From this, we see that for any given $p \in \mathcal{R}_\infty = \mathcal{P}_{1\underline{0}}$ and any given function a, there exists a unique b such that p can be represented as $R_\infty[a, b]$. For instance, we can choose $a \equiv 0$. In that case, (5.18) and (5.19) reduce to

$$p(n) = \frac{1}{n} \sum_{i=1}^{n} b(i) p(n-i); \qquad b(n) = \frac{1}{p(0)} \left(np(n) - \sum_{i=1}^{n-1} b(i) p(n-i)\right).$$

$$(n = 1, 2, \dots)$$

This looks familiar; by comparison with (4.14), we see that with this choice of a, b becomes the De Pril transform φ_p of p, so that any $p \in \mathcal{P}_{1\underline{0}}$ can be represented as $R_\infty[0, \varphi_p]$. Thus, many of the results that we prove for distributions in the $\mathcal{R}_k$ classes, can also be applied in connection with De Pril transforms. We shall return to that in Chap. 6.

5.3.2 Compound Distributions

In the following theorem, we reformulate the recursion (5.13) with our new notation, also taking into account the effect of finite support for h.

Theorem 5.4 *The compound distribution with counting distribution $R_k[a, b]$ and severity distribution $h \in \mathcal{P}_{10}$ is $R_{mk}[\tilde{a}, \tilde{b}]$ with $m = \sup\{y : h(y) > 0\}$ and $\tilde{a}$ and $\tilde{b}$ given by* (5.10).

In the following corollary, we study the effect of thinning on $R_k[a, b]$.

Corollary 5.3 *The π-thinning of $R_k[a, b]$ is $R_k[a_\pi, b_\pi]$ with*

$$a_\pi(y) = \frac{\pi^y}{1 - \tau_a(1-\pi)} \sum_{i=y}^{k} a(i) \binom{i}{y} (1-\pi)^{i-y}$$

$$b_\pi(y) = \frac{\pi^y}{1 - \tau_a(1-\pi)} \sum_{i=y}^{k} b(i) \binom{i-1}{y-1} (1-\pi)^{i-y}$$

for $y = 1, 2, \ldots, k$.

Proof We let h be the Bernoulli distribution Bern(π). Then, for $i = 1, 2, \ldots, h^{i*}$ is the binomial distribution bin(i, π), and the corollary follows from Theorem 5.4. □

From this corollary, we see that $\mathcal{R}_k$ is closed under thinning.

For $k = 1$, thinning was discussed in Sect. 2.7.2.

By letting the severity distribution h be concentrated in some positive integer m and using that then its i-fold convolution h^{i*} is concentrated in im for all positive integers i, we obtain the following corollary to Theorem 5.4.

Corollary 5.4 *If the random variable N has distribution $R_k[a, b]$ and m is a positive integer, then mN has distribution $R_{mk}[a_m, b_m]$ with*

$$(a_m(y), b_m(y)) = \begin{cases} (a(y/m), mb(y/m)) & (y = m, 2m, \ldots, km) \\ (0, 0). & (otherwise) \end{cases}$$

The following theorem is a reformulation of Theorem 2.3.

Theorem 5.5 *A compound Poisson distribution with Poisson parameter λ and severity distribution $h \in \mathcal{P}_{10}$ that satisfies the conditions of Theorem 2.3, can be expressed as $R_r[\chi, \lambda\eta - \Phi\chi]$.*

5.3.3 Distributions in $\mathcal{P}_{1\underline{0}}$ on the Range $\{0, 1, 2, \ldots, k\}$

The following theorem shows that any distribution in $\mathcal{P}_{1\underline{0}}$ on the range $\{0, 1, 2, \ldots, k\}$ can be expressed as a distribution in $\mathcal{R}_k$.

Theorem 5.6 *Any distribution $p \in \mathcal{P}_{1\underline{0}}$ on the range $\{0, 1, 2, \ldots, k\}$ where k is a positive integer or infinity, can be expressed as $R_k[-p/p(0), 2\Phi p/p(0)]$.*

Proof From (5.16), we obtain

$$\rho_p(s) = \frac{\sum_{i=1}^k i s^{i-1} p(i)}{\sum_{i=0}^k s^i p(i)} = \frac{\sum_{i=1}^k (i(-\frac{p(i)}{p(0)}) + 2i\frac{p(i)}{p(0)}) s^{i-1}}{1 - \sum_{i=1}^k (-\frac{p(i)}{p(0)}) s^i}.$$

From Theorem 5.2 follows that p is $R_k[-p/p(0), 2\Phi p/p(0)]$. □

5.3.4 Convolutions

In Sect. 5.3.4, we shall prove some results on convolutions. We start softly with a lemma.

Lemma 5.1 *If $p, q \in \mathcal{P}_{1\underline{0}}$, then $\rho_{p*q} = \rho_p + \rho_q$.*

Proof We have

$$\rho_{p*q}(s) = \frac{\mathrm{d}}{\mathrm{d}s} \ln \tau_{p*q}(s) = \frac{\mathrm{d}}{\mathrm{d}s} \ln(\tau_p(s)\tau_q(s)) = \frac{\mathrm{d}}{\mathrm{d}s} (\ln \tau_p(s) + \ln \tau_q(s))$$
$$= \frac{\mathrm{d}}{\mathrm{d}s} \ln \tau_p(s) + \frac{\mathrm{d}}{\mathrm{d}s} \ln \tau_q(s) = \rho_p(s) + \rho_q(s).$$ □

From this theorem, we realise that if ρ_p and ρ_q are represented in the form (5.17), then we can find an analogous representation of ρ_{p*q} by arranging the sum of these two fractions as one fraction and sorting out the coefficients of the polynomials of the denominator and the numerator. The simplest case is obviously when the two initial fractions have the same denominator. Then the sum of these fractions has the same denominator, and the numerator is the sum of the numerators of the initial fractions. We start with that case.

Theorem 5.7 *The convolution of the* M *distributions* $R_k[a, b_1], R_k[a, b_2], \ldots, R_k[a, b_M]$ *can be expressed as* $R_k[a, (M-1)\Phi a + b_{\bullet M}]$.

Proof For $i = 1, 2, \ldots, M$, let p_i denote the distribution $R_k[a, b_i]$. By application of Lemma 5.1 and Theorem 5.2, we obtain

$$\rho_{*_{j=1}^M p_j}(s) = \sum_{j=1}^M \rho_{p_j}(s) = \sum_{j=1}^M \frac{\sum_{i=1}^k (ia(i) + b_j(i))s^{i-1}}{1 - \sum_{i=1}^k a(i)s^i}$$

$$= \frac{\sum_{i=1}^k (Mia(i) + b_{\bullet M}(i))s^{i-1}}{1 - \sum_{i=1}^k a(i)s^i}$$

$$= \frac{\sum_{i=1}^k (ia(i) + ((M-1)\Phi a + b_{\bullet M})(i))s^{i-1}}{1 - \sum_{i=1}^k a(i)s^i},$$

which gives that $*_{j=1}^M p_j$ is $R_k[a, (M-1)\Phi a + b_{\bullet M}]$. □

By letting all the b_js be equal, we obtain the following corollary.

Corollary 5.5 *The* M*-fold convolution of* $R_k[a, b]$ *is* $R_k[a, (M-1)\Phi a + Mb]$.

By combining this with Theorem 5.6, we obtain the following corollary.

Corollary 5.6 *For any distribution* $p \in \mathcal{P}_{1\underline{0}}$ *on the range* $\{0, 1, 2, \ldots, k\}$ *where* k *is a positive integer or infinity,* p^{M*} *can be expressed as* $R_k[-p/p(0), (M+1)\Phi p/p(0)]$, *so that*

$$p^{M*}(n) = \frac{1}{p(0)} \sum_{i=1}^k \left((M+1)\frac{i}{n} - 1\right) p(i) p^{M*}(n-i). \qquad (n = 1, 2, \ldots)$$

This is the same recursion that we got in Theorem 2.8.

Example 5.2 Let $p \in \mathcal{P}_{1\underline{0}}$ be the discrete uniform distribution given by (2.24). Then Theorem 5.6 gives that p can be expressed as $R_k[a, b]$ with

$$a(i) = -1; \qquad b(i) = 2i. \qquad (i = 1, 2, \ldots, k) \tag{5.20}$$

If $h \in \mathcal{P}_{10}$, then Theorem 5.4 gives

$$(p \vee h)(x) = \frac{1 - h(0)}{x(1 - h(0)^{k+1})} \sum_{y=1}^x (2y - x)(p \vee h)(x - y) \sum_{i=1}^k h^{i*}(y),$$
$$(x = 1, 2, \ldots)$$

and in Example 2.6, we presented a recursion for p^{M*}. In the following, we shall deduce simpler recursions for p^{M*} and $p \vee h$ in the present case.

Insertion of (2.25) and (2.26) in the last expression in (5.16) gives

$$\rho_p(s) = \frac{1-(k+1)s^k + ks^{k+1}}{(1-s^{k+1})(1-s)} = \frac{1-(k+1)s^k + ks^{k+1}}{1-(s+s^{k+1}-s^{k+2})}. \tag{5.21}$$

Hence, p can be expressed as $R_{k+2}[\alpha, \beta]$ with

$$\alpha(1) = \alpha(k+1) = 1; \qquad \alpha(k+2) = -1$$
$$\beta(1) = 0; \qquad \beta(k+1) = -2(k+1); \qquad \beta(k+2) = 2(k+1)$$
$$\alpha(i) = \beta(i) = 0, \qquad (i = 2, 3, \dots, k)$$

that is,

$$p(n) = p(n-1) + \left(1 - 2\frac{k+1}{n}\right)(p(n-k-1) - p(n-k-2)).$$
$$(n = 1, 2, \dots)$$

Application of Corollary 5.5 gives that p^{M*} can be expressed as $R_{k+2}[\alpha, \beta_M]$ with

$$\beta_M(1) = M - 1$$
$$\beta_M(i) = 0 \qquad (i = 2, 3, \dots, k)$$
$$\beta_M(k+1) = -(M+1)(k+1); \qquad \beta_M(k+2) = (M+1)k + 2$$

that is,

$$p^{M*}(n) = \left(1 + \frac{M-1}{n}\right)p^{M*}(n-1) + \left(1 - \frac{(M+1)(k+1)}{n}\right)p^{M*}(n-k-1) + \left(\frac{(M+1)k+2}{n} - 1\right)p^{M*}(n-k-2), \quad (n = 1, 2, \dots) \tag{5.22}$$

and Theorem 5.4 gives that

$$(p \vee h)(x) = \frac{1}{(1-h(0)^{k+1})(1-h(0))} \sum_{y=1}^{x} \left(h(y) + \left(1 - 2\frac{y}{x}\right)h^{(k+1)*}(y) + \left(2\frac{k+1}{k+2}\frac{y}{x} - 1\right)h^{(k+2)*}(y)\right)(p \vee h)(x-y). \qquad (x = 1, 2, \dots) \tag{5.23}$$

If k is large, then it seems reasonable to evaluate $h^{(k+1)*}$ recursively by Theorem 2.10 and $h^{(k+2)*}$ by

$$h^{(k+2)*}(y) = \sum_{z=0}^{y} h(z)h^{(k+1)*}(y-z). \qquad (y = 0, 1, 2, \dots)$$

It is interesting to note that in the present case, although p can be expressed as a distribution in $\mathcal{R}_k$, it seems more efficient to express it as a distribution in $\mathcal{R}_{k+2}$ when k is large.

In Example 5.4, we shall deduce a much simpler recursion for $p \vee h$. □

When $k = 1$, Theorem 5.7 reduces to the following corollary, which is equivalent with Theorem 2.7.

Corollary 5.7 *The convolution of the M distributions $R_1[a, b_1], R_1[a, b_2], \ldots, R_1[a, b_M]$ can be expressed as $R_1[a, (M-1)a + b_{\bullet M}]$.*

In particular, we obtain that the M-fold convolution of $R_1[a, b]$ can be expressed as $R_1[a, (M-1)a + Mb]$.

Till now, we have discussed convolutions of distributions with the same function a. Unfortunately, the situation gets much more complicated when they do not have the same a. We turn to that case now.

Theorem 5.8 *The convolution of $R_{k_1}[a_1, b_1]$ and $R_{k_2}[a_2, b_2]$ can be expressed as $R_{k_1+k_2}[a, b]$ with*

$$a(i) = a_1(i) + a_2(i) - \sum_{j=\max(1,i-k_2)}^{\min(i-1,k_1)} a_1(j)a_2(i-j) \tag{5.24}$$

$$b(i) = b_1(i) + b_2(i) - \sum_{j=\max(1,i-k_2)}^{\min(i-1,k_1)} (b_1(j)a_2(i-j) + a_1(j)b_2(i-j))$$

for $i = 1, 2, \ldots, k_1 + k_2$.

Proof Let p_j be $R_{k_j}[a_j, b_j]$ for $j = 1, 2$. By using Lemma 5.1 and Theorem 5.2, we obtain

$$\rho_{p_1 * p_2}(s) = \sum_{j=1}^{2} \rho_{p_j}(s) = \sum_{j=1}^{2} \frac{\sum_{i=1}^{k_j}(ia_j(i) + b_j(i))s^{i-1}}{1 - \sum_{i=1}^{k_j} a_j(i)s^i},$$

which gives

$$\rho_{p_1 * p_2}(s) = \left(\sum_{i=1}^{k_1}(ia_1(i) + b_1(i))s^{i-1}\left(1 - \sum_{j=1}^{k_2} a_2(j)s^j\right)\right.$$
$$\left. + \sum_{i=1}^{k_2}(ia_2(i) + b_2(i))s^{i-1}\left(1 - \sum_{j=1}^{k_1} a_1(j)s^j\right)\right)$$

$$\Bigg/\left(\left(1-\sum_{i=1}^{k_1}a_1(i)s^i\right)\left(1-\sum_{j=1}^{k_2}a_2(j)s^j\right)\right)$$

$$=\frac{\sum_{i=1}^{k_1+k_2}(ic(i)+d(i))s^{i-1}}{1-\sum_{i=1}^{k_1+k_2}c(i)s^i} \tag{5.25}$$

for some functions c and d. It remains to show that $c=a$ and $d=b$.

By equating the denominators in (5.25), we obtain

$$\begin{aligned}\sum_{i=1}^{k_1+k_2}c(i)s^i &= 1-\left(1-\sum_{i=1}^{k_1}a_1(i)s^i\right)\left(1-\sum_{i=1}^{k_2}a_2(i)s^i\right)\\ &=\sum_{i=1}^{k_1}a_1(i)s^i+\sum_{i=1}^{k_2}a_2(i)s^i-\sum_{j=1}^{k_1}\sum_{i=1}^{k_2}a_1(j)a_2(i)s^{i+j}\\ &=\sum_{i=1}^{k_1+k_2}\left(a_1(i)+a_2(i)-\sum_{j=\max(1,i-k_2)}^{\min(i-1,k_1)}a_1(j)a_2(i-j)\right)s^i\\ &=\sum_{i=1}^{k_1+k_2}a(i)s^i.\end{aligned}$$

Hence, $c=a$.

By using this and equating the numerators in (5.25), we obtain

$$\begin{aligned}&\sum_{i=1}^{k_1+k_2}(ia(i)+d(i))s^{i-1}\\ &\quad=\sum_{i=1}^{k_1+k_2}(ic(i)+d(i))s^{i-1}=\sum_{i=1}^{k_1}(ia_1(i)+b_1(i))s^{i-1}\left(1-\sum_{j=1}^{k_2}a_2(j)s^j\right)\\ &\qquad+\sum_{j=1}^{k_2}(ja_2(j)+b_2(j))s^{j-1}\left(1-\sum_{i=1}^{k_1}a_1(i)s^i\right)\\ &\quad=\sum_{i=1}^{k_1+k_2}\Bigg(i(a_1(i)+a_2(i))+b_1(i)+b_2(i)\\ &\qquad-\sum_{j=\max(1,i-k_2)}^{\min(i-1,k_1)}(ia_1(j)a_2(i-j)+b_1(j)a_2(i-j)+a_1(j)b_2(i-j))\Bigg)s^{i-1}\\ &\quad=\sum_{i=1}^{k_1+k_2}(ia(i)+b(i))s^{i-1}.\end{aligned}$$

Hence, $d=b$.

This completes the proof of Theorem 5.8. □

We can obviously deduce a representation of lower order for the convolution when the denominators $1-\sum_{i=1}^{k_1} a_1(i)s^i$ and $1-\sum_{j=1}^{k_2} a_2(j)s^j$ have a common factor. The extreme case is when $a_1 = a_2$; that is the case treated in Theorem 5.7.

Theorem 5.8 gives in particular that the convolution of $R_1[a_1, b_1]$ and $R_1[a_2, b_2]$ can be expressed as $R_2[a, b]$ with

$$a(1) = a_1 + a_2; \qquad a(2) = -a_1 a_2 \tag{5.26}$$

$$b(1) = b_1 + b_2; \qquad b(2) = -a_1 b_2 - a_2 b_1. \tag{5.27}$$

We see that $a(2) = 0$ iff at least one of these distributions is Poisson.

Example 5.3 Let p be a mixed Poisson distribution with shifted Gamma mixing distribution with density

$$u(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}(\theta-\lambda)^{\alpha-1}e^{-\beta(\theta-\lambda)}. \qquad (\theta > \lambda;\, \beta, \alpha, \lambda > 0)$$

In this case, p is called a *Delaporte distribution*.

Our shifted Gamma distribution can be interpreted as the convolution of the distribution concentrated in λ and the Gamma distribution $\text{Gamma}(\alpha, \beta)$. Hence, Theorem 3.1 gives that p is the convolution of the Poisson distribution $\text{Po}(\lambda)$, that is, $R_1[0, \lambda]$, and a mixed Poisson distribution with mixing distribution $\text{Gamma}(\alpha, \beta)$. In Sect. 3.2, we showed that that mixed distribution is the negative binomial distribution $\text{NB}(\alpha, (\beta+1)^{-1})$, and, from Table 2.1, we obtain that that distribution is $R_1[(\beta+1)^{-1}, (\alpha-1)(\beta+1)^{-1}]$. Application of (5.26) and (5.27) gives that p is $R_2[a, b]$ with

$$a(1) = \frac{1}{\beta+1}; \qquad a(2) = 0; \qquad b(1) = \frac{\alpha-1}{\beta+1} + \lambda; \qquad b(2) = -\frac{\lambda}{\beta+1}. \tag{5.28}$$

Hence,

$$p(n) = \frac{1}{\beta+1}\left(\left(1 + \frac{\alpha+\beta\lambda+\lambda-1}{n}\right)p(n-1) - \frac{\lambda}{n}p(n-2)\right).$$
$$(n = 1, 2, \dots)$$

If $f = p \vee h$ with $h \in \mathcal{P}_{10}$, then (5.13) gives that

$$f(x) = \frac{1}{\beta+1-h(0)} \times \sum_{y=1}^{x} f(x-y)\left(\left(1 + (\alpha+\beta\lambda+\lambda-1)\frac{y}{x}\right)h(y) - \frac{\lambda}{2}\frac{y}{x}h^{2*}(y)\right).$$
$$(x = 1, 2, \dots)$$

As p is the convolution between a Poisson distribution and a negative binomial distribution, we could also have evaluated p and f by the algorithms of Example 4.4. □

To extend Theorem 5.8 to the convolution of M distributions seems rather messy. However, let us have a go on the special case of M distributions in $\mathcal{R}_1$.

Theorem 5.9 *The convolution of the M distributions $R_1[a_1, b_1], R_1[a_2, b_2], \ldots, R_1[a_M, b_M]$ can be expressed as $R_M[a, b]$ with*

$$a(i) = (-1)^{i+1} \sum_{1 \le j_1 < j_2 < \cdots < j_i \le M} \prod_{k=1}^{i} a_{j_k} \qquad (i = 1, 2, \ldots, M)$$

$$\begin{aligned} b(i) &= (-1)^{i+1} \sum_{l=1}^{M} b_l \sum_{\substack{1 \le j_1 < j_2 < \cdots < j_{i-1} \le M \\ j_k \ne l \; (k=1,2,\ldots,i-1)}} \prod_{k=1}^{i-1} a_{j_k} - (i-1)a(i) \\ &= (-1)^{i+1} \sum_{l=1}^{M} (b_l - (i-1)a_l) \sum_{\substack{1 \le j_1 < j_2 < \cdots < j_{i-1} \le M \\ j_k \ne l \; (k=1,2,\ldots,i-1)}} \prod_{k=1}^{i-1} a_{j_k} \end{aligned}$$

$$(i = 2, 3, \ldots, M)$$

$$b(1) = b_{\bullet M}.$$

Proof That the convolution of M distributions in $\mathcal{R}_1$ is in $\mathcal{R}_M$, follows from Theorem 5.8. Hence, there exist functions c and d such that the convolution is $R_M[c, d]$. It remains to show that $c = a$ and $b = d$.

Application of Theorem 5.2 and Lemma 5.1 gives

$$\frac{\sum_{i=1}^{M} (ic(i) + d(i))s^{i-1}}{1 - \sum_{i=1}^{M} c(i)s^i} = \sum_{i=1}^{M} \frac{a_i + b_i}{1 - a_i s} = \frac{\sum_{i=1}^{M} (a_i + b_i) \prod_{j \ne i} (1 - a_j s)}{\prod_{i=1}^{M} (1 - a_i s)}.$$

By equating the denominators, we obtain

$$\sum_{i=1}^{M} c(i)s^i = 1 - \prod_{i=1}^{M} (1 - a_i s) = -\sum_{i=1}^{M} (-s)^i \sum_{1 \le j_1 < j_2 < \cdots < j_i \le M} \prod_{k=1}^{i} a_{j_k} = \sum_{i=1}^{M} a(i)s^i,$$

that is, $c = a$, and equating the numerators gives

$$\begin{aligned} &\sum_{i=1}^{M} (ia(i) + d(i))s^{i-1} \\ &\quad = \sum_{i=1}^{M} (ic(i) + d(i))s^{i-1} = \sum_{i=1}^{M} (a_i + b_i) \prod_{j \ne i} (1 - a_j s) \end{aligned}$$

$$= \sum_{l=1}^{M}(a_l + b_l) + \sum_{i=2}^{M}(-s)^{i-1}\sum_{l=1}^{M}(a_l + b_l) \sum_{\substack{1 \le j_1 < j_2 < \cdots < j_{i-1} \le M \\ j_k \ne l \ (k=1,2,\ldots,i-1)}} \prod_{k=1}^{i-1} a_{j_k}$$

$$= \sum_{i=1}^{M}(ia(i) + b(i))s^{i-1},$$

that is, $d = b$.

This completes the proof of Theorem 5.9. □

In particular, we obtain $a(1) = \sum_{i=1}^{M} a_i$.

Even in the present case, it seems more convenient to evaluate a and b recursively by Theorem 5.8 than by using Theorem 5.9. Such a procedure is described in the following corollary that immediately follows from Theorem 5.8.

Corollary 5.8 *For $k = 1, 2, \ldots,$ the convolution of the k distributions $R_1[a_1, b_1]$, $R_1[a_2, b_2], \ldots, R_1[a_k, b_k]$ can be expressed as $R_k[\alpha_k, \beta_k]$, where the functions α_k and β_k can be evaluated recursively by*

$$\alpha_k(i) = \begin{cases} \alpha_{k-1}(i) - a_k\alpha_{k-1}(i-1) & (i = 2, 3, \ldots, k) \\ \alpha_{k-1}(1) + a_k & (i = 1) \end{cases} \tag{5.29}$$

$$\beta_k(i) = \begin{cases} \beta_{k-1}(i) - a_k\beta_{k-1}(i-1) - b_k\alpha_{k-1}(i-1) & (i = 2, 3, \ldots, k) \\ \beta_{k-1}(1) + b_k. & (i = 1) \end{cases} \tag{5.30}$$

with initial values $\alpha_1(1) = a_1$ and $\beta_1(1) = b_1$.

5.3.5 Cumulants

The following theorem gives a recursion for the cumulants of a compound $\mathcal{R}_k$ distribution.

Theorem 5.10 *If $F = p \vee H$ where p is $R_k[a, b]$ and H is a univariate distribution, then*

$$\kappa_F(j) = \frac{1}{1 - \mu_a(0)}\left(\mu_{(a+\Psi b)\vee H}(j) + \sum_{u=1}^{j-1}\binom{j-1}{u}\mu_{a\vee H}(u)\kappa_F(j-u)\right),$$
$$(j = 1, 2, \ldots) \tag{5.31}$$

provided that the quantities in this formula exist.

Proof By application of (1.11), (1.31), (5.16), and (5.17), we obtain

$$\theta_F'(s) = \frac{d}{ds}\ln\omega_F(s) = \frac{d}{ds}\ln\tau_p(\omega_H(s)) = \frac{\tau_p'(\omega_H(s))}{\tau_p(\omega_H(s))}\omega_H'(s)$$
$$= \rho_p(\omega_H(s))\omega_H'(s) = \frac{\sum_{i=1}^k (ia(i)+b(i))\omega_H(s)^{i-1}}{1-\sum_{i=1}^k a(i)\omega_H(s)^i}\omega_H'(s),$$

which gives

$$\theta_F'(s) = \sum_{i=1}^k (ia(i)+b(i))\omega_H(s)^{i-1}\omega_H'(s) + \theta_F'(s)\sum_{i=1}^k a(i)\omega_H(s)^i$$
$$= \frac{d}{ds}\sum_{i=1}^k \left(a(i)+\frac{b(i)}{i}\right)\omega_H(s)^i + \omega_{a\vee H}(s)\theta_F'(s)$$
$$= \omega'_{(a+\Psi b)\vee H}(s) + \omega_{a\vee H}(s)\theta_F'(s),$$

that is,

$$\sum_{j=1}^\infty \frac{\kappa_F(j)}{j!}js^{j-1} = \sum_{j=1}^\infty \frac{\mu_{(a+\Psi b)\vee H}(j)}{j!}js^{j-1} + \sum_{u=0}^\infty \frac{\mu_{a\vee H}(u)}{u!}s^u \sum_{j=1}^\infty \frac{\kappa_F(j)}{j!}js^{j-1},$$

which we rewrite as

$$\sum_{j=1}^\infty \frac{\kappa_F(j)}{(j-1)!}s^{j-1} = \sum_{j=1}^\infty \frac{1}{(j-1)!}\Bigg(\mu_{(a+\Psi b)\vee H}(j)$$
$$+ \sum_{u=0}^{j-1}\binom{j-1}{u}\mu_{a\vee H}(u)\kappa_F(j-u)\Bigg)s^{j-1}.$$

Comparison of coefficients gives that for $j = 1, 2, \ldots$, we have

$$\kappa_F(j) = \mu_{(a+\Psi b)\vee H}(j) + \sum_{u=0}^{j-1}\binom{j-1}{u}\mu_{a\vee H}(u)\kappa_F(j-u).$$

Solving for $\kappa_F(j)$ gives

$$\kappa_F(j) = \frac{1}{1-\mu_{a\vee H}(0)}\left(\mu_{(a+\Psi b)\vee H}(j) + \sum_{u=1}^{j-1}\binom{j-1}{u}\mu_{a\vee H}(u)\kappa_F(j-u)\right).$$

As

$$\mu_{a\vee H}(0) = \sum_{i=1}^k a(i)\mu_{H^{i*}}(0) = \sum_{i=1}^k a(i) = \mu_a(0),$$

(5.31) follows. □

For H concentrated in one, (5.31) gives

$$\kappa_p(j) = \frac{1}{1-\mu_a(0)}\left(\mu_a(j) + \mu_b(j-1) + \sum_{u=1}^{j-1}\binom{j-1}{u}\mu_a(u)\kappa_p(j-u)\right).$$
$$(j = 1, 2, \ldots) \tag{5.32}$$

In the special case when p is $R_1[a,b]$, (5.31) and (5.32) reduce to

$$\kappa_F(j) = \frac{1}{1-a}\left((a+b)\mu_H(j) + a\sum_{u=1}^{j-1}\binom{j-1}{u}\mu_H(u)\kappa_F(j-u)\right) \tag{5.33}$$

$$\kappa_p(j) = \frac{1}{1-a}\left(a+b+a\sum_{u=1}^{j-1}\binom{j-1}{u}\kappa_p(j-u)\right) \tag{5.34}$$

for $j = 1, 2, \ldots$.

Let us apply the recursions (5.33) and (5.34) to the three main classes in $\mathcal{R}_1$. The values of a and b are found in Table 2.1.

1. *Poisson distribution* $\mathrm{Po}(\lambda)$. In this case, we obtain

$$\kappa_F = \lambda\mu_H \tag{5.35}$$

and $\kappa_p \equiv \lambda$. By combining this with Theorem 4.2, we see that an infinitely divisible distribution in $\mathcal{P}_{1\underline{0}}$ whose cumulant generating function exists, has non-negative cumulants of all orders.

2. *Binomial distribution* $\mathrm{bin}(M, \pi)$. For $j = 1, 2, \ldots$, we obtain

$$\kappa_F(j) = \pi\left(M\mu_H(j) - \sum_{u=1}^{j-1}\binom{j-1}{u}\mu_H(u)\kappa_F(j-u)\right)$$

$$\kappa_p(j) = \pi\left(M - \sum_{u=1}^{j-1}\binom{j-1}{u}\kappa_p(j-u)\right). \tag{5.36}$$

3. *Negative binomial distribution* $\mathrm{NB}(\alpha, \pi)$. For $j = 1, 2, \ldots$, we obtain

$$\kappa_F(j) = \frac{\pi}{1-\pi}\left(\alpha\mu_H(j) + \sum_{u=1}^{j-1}\binom{j-1}{u}\mu_H(u)\kappa_F(j-u)\right)$$

$$\kappa_p(j) = \frac{\pi}{1-\pi}\left(\alpha + \sum_{u=1}^{j-1}\binom{j-1}{u}\kappa_p(j-u)\right).$$

Theorem 5.11 *If $h \in \mathcal{P}_{10}$ satisfies the conditions of Theorem 2.3, then its non-central moments satisfy the recursion*

$$\mu_h(j) = \frac{1}{1-\mu_\chi(0)}\left(\mu_\eta(j-1) + \sum_{u=1}^{j-1}\binom{j-1}{u}\mu_\chi(u)\mu_h(j-u)\right),$$
$$(j = 1, 2, \dots) \tag{5.37}$$

provided that the quantities in this formula exist.

Proof Let f be a compound Poisson distribution with Poisson parameter one and severity distribution h. By Theorem 5.5, f can be represented in the form $R_r[\chi, \eta - \Phi\chi]$, and (5.35) gives that $\kappa_F = \mu_h$. By application of these results in (5.32), we obtain (5.37). □

With $r = \infty$, $\eta = \Phi h$, and $\chi \equiv 0$, the right-hand side of (5.37) reduces to $\mu_h(j)$.
In Sect. 9.2, we shall study recursions for moments of compound distributions more generally.

5.4 Counting Distributions with Rational Generating Function

We assume that $p \in \mathcal{P}_{1\underline{0}}$ satisfies the condition

$$\tau_p(s) = \frac{\sum_{i=0}^m \alpha(i)s^i}{1 - \sum_{i=1}^m \beta(i)s^i} \tag{5.38}$$

for some positive integer m and functions α and β. Then

$$\rho_p(s) = \frac{\sum_{i=1}^m i\alpha(i)s^{i-1}(1-\sum_{j=1}^m \beta(j)s^j) + \sum_{i=0}^m \alpha(i)s^i \sum_{j=1}^m j\beta(j)s^{j-1}}{\sum_{i=0}^m \alpha(i)s^i(1-\sum_{j=1}^m \beta(j)s^j)}, \tag{5.39}$$

and Theorem 5.2 gives that $p \in \mathcal{R}_{2m}$. Thus, p can be evaluated recursively by (5.12). However, like in Sect. 2.3.3, we obtain that p satisfies the relation $p = \alpha + \beta * p$, from which we obtain the alternative recursion

$$p(n) = \alpha(n) + \sum_{i=1}^m \beta(i)p(n-i). \qquad (n = 0, 1, 2, \dots) \tag{5.40}$$

If $f = p \vee h$ with $h \in \mathcal{P}_{10}$, then application of (5.40) and (5.8) gives

$$f(x) = \frac{1}{1-\tau_\beta(h(0))}\left(\sum_{i=1}^m \alpha(i)h^{i*}(x) + \sum_{y=1}^x f(x-y)\sum_{i=1}^m \beta(i)h^{i*}(y)\right).$$
$$(x = 1, 2, \dots) \tag{5.41}$$

Example 5.4 Let $p \in \mathcal{P}_{1\underline{0}}$ be the discrete uniform distribution given by (2.24). Then

$$p(n) = \frac{I(n=0) - I(n=k+1)}{k+1} + p(n-1), \qquad (n = 1, 2, \dots)$$

that is, (5.40) is satisfied with

$$m = k+1; \qquad \alpha(0) = \frac{1}{k+1}; \qquad \alpha(k+1) = -\frac{1}{k+1}; \qquad \beta(1) = 1,$$

and $\alpha(i)$ and $\beta(i)$ equal to zero for all other values of i. Application of (5.41) gives the recursion

$$f(x) = \frac{1}{1-h(0)} \left(\sum_{y=1}^{x} h(y) f(x-y) - \frac{h^{(k+1)*}(x)}{k+1} \right), \qquad (x = 1, 2, \dots)$$

which is much simpler than (5.23). The convolution $h^{(k+1)*}$ can be evaluated recursively by Theorem 2.10. □

Further Remarks and References

Schröter (1990) characterised the subclass of distributions $R_2[a, b] \in \mathcal{R}_2$ with $a(2) = 0$. For compound distributions with counting distribution in this class and severity distribution in $\mathcal{P}_{10}$, he deduced the recursion (5.13), and in that connection he indicated (5.4) for general i.

In particular, Schröter showed that his class contained the convolutions of a Poisson distribution and another distribution in the Panjer class. That was a strong indication that convolutions could be a clue for deducing more general results. This was followed up in Sundt (1992), from which most of the results and notation in this chapter originate.

The Delaporte distribution was introduced by Delaporte (1959) for fitting the numbers of claims in an automobile insurance portfolio and has later been studied by Delaporte (1960, 1972a, 1972b), Philipson (1960), Kupper (1962), Albrecht (1981, 1984), Ruohonen (1983, 1988), Willmot and Sundt (1989), and Schröter (1990). The recursions of Example 5.3 were presented by Schröter (1990).

Corollary 5.8 was deduced by Sundt and Vernic (2006).

Example 5.2 is based on Sundt (1999a). Recursive evaluation of the M-fold convolution of a discrete uniform distribution had earlier been discussed by Sundt (1988).

Sundt et al. (1998) deduced the recursion (5.32).

Sundt (2002) presented the procedure for recursive evaluation of a compound distribution with counting distribution given by (2.73) in a more general setting with counting distribution satisfying a recursion of higher order.

Gerhold et al. (2008) presented an extended version of the recursion (5.8).

Kitano et al. (2005) studied the generalised Charlier distribution $p \in \mathcal{P}_{10}$, which satisfies a recursion on the form

$$p(n) = \left(a + \frac{b}{n}\right) p(n-1) + \left(c + \frac{d}{n} + \frac{f}{n-1}\right) p(n-2). \qquad (n = 2, 3, \dots)$$

In particular, they proved Theorem 5.1 for this p.

Sundt and Dickson (2000) discussed the procedure for evaluation of $h^{2*}, h^{3*}, \dots, h^{k*}$ outlined in Sect. 5.2.

With a somewhat different parameterisation, Eisele (2006) deduced the recursion (5.41) within the framework of phase distributions and compared it with the recursion based on the $\mathcal{R}_{2m}$ representation (5.39) of the counting distribution. He also considered the case with continuous severity distribution.

Wang and Sobrero (1994) presented an algorithm for recursive evaluation of compound distributions with severity distribution in $\mathcal{P}_{10}$ and counting distribution $p \in \mathcal{P}_{10}$ that satisfies a recursion in the form

$$p(n) = \sum_{i=1}^{k} \frac{\sum_{j=0}^{t} c(i, j) n^j}{\sum_{j=0}^{t} d(j) n^j} p(n-i). \qquad (n = l+1, l+2, \dots)$$

The recursion (5.11) is obtained by letting

$$t = 1; \qquad d(0) = 0; \qquad d(1) = 1$$
$$c(i, 0) = b(i); \qquad c(i, 1) = a(i). \qquad (i = 1, 2, \dots)$$

Doray and Haziza (2004) discussed statistical inference in the $\mathcal{R}_k$ classes.

Note that in (5.9) and (5.10), we defined $\tilde{q}(y)$, $\tilde{a}(y)$, and $\tilde{b}(y)$ for $y = 1, 2, \dots$; we did not give the definitions

$$\begin{aligned} \tilde{q} &= \frac{q \vee h}{1 - \tau_a(h(0))} \\ \tilde{a} &= \frac{a \vee h}{1 - \tau_a(h(0))} \\ \tilde{b} &= \frac{\Phi(\Psi b \vee h)}{1 - \tau_a(h(0))}. \end{aligned} \tag{5.42}$$

The reason is that when $h \in \mathcal{P}_{10}$, the latter definitions could give a non-zero mass at zero. This could create problems. Let us look at an example. Let p be $R_k[a, b]$ and $h, g \in \mathcal{P}_{10}$. We want to evaluate $(p \vee h) \vee g$. With $\tilde{a}$ and $\tilde{b}$ given by (5.10), two applications of (5.13) give

$$((p \vee h) \vee g)(x) = \frac{1}{1 - \tau_{\tilde{a}}(g(0))} \sum_{y=1}^{x} \left((\tilde{a} \vee g)(y) + \frac{\Phi(\Psi \tilde{b} \vee g)(y)}{x} \right) \times ((p \vee h) \vee g)(x - y).$$

Here, $\tau_{\tilde{a}}(g(0)) = \sum_{y=1}^{\infty} \tilde{a}(y)g(0)^y$, whereas with the definition (5.42), we get $\tau_{\tilde{a}}(g(0)) = \sum_{y=0}^{\infty} \tilde{a}(y)g(0)^y$ with

$$\tilde{a}(0) = \frac{(a \vee h)(0)}{1 - \tau_a(h(0))} = \frac{\tau_a(h(0))}{1 - \tau_a(h(0))}.$$

Chapter 6
De Pril Transforms of Distributions in $\mathcal{P}_{1\underline{0}}$

Summary

Within the context of infinitely divisible distributions in $\mathcal{P}_{1\underline{0}}$, we defined the De Pril transform and studied some of its properties in Sect. 4.4. In the present chapter, we shall study the De Pril transform more generally for all distributions in $\mathcal{P}_{1\underline{0}}$.

In Sect. 6.1, we recapitulate some results from Sect. 4.4, define the De Pril transform of a distribution in $\mathcal{P}_{1\underline{0}}$, and present some of its properties. Most of these properties follow immediately as corollaries to results from Sect. 5.3.

In Sect. 6.2, we more specifically consider De Pril transforms of distributions within the $\mathcal{R}_k$ classes.

6.1 General Results

In Sect. 4.4, we defined the De Pril transform φ_f of a distribution $f \in \mathcal{P}_{1\underline{0}}$ by (4.14) and discussed some of its properties for infinitely divisible distributions. In particular, we

- showed that the De Pril transform of the convolution of M infinitely divisible distributions in $\mathcal{P}_{1\underline{0}}$ is the sum of the De Pril transforms of these distributions
- found the expression (4.19) for the De Pril transform of a compound distribution with severity distribution in $\mathcal{P}_{11}$ and infinitely divisible counting distribution in terms of the severity distribution and the De Pril transform of the counting distribution
- found the expressions (4.17) and (4.20) for the De Pril transform of a Poisson distribution and a negative binomial distribution respectively
- presented De Pril's first method for evaluating the convolution of M infinitely divisible distributions in $\mathcal{P}_{1\underline{0}}$
- showed that a distribution in $\mathcal{P}_{1\underline{0}}$ is infinitely divisible iff its De Pril transform is non-negative.

Consistent with the definition of the De Pril transform of an infinitely divisible distribution in $\mathcal{P}_{1\underline{0}}$ in Sect. 4.4.1, for any distribution $f \in \mathcal{P}_{1\underline{0}}$, we now define its *De Pril transform* $\varphi_f \in \mathcal{P}_{11}$ recursively by

$$\varphi_f(x) = \frac{1}{f(0)}\left(xf(x) - \sum_{y=1}^{x-1}\varphi_f(y)f(x-y)\right). \qquad (x = 1, 2, \dots) \tag{6.1}$$

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_6, © Springer-Verlag Berlin Heidelberg 2009

By solving for $f(x)$, we obtain the recursion

$$f(x)=\frac{1}{x}\sum_{y=1}^{x}\varphi_f(y)f(x-y). \qquad (x=1,2,\dots) \tag{6.2}$$

Multiplication by x gives the more compact representation

$$\Phi f=\varphi_f * f. \tag{6.3}$$

In Sect. 5.3.1, we showed that any distribution $f\in\mathcal{P}_{1\underline{0}}$ can be represented in the form $R_\infty[0,\varphi_f]$. This means that we can obtain results for De Pril transforms as special cases of results that we proved in Chap. 5 for distributions in the $\mathcal{R}_k$ classes; the following corollaries follow immediately from Theorem 5.7, Corollary 5.5, Theorem 5.4, Corollary 5.3, Corollary 5.4, Theorem 5.2, and (5.32) respectively. For infinitely divisible distributions, Corollaries 6.1 and 6.3 were proved already in Sect. 4.4.1.

Corollary 6.1 *The De Pril transform of the convolution of a finite number of distributions in $\mathcal{P}_{1\underline{0}}$ is the sum of the De Pril transforms of these distributions.*

Corollary 6.2 *The De Pril transform of the M-fold convolution of a distribution in $\mathcal{P}_{1\underline{0}}$ is M times the De Pril transform of that distribution.*

Corollary 6.3 *If $p\in\mathcal{P}_{1\underline{0}}$ and $h\in\mathcal{P}_{11}$, then*

$$\varphi_{p\vee h}=\Phi(\Psi\varphi_p\vee h). \tag{6.4}$$

Corollary 6.4 *If p_π is the π-thinning of a distribution $p\in\mathcal{P}_{1\underline{0}}$, then*

$$\varphi_{p_\pi}(x)=\pi^x\sum_{y=x}^{\infty}\varphi_p(y)\binom{y-1}{x-1}(1-\pi)^{y-x}. \qquad (x=1,2,\dots)$$

Corollary 6.5 *If the random variable N has distribution $p\in\mathcal{P}_{1\underline{0}}$ and p_m is the distribution of mN for some positive integer m, then*

$$\varphi_{p_m}(x)=\begin{cases} m\varphi_p(x/m) & (x=m,2m,\dots)\\ 0. & (\textit{otherwise})\end{cases}$$

Corollary 6.6 *If $f\in\mathcal{P}_{1\underline{0}}$, then*

$$\tau_{\varphi_f}=\Phi\rho_f. \tag{6.5}$$

Corollary 6.7 *If $f\in\mathcal{P}_{1\underline{0}}$, then*

$$\kappa_f(j)=\mu_{\varphi_f}(j-1). \qquad (j=1,2,\dots) \tag{6.6}$$

In Sect. 4.4.1, we proved Corollary 6.1 for infinitely divisible distributions in $\mathcal{P}_{1\underline{0}}$. Now that we know that it holds for all distributions in $\mathcal{P}_{1\underline{0}}$, we see that De Pril's first method for evaluating the convolution of M distributions holds for all distributions in $\mathcal{P}_{1\underline{0}}$, not only for the infinitely divisible ones.

Example 6.1 Let $f \in \mathcal{P}_{1\underline{0}}$ be the discrete uniform distribution given by (2.24). Insertion of (2.24) in (6.1) gives

$$\varphi_f(x) = xI(x \le k) - \sum_{y=\max(1,x-k)}^{x-1} \varphi_f(y). \qquad (x = 1, 2, \dots)$$

From this, we obtain that $\varphi_f(x) = 1$ for $x = 1, 2, \dots, k$. For $x = k+1, k+2, \dots$, we have $\varphi_f(x) = -\sum_{y=x-k}^{x-1} \varphi_f(y)$, from which we see that $\varphi_f(y(k+1)) = -k$ for $y = 1, 2, 3, \dots$, and $\varphi_f(x) = 1$ otherwise. Hence,

$$\varphi_f = 1 - (k+1)\delta_k \tag{6.7}$$

with

$$\delta_k(x) = \begin{cases} 1 & (x = k+1, 2(k+1), \dots) \\ 0. & (\text{otherwise}) \end{cases}$$

For $j = 1, 2, \dots, M$, let f_j be the discrete uniform distribution on $0, 1, 2, \dots, k_j$. Then application of (6.7) and Corollary 6.1 gives

$$\varphi_{*_{j=1}^M f_j} = \sum_{j=1}^M \varphi_{f_j} = M - \sum_{j=1}^M (k_j + 1)\delta_{k_j}, \tag{6.8}$$

and by insertion in (6.2), we obtain

$$\begin{aligned}(*_{j=1}^M f_j)(x) &= \frac{1}{x}\sum_{y=1}^x \left(M - \sum_{j=1}^M (k_j+1)\delta_{k_j}(y)\right)(*_{j=1}^M f_j)(x-y) \\ &= \frac{1}{x}\Bigg(M\Gamma(*_{j=1}^M f_j)(x-1) \\ &\quad - \sum_{j=1}^M (k_j+1) \sum_{z=1}^{[x/(k_j+1)]} (*_{j=1}^M f_j)(x - z(k_j+1))\Bigg).\end{aligned}$$

$(x = 1, 2, \dots)$

In particular,

$$f^{M*}(x) = \frac{M}{x}\sum_{y=1}^x (1 - (k+1)\delta_k(y)) f^{M*}(x-y)$$

$$= \frac{M}{x}\left(\Gamma f^{M*}(x-1) - (k+1)\sum_{z=1}^{[x/(k+1)]} f^{M*}(x-z(k+1))\right),$$
$$(x = 1, 2, \dots)$$

but in that case the recursion (5.22) is more efficient. □

Insertion of (5.16) in (6.5) gives

$$\tau_{\varphi_f}(s) = s\frac{\mathrm{d}}{\mathrm{d}s}\ln\tau_f(s) = s\frac{\tau_f'(s)}{\tau_f(s)}, \tag{6.9}$$

which defines the De Pril transform of f in terms of the generating function of f. From (1.34), the last expression in (6.9), and (1.20) we obtain

$$\tau_{\Phi f} = \Phi(\tau_f') = \tau_{\varphi_f}\tau_f = \tau_{\varphi_f * f},$$

which brings us back to (6.3).

From (4.29), we know that if p is infinitely divisible, then

$$\mu_{\varphi_p}(-1) = -\ln p(0). \tag{6.10}$$

We shall now show that this relation holds more generally.

Theorem 6.1 *If $p \in \mathcal{F}_{1\underline{0}}$ and $\mu_{|\varphi_p|}(-1) < \infty$, then*

$$\kappa_p(0) = \ln p(0) + \mu_{\varphi_p}(-1). \tag{6.11}$$

Proof By application of (1.27), (1.11), (5.16), (1.28), and Corollary 6.6, we obtain

$$\kappa_p(0) = \theta_p(0) = \ln\tau_p(1) = \ln\tau_p(0) + \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\ln\tau_p(s)\,\mathrm{d}s$$
$$= \ln p(0) + \int_0^1 \rho_p(s)\,\mathrm{d}s = \ln p(0) + \int_0^1\left(\sum_{n=1}^{\infty}\varphi_p(n)s^{n-1}\right)\mathrm{d}s$$
$$= \ln p(0) + \sum_{n=1}^{\infty}\frac{\varphi_p(n)}{n} = \ln p(0) + \mu_{\varphi_p}(-1).$$
□

We know that $\kappa_p(0) = 0$ when $p \in \mathcal{P}_{1\underline{0}}$. Thus, (6.10) is satisfied if in addition $\mu_{|\varphi_p|}(-1) < \infty$.

6.2 The $\mathcal{R}_k$ Classes

The recursion (4.23) is trivially extended to the case when we drop the assumption of infinite divisibility. Furthermore, because a distribution in $\mathcal{P}_{1\underline{0}}$ uniquely determines its De Pril transform, we also have that that recursion implies the recursion (4.22). Hence, we obtain the following theorem.

Theorem 6.2 *A distribution $f \in \mathcal{P}_{1\underline{0}}$ satisfies the recursion*

$$f(x) = \sum_{y=1}^{r} \left(\frac{\xi(y)}{x} + \left(1 - \frac{y}{x}\right)\chi(y)\right) f(x-y) \qquad (x = 1, 2, \dots) \tag{6.12}$$

for functions ξ and χ on $\{1, 2, \dots, r\}$ with r being a positive integer or infinity iff its De Pril transform satisfies the recursion

$$\varphi_f(x) = \xi(x) + \sum_{y=1}^{r} \chi(y)\varphi_f(x-y). \qquad (x = 1, 2, \dots) \tag{6.13}$$

By letting $f = p$, $r = k$, $\chi = a$, and $\xi = \Phi a + b$, we obtain the following corollary.

Corollary 6.8 *If p is $R_k[a, b]$ with k being a positive integer or infinity, then*

$$\varphi_p(n) = na(n) + b(n) + \sum_{i=1}^{k} a(i)\varphi_p(n-i). \qquad (n = 1, 2, \dots) \tag{6.14}$$

From Theorem 5.6, we know that p can be expressed as $R_\infty[-p/p(0), 2\Phi p/p(0)]$. Insertion of $a = -p/p(0)$ and $b = 2\Phi p/p(0)$ in (6.14) brings us back to the recursion (6.1).

By letting $k = \infty$, $a \equiv 0$, and $b = \varphi_p$ in (6.14), we get the obvious result $\varphi_p(n) = \varphi_p(n)$.

By combining Theorem 5.4 and Corollary 6.8, we obtain the following corollary.

Corollary 6.9 *If $f = p \vee h$ where p is $R_k[a, b]$ and $h \in \mathcal{P}_{10}$, then*

$$\varphi_f(x) = \frac{1}{1 - \tau_a(h(0))}\left(x((a + \Psi b) \vee h)(x) + \sum_{y=1}^{x-1} (a \vee h)(y)\varphi_f(x-y)\right).$$
$$(x = 1, 2, \dots) \tag{6.15}$$

Now let $k = 1$. In this case, (6.15) reduces to

$$\varphi_f(x) = \frac{1}{1 - ah(0)}\left((a + b)xh(x) + a\sum_{y=1}^{x-1} h(y)\varphi_f(x-y)\right). \tag{6.16}$$
$$(x = 1, 2, \dots)$$

In particular, with h concentrated in one, this gives

$$\varphi_p(n) = \begin{cases} a+b & (n=1) \\ a\varphi_p(n-1), & (n=2,3,\dots) \end{cases}$$

from which we obtain

$$\varphi_p(n) = (a+b)a^{n-1}. \qquad (n=1,2,\dots) \tag{6.17}$$

If $h \in \mathcal{P}_{11}$, then insertion in (6.4) gives

$$\varphi_f(x) = x(a+b)\sum_{n=1}^{x}\frac{a^{n-1}}{n}h^{n*}(x). \qquad (x=1,2,\dots) \tag{6.18}$$

Let us apply (6.16)–(6.18) to the three main classes in $\mathcal{R}_1$. The values of a and b are found in Table 2.1.

1. *Poisson distribution* $\mathrm{Po}(\lambda)$. In this case, both (6.16) and (6.18) give (4.13), and from (6.17), we obtain (4.17).
2. *Binomial distribution* $\mathrm{bin}(M,\pi)$. We obtain

$$\varphi_f(x) = \frac{\pi}{1-\pi+\pi h(0)}\left(Mxh(x) - \sum_{y=1}^{x-1}h(y)\varphi_f(x-y)\right) \tag{6.19}$$

$$(x=1,2,\dots)$$

$$\varphi_p(n) = -M\left(\frac{\pi}{\pi-1}\right)^n \qquad (n=1,2,\dots) \tag{6.20}$$

$$\varphi_f(x) = -Mx\sum_{n=1}^{x}\frac{1}{n}\left(\frac{\pi}{\pi-1}\right)^n h^{n*}(x). \qquad (x=1,2,\dots) \tag{6.21}$$

When $M=1$ and $h \in \mathcal{P}_{11}$, insertion of (2.55) and (2.56) in (6.19) brings us back to (6.1).

3. *Negative binomial distribution* $\mathrm{NB}(\alpha,\pi)$. We obtain

$$\varphi_f(x) = \frac{\pi}{1-\pi h(0)}\left(\alpha x h(x) + \sum_{y=1}^{x-1}h(y)\varphi_f(x-y)\right), \qquad (x=1,2,\dots)$$

which also follows from (4.27) when $h \in \mathcal{P}_{11}$, and

$$\varphi_p(n) = \alpha\pi^n \qquad (n=1,2,\dots)$$

$$\varphi_f(x) = \alpha x\sum_{y=1}^{x}\frac{\pi^n}{n}h^{n*}(x), \qquad (x=1,2,\dots)$$

which also follow from (4.20) and (4.21) respectively.

Further Remarks and References

Theorem 6.1 was proved by Sundt et al. (1998). Formula (6.10) was deduced by Dhaene and De Pril (1994).

Example 6.1 is based on Sundt (1999a).

Corollary 6.9 was proved by Sundt and Ekuma (1999).

Most of the remaining results in this chapter were presented by Sundt (1995).

Chapter 7
Individual Models

Summary

In Sect. 4.4.1, we discussed recursive evaluation of the aggregate claims distribution of an individual model where each policy had an infinitely divisible aggregate claims distribution in $\mathcal{P}_{1\underline{0}}$. In the present chapter, we extend this setting by dropping the assumption of infinite divisibility.

We start Sect. 7.1 with a recapitulation of De Pril's first method from Sect. 4.4.1, extended to our present setting. Then we present De Pril's second method within that setting. The difference between these two methods is that whereas in the first method, the De Pril transform of the aggregate claims distribution of each policy is evaluated recursively, we use a closed-form expression for this De Pril transform in the second method. Within the same setting, Sect. 7.2 is devoted to Dhaene–Vandebroek's method, which is more efficient than De Pril's methods in some situations.

The methods of De Pril and Dhaene–Vandebroek are sometimes presented in a two-way model referred to as De Pril's individual model. This model is discussed in Sect. 7.3, and, in addition to those methods, we introduce two other recursive methods.

Section 7.4 is devoted to collective approximations to the individual model, and we also discuss a semi-collective model where "normal" policies are treated collectively and some "special" policies individually.

7.1 De Pril's Methods

Let us assume that we want to evaluate $f = *_{j=1}^{M} f_j$ with $f_1, f_2, \dots, f_M \in \mathcal{P}_{1\underline{0}}$. As outlined in Sect. 4.4.1, in De Pril's first method, we first evaluate the De Pril transform of each f_j by (6.1), then we find the De Pril transform of f by summing these De Pril transforms, and finally we evaluate f by (6.2); we can express these last two steps by

$$f(x) = \frac{1}{x} \sum_{y=1}^{x} f(x-y) \sum_{j=1}^{M} \varphi_{f_j}(y). \qquad (x = 1, 2, \dots) \tag{7.1}$$

In De Pril's *second method*, we evaluate the De Pril transform of each f_j from a closed-form expression instead of using the recursion (6.1). To derive this expression, we express f_j as a compound Bernoulli distribution $p_j \vee h_j$ where p_j is

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_7, © Springer-Verlag Berlin Heidelberg 2009

the Bernoulli distribution Bern(π_j) with $\pi_j = 1 - f_j(0)$ and severity distribution $h_j \in \mathcal{P}_{11}$ given by $h_j(y) = f_j(y)/\pi_j$ for $y = 1, 2, \ldots$. From (6.21), we obtain that

$$\varphi_{f_j}(x) = -x \sum_{n=1}^{x} \frac{1}{n} \left(\frac{\pi_j}{\pi_j - 1} \right)^n h_j^{n*}(x). \qquad (x = 1, 2, \ldots)$$

Application of Corollary 6.1 gives

$$\varphi_f(x) = -x \sum_{n=1}^{x} \frac{1}{n} \sum_{j=1}^{M} \left(\frac{\pi_j}{\pi_j - 1} \right)^n h_j^{n*}(x), \qquad (x = 1, 2, \ldots) \tag{7.2}$$

which should be inserted in (6.2).

If

$$s_j = \max\{x : f_j(x) > 0\} < \infty, \qquad (j = 1, 2, \ldots, M)$$

then

$$s = \max\{x : f(x) > 0\} = \sum_{j=1}^{M} s_j,$$

and we can evaluate f by a backward recursion similar to Theorem 2.9. Let $X = X_{\bullet M}$ where $X_1, X_2, \ldots, X_M$ are independent random variables, X_j with distribution f_j for $j = 1, 2, \ldots, M$. For $j = 1, 2, \ldots, M$, we introduce $\tilde{X}_j = s_j - X_j$. Its distribution $\tilde{f}_j$ belongs to $\mathcal{P}_{1\underline{0}}$ and is given by

$$\tilde{f}_j(x) = f_j(s_j - x)$$

for $x = 0, 1, 2, \ldots$. The distribution of $\tilde{X} = \tilde{X}_{\bullet M} = s - X$ is $\tilde{f} = *_{j=1}^{M} \tilde{f}_j$ and can be expressed by

$$\tilde{f}(x) = f(s - x)$$

for $x = 0, 1, 2, \ldots$. We can evaluate $\tilde{f}$ by the recursion of the previous paragraph. This gives a backwards recursion for f.

If each h_j is concentrated in a positive integer s_j, then h_j^{n*} is concentrated in ns_j. Then (7.2) can be written as

$$\varphi_f(x) = -\sum_{j=1}^{M} s_j \left(\frac{\pi_j}{\pi_j - 1} \right)^{x/s_j} I(x = [x/s_j]s_j > 0), \quad (x = 1, 2, \ldots) \tag{7.3}$$

and insertion in (6.2) gives

$$f(x) = -\frac{1}{x} \sum_{j=1}^{M} s_j \sum_{n=1}^{[x/s_j]} \left(\frac{\pi_j}{\pi_j - 1} \right)^n f(x - ns_j). \qquad (x = 1, 2, \ldots) \tag{7.4}$$

If we let $s_j = 1$ for $j = 1, 2, \dots, M$, then f becomes the distribution of the number of policies with claims, and we obtain

$$\varphi_f(x) = -\sum_{j=1}^{M}\left(\frac{\pi_j}{\pi_j - 1}\right)^x; \qquad f(x) = -\frac{1}{x}\sum_{n=1}^{x} f(x-n)\sum_{j=1}^{M}\left(\frac{\pi_j}{\pi_j - 1}\right)^n.$$
$$(x = 1, 2, \dots)$$

For the recursion for $\tilde{f}$ in the previous paragraph, we use the same recursion as for f with π_j replaced with $1 - \pi_j$ for $j = 1, 2, \dots, M$.

7.2 Dhaene–Vandebroek's Method

As we shall see from the following results, it can sometimes be useful to express (7.1) as

$$f(x) = \frac{1}{x}\sum_{j=1}^{M}\sigma_j(x) \qquad (x = 1, 2, \dots) \tag{7.5}$$

with

$$\sigma_j(x) = \sum_{y=1}^{x}\varphi_{f_j}(y) f(x-y). \qquad (x = 1, 2, \dots;\ j = 1, 2, \dots, M) \tag{7.6}$$

Theorem 7.1 *Let* $f_1, f_2, \dots, f_M \in \mathcal{P}_{1\underline{0}}$ *and* $f = *_{j=1}^{M} f_j$. *If* f_j *is* $R_k[a, b]$ *for some* j, *then*

$$\sigma_j(x) = \sum_{y=1}^{k}((ya(y) + b(y)) f(x-y) + a(y)\sigma_j(x-y)). \tag{7.7}$$
$$(x = 1, 2, \dots)$$

Proof By starting with (7.6) and successive application of Corollary 6.8 and (7.6), we obtain

$$\begin{aligned}
\sigma_j(x) &= \sum_{y=1}^{x}\varphi_{f_j}(y) f(x-y) \\
&= \sum_{y=1}^{x}\left(ya(y) + b(y) + \sum_{i=1}^{k} a(i)\varphi_{f_j}(y-i)\right) f(x-y) \\
&= \sum_{y=1}^{k}(ya(y) + b(y)) f(x-y) + \sum_{i=1}^{k} a(i)\sum_{z=1}^{x-i}\varphi_{f_j}(z) f(x-i-z)
\end{aligned}$$

$$= \sum_{y=1}^{k}(ya(y)+b(y))f(x-y)+\sum_{i=1}^{k}a(i)\sigma_j(x-i)$$

$$= \sum_{y=1}^{k}((ya(y)+b(y))f(x-y)+a(y)\sigma_j(x-y)).$$

□

With $k=\infty$, $a=0$, and $b=\varphi_{f_j}$, (7.7) brings us back to the definition (7.6). By inserting (5.10) in (7.7), we obtain the following corollary.

Corollary 7.1 *Let $f_1, f_2, \dots, f_M \in \mathcal{P}_{1\underline{0}}$ and $f = *_{j=1}^M f_j$. If f_j is a compound distribution with counting distribution $R_k[a,b]$ and severity distribution $h \in \mathcal{P}_{10}$ for some j, then*

$$\sigma_j(x) = \frac{1}{1-\tau_a(h(0))}\sum_{y=1}^{x}(y((a+\Psi b)\vee h)(y)f(x-y)+(a\vee h)(y)\sigma_j(x-y)).$$

$$(x = 1,2,\dots) \tag{7.8}$$

The following corollary is obtained by application of Theorem 5.6 in (7.7).

Corollary 7.2 *Let $f_1, f_2, \dots, f_M \in \mathcal{P}_{1\underline{0}}$ and $f = *_{j=1}^M f_j$. Then*

$$\sigma_j(x) = \frac{1}{f_j(0)}\sum_{y=1}^{x}(yf(x-y)-\sigma_j(x-y))f_j(y). \tag{7.9}$$

$$(x = 1,2,\dots;\ j=1,2,\dots,M)$$

We see that by evaluating the σ_js recursively by (7.9), we easily find $f(x)$ by (7.5) instead of using the recursion (6.2); this is *Dhaene–Vandebroek's method*. On the other hand, the recursion (7.9) is somewhat more complicated than (6.1).

In Sect. 4.4.1, we pointed out that when we know the De Pril transforms of the aggregate claims distributions of the various types of policies, then it is easy to evaluate the aggregate claims distribution of the portfolio under changes of the composition of the portfolio. As the σ_js depend on the aggregate claims distribution of the portfolio, and, hence, the composition of the portfolio, Dhaene–Vandebroek's method is less convenient for studying changes in the composition of the portfolio than De Pril's methods.

7.3 De Pril's Individual Model

The methods of De Pril and Dhaene–Vandebroek are sometimes presented in a two-way model for an insurance portfolio of independent policies where each policy can have at most one claim during the period under consideration. It is assumed that there are classes (i, j) with $i = 1, 2, \ldots, I$ and $j = 1, 2, \ldots, J$. In class (i, j), there are M_{ij} policies and each policy has claim probability $\pi_j \in (0, 1)$ and severity distribution $h_i \in \mathcal{P}_{11}$, that is, the aggregate claims distribution for a policy in this class is $f_{ij} = p_j \vee h_i$ where p_j is the Bernoulli distribution $\mathrm{Bern}(\pi_j)$. Then we have $f_{ij}(0) = 1 - \pi_j$ and $f_{ij}(x) = \pi_j h_i(x)$ for $x = 1, 2, \ldots$. We shall call this model *De Pril's individual model*.

With De Pril's first method, we first evaluate the De Pril transform of each f_{ij} by (6.1), then the De Pril transform of the aggregate claims distribution $f = *_{i=1}^{I} *_{j=1}^{J} f_{ij}^{M_{ij}*}$ by $\varphi_f = \sum_{i=1}^{I} \sum_{j=1}^{J} M_{ij} \varphi_{f_{ij}}$, and finally f by (6.2). With Dhaene–Vandebroek's method, for $x = 1, 2, \ldots$, we first evaluate

$$\sigma_{ij}(x) = \sum_{y=1}^{x} M_{ij} \varphi_{f_{ij}}(y) f(x-y)$$

by the recursion (7.9) for each cell (i, j), and then we find $f(x)$ by

$$f(x) = \frac{1}{x} \sum_{i=1}^{I} \sum_{j=1}^{J} M_{ij} \sigma_{ij}(x). \tag{7.10}$$

For De Pril's second method, (7.2) gives

$$\varphi_f(x) = -x \sum_{n=1}^{x} \frac{1}{n} \sum_{j=1}^{J} \left(\frac{\pi_j}{\pi_j - 1} \right)^n \sum_{i=1}^{I} M_{ij} h_i^{n*}(x). \quad (x = 1, 2, \ldots) \tag{7.11}$$

We see that for De Pril's second method, the knowledge of the two-way structure gives a computational advantage whereas this is not the case for the two other methods. On the other hand, the need for all the convolutions in (7.11) will often make De Pril's second method rather inefficient in practice. However, this method has been used as basis for development of approximations based on approximating each p_j with a function $p_j^{(r)} \in \mathcal{F}_{1\underline{0}}$ with $\varphi_{p_j^{(r)}}(n) = 0$ for each n greater than some integer r. Then we are in the unusual situation of finding it efficient to approximate one of the simplest of all types of non-degenerate distributions, the Bernoulli distribution, with a much more complex function. We shall study such approximations in Chap. 10.

In the individual life model with h_i concentrated in a positive integer s_i for $i = 1, 2, \ldots, I$, (7.3) and (7.4) give that for $x = 1, 2, \ldots$,

$$\varphi_f(x) = -\sum_{i=1}^{I} s_i I(x = [x/s_i] s_i > 0) \sum_{j=1}^{J} M_{ij} \left(\frac{\pi_j}{\pi_j - 1} \right)^{x/s_i}$$

$$f(x) = -\frac{1}{x} \sum_{i=1}^{I} s_i \sum_{n=1}^{[x/s_i]} f(x - n s_i) \sum_{j=1}^{J} M_{ij} \left(\frac{\pi_j}{\pi_j - 1} \right)^n .$$

Returning to general severity distributions in $\mathcal{P}_{11}$, let us now look at yet another method for evaluating f. For each cell (i, j), the aggregate claims distribution can be expressed as a compound binomial distribution with counting distribution $\text{bin}(M_{ij}, \pi_j)$ and severity distribution h_i. Hence, the aggregate claims distribution f_i for all the cells (i, j) for fixed i is a compound distribution f_i with severity distribution h_i and counting distribution p_i being the convolution of $\text{bin}(M_{i1}, \pi_1), \text{bin}(M_{i2}, \pi_2), \dots, \text{bin}(M_{iJ}, \pi_J)$. As each of these binomial distributions is in $\mathcal{R}_1$, Theorem 5.8 gives that their convolution p_i can be represented in the form $R_J[a_i, b_i]$. From Table 2.1, we obtain that for each j, $\text{bin}(M_j, \pi_j)$ is $R_1[-\pi_j/(1-\pi_j), (M_j+1)\pi_j/(1-\pi_j)]$, and insertion in Corollary 5.8 gives the following corollary.

Corollary 7.3 *Let* $R_k[\alpha_k, \beta_k]$ *be the convolution of the binomial distributions* $\text{bin}(M_1, \pi_1), \text{bin}(M_2, \pi_2), \dots, \text{bin}(M_k, \pi_k)$ *for* $k = 1, 2, \dots$. *Then the functions* α_k *and* β_k *can be evaluated recursively by*

$$\alpha_k(u) = \begin{cases} \alpha_{k-1}(u) + \frac{\pi_k}{1-\pi_k}\alpha_{k-1}(u-1) & (u = 2, 3, \dots, k) \\ \\ \alpha_{k-1}(1) - \frac{\pi_k}{1-\pi_k} & (u = 1) \end{cases}$$

$$\beta_k(u) = \begin{cases} \beta_{k-1}(u) + \frac{\pi_k}{1-\pi_k}(\beta_{k-1}(u-1) - (M_k+1)\alpha_{k-1}(u-1)) \\ \qquad\qquad\qquad\qquad\qquad\qquad (u = 2, 3, \dots, k) \\ \beta_{k-1}(1) + (M_k+1)\frac{\pi_k}{1-\pi_k} \quad (u = 1) \end{cases}$$

for $k = 2, 3, \dots$ *with*

$$\alpha_1(1) = -\frac{\pi_1}{1-\pi_1}; \qquad \beta_1(1) = (M_1+1)\frac{\pi_1}{1-\pi_1}. \tag{7.12}$$

When we have found a_i and b_i by this procedure, then we can evaluate f_i recursively by (5.13), and finally we evaluate the convolution $f = *_{i=1}^{I} f_i$ by brute force.

We could also insert a_i and b_i in (7.8) and evaluate f by (7.10).

When the De Pril transform was introduced, some people argued that there was no need for introducing it as what one could show by De Pril transforms, could also be shown by generating functions. Against this criticism, it should be emphasised that whereas the generating function is primarily a tool for proving analytical results, the De Pril transform is primarily a tool for evaluation of e.g. compound distributions and convolutions. Although, as we have seen, it can also be applied to prove analytical results, that is not its primary purpose.

7.4 Collective Approximations

We have now considered various methods for recursive exact evaluation of the aggregate claims distribution in an individual model. In this model, the aggregate claims distribution is a convolution. In a collective model, the aggregate claims distribution is a compound distribution, and we have developed recursive methods for compound distributions under various assumptions. These methods often seem much simpler than the methods we have developed for the individual model. One of the simplest cases is when the counting distribution of the compound distribution is a Poisson distribution. We shall now consider approximating the aggregate claims distribution of the individual model in the setting of Sect. 7.1 with a compound Poisson distribution.

Let us consider each f_j as a compound Bernoulli distribution like we did in connection with De Pril's second method, and approximate the Bernoulli counting distribution p_j with a Poisson distribution q_j with parameter equal to the Bernoulli parameter $\pi_j = 1 - f_j(0)$. Thus, we approximate $f_j = p_j \vee h_j$ with $g_j = q_j \vee h_j$ and the aggregate claims distribution $f = *_{j=1}^{M} f_j$ with $g = *_{j=1}^{M} g_j$. From Theorem 4.4, we obtain that $g = q \vee h$ where q is the Poisson distribution $\mathrm{Po}(\lambda)$ with $\lambda = \pi_{\bullet M}$ and

$$h = \frac{1}{\lambda}\sum_{j=1}^{M} \pi_j h_j.$$

Thus, we can evaluate g recursively by Theorem 2.2.

For $x = 1, 2, \ldots$, we have

$$h(x) = \frac{1}{\lambda}\sum_{j=1}^{M} \pi_j h_j(x) = \frac{\sum_{j=1}^{M} \pi_j h_j(x)}{\sum_{j=1}^{M} \pi_j} = \frac{\sum_{j=1}^{M} f_j(x)}{\sum_{j=1}^{M}(1 - f_j(0))} = \frac{\tilde{f}(x)}{1 - \tilde{f}(0)}$$

with

$$\tilde{f} = \frac{1}{M}\sum_{j=1}^{M} f_j,$$

and we have $\lambda = M(1 - \tilde{f}(0))$, that is, the approximation g depends on $f_1, f_2, \ldots, f_M$ only through $\tilde{f}$ and M. Thus, the transition from f to g can be considered as a two-step operation. In the first step, we approximate each f_j with $\tilde{f}$, which implies that f is approximated with $\tilde{f}^{M*}$, and in the second step, we approximate that aggregate claims distribution with a compound Poisson distribution. When considering the transition to the compound Poisson distribution like that, it seems likely that we would get a better approximation by only carrying through the first step. Then we can evaluate $\tilde{f}^{M*}$ by the recursion of Theorem 2.8. This approximation is sometimes called the *natural approximation*.

The distribution $\tilde{f}$ can be interpreted as a mixed distribution. To simplify the evaluation of the aggregate claims distribution in an insurance setting, we consider the individual risk characteristics of each individual policy as random; we incorporate these characteristics in an abstract random risk parameter and consider these parameters of different policies as independent and identically distributed. We estimate their distribution by the empirical distribution of the distributions f_j within the portfolio. In this setting, $\tilde{f}$ becomes the unconditional aggregate claims distribution of an arbitrary policy from the portfolio.

Our compound approximation g of f can be interpreted as if we approximate the distribution $p = *_{j=1}^{M} p_j$ of the number of policies with claims by q. Furthermore, it is assumed that the aggregate claims of policies with claims are independent of the number of such policies and mutually independent and identically distributed with distribution h. Although the total number of policies is M, the approximation q gives positive probability of having more than M policies with claims. If we assume that each policy can have at most one claim, then q can be interpreted as the claim number distribution and h as the claim amount distribution.

We shall show that $\Pi_f \leq \Pi_g$, that is, our approximation gives an upper bound to the exact stop loss transform. We shall prove this result for a more general class of distributions. First we consider the case $M = 1$.

Lemma 7.1 *Let H be a univariate distribution*, $0 < \pi < 1$, *p the Bernoulli distribution* $\mathrm{Bern}(\pi)$, *and q the Poisson distribution* $\mathrm{Po}(\pi)$. *Then* $\Pi_{q \vee H} \geq \Pi_{p \vee H}$.

Proof Application of Theorem 1.7 gives

$$\Pi_{q \vee H} \geq \mu_q(1)\Pi_H = \pi \Pi_H = \Pi_{p \vee H}. \qquad \square$$

We now turn to the general case.

Theorem 7.2 *For $j = 1, 2, \ldots, M$, let H_j be a univariate distribution, p_j the Bernoulli distribution* $\mathrm{Bern}(\pi_j)$ *with* $0 < \pi_j < 1$, *and q_j the Poisson distribution* $\mathrm{Po}(\pi_j)$. *Then* $\Pi_{q \vee H} \geq \Pi_{*_{j=1}^{M}(p_j \vee H_j)}$ *where q is the Poisson distribution* $\mathrm{Po}(\lambda)$ *with* $\lambda = \sum_{j=1}^{M} \pi_j$ *and*

$$H = \frac{1}{\lambda} \sum_{j=1}^{M} \pi_j H_j.$$

Proof By application of Theorem 4.4, Lemma 7.1, and Theorem 1.5, we obtain

$$\Pi_{q \vee H} = \Pi_{*_{j=1}^{M}(q_j \vee H_j)} \geq \Pi_{*_{j=1}^{M}(p_j \vee H_j)}. \qquad \square$$

As a special case, we obtain that $\Pi_f \leq \Pi_g$. We do not have such a result for the natural approximation $\tilde{f}^{M*}$.

Now let us consider a mixed setting with mixing variable Θ with distribution U. Conditionally given that $\Theta = \theta$, we have the same assumptions as in Theorem 7.2 apart from replacing π_j with $\theta \pi_j$ for $j = 1, 2, \ldots, M$. Letting now

$$q(n) = \int_{(0,\infty)} \frac{(\theta\lambda)^n}{n!} \mathrm{e}^{-\theta\lambda} \, \mathrm{d}U(\theta), \qquad (n = 0, 1, 2, \ldots)$$

we have that $\Pi_{q \vee h}$ gives an upper bound of the stop loss transform of the exact unconditional distribution. In an insurance setting, we can interpret Θ as representing unknown random conditions that will affect the risk of all policies in the portfolio, e.g. windstorms. This is a way of modelling dependence between policies within the individual model.

Let us now return to the original setting of our individual model. We have considered various ways of exact evaluation as well as a collective compound Poisson approximation for evaluation of the aggregate claims distribution in this model. If we feel that the collective approximation is not sufficiently accurate, then we could apply a *semi-collective* approximation where we use the collective approximation on a part of the portfolio and the individual model on the rest. The approximation might work well on the normal policies, but it might improve the accuracy significantly to apply exact evaluation on some special policies. But which of the policies should we consider as special? To answer that, we need a measure of the quality of the approximation.

Let $A \subseteq \{1, 2, \dots, M\}$ denote the normal policies. The idea is then to approximate f by $f_A = g_A * (*_{j \notin A} f_j)$ with

$$g_A = *_{j \in A}(q_j \vee h_j) = q_A \vee h_A$$

where q_A is the Poisson distribution $\mathrm{Po}(\lambda_A)$ with $\lambda_A = \sum_{j \in A} \pi_j$ and the severity distribution h_A is given by

$$h_A = \frac{1}{\lambda_A} \sum_{j \in A} \pi_j h_j.$$

We then have $\Pi_f \leq \Pi_{f_A} \leq \Pi_g$. As $\mu_{p_j}(1) = \mu_{q_j}(1) = \pi_j$ for $j = 1, 2, \dots$, (1.32) gives that

$$\mu_f(1) = \mu_{f_A}(1) = \mu_g(1), \tag{7.13}$$

that is, $\Pi_f(0) = \Pi_{f_A}(0) = \Pi_g(0)$.

Application of (4.13) gives that for $x = 1, 2, \dots$,

$$\varphi_{g_A}(x) = \lambda_A x h_A(x) = x \sum_{j \in A} \pi_j h_j(x) = x \sum_{j \in A} f_j(x),$$

and by Corollary 6.1 we obtain

$$\begin{aligned}\varphi_{f_A}(x) &= \varphi_{g_A}(x) + \sum_{j \notin A} \varphi_{f_j}(x) = \lambda_A \Phi h_A(x) + \sum_{j \notin A} \varphi_{f_j}(x) \\ &= x \sum_{j \in A} f_j(x) + \sum_{j \notin A} \varphi_{f_j}(x) = \sum_{j \in A} \Phi f_j(x) + \sum_{j \notin A} \varphi_{f_j}(x).\end{aligned}$$

We see that the compound Poisson approximation of an f_j actually consists of approximating $\varphi_{f_j}(x)$ with $\Phi f_j(x)$ for $x = 1, 2, \dots$.

As a measure of the quality of the approximation f_A compared with the approximation g, we introduce the efficiency

$$e_A = \frac{\sum_{x=0}^{\infty}(\Pi_g - \Pi_{f_A})(x)}{\sum_{x=0}^{\infty}(\Pi_g - \Pi_f)(x)}. \tag{7.14}$$

In particular, we have $e_{\{1,2,\dots,M\}} = 0$ and $e_\emptyset = 1$.

Application of (1.47) and (1.41) gives

$$\begin{aligned}\sum_{x=0}^{\infty}(\Pi_g - \Pi_{f_A})(x) &= \sum_{x=0}^{\infty}\Lambda^2(g - f_A)(x-1) = \Lambda^3(g - f_A)(-2)\\ &= \sum_{y=1}^{\infty}\binom{y+1}{2}(g - f_A)(y) = \frac{1}{2}\sum_{y=1}^{\infty}(y^2 + y)(g - f_A)(y)\\ &= \frac{1}{2}(\mu_g(2) + \mu_g(1) - \mu_{f_A}(2) - \mu_{f_A}(1)).\end{aligned}$$

By (7.13), we obtain

$$\begin{aligned}2\sum_{x=0}^{\infty}(\Pi_g(x) - \Pi_{f_A}(x)) &= \mu_g(2) - \mu_{f_A}(2)\\ &= \mu_g(2) - \mu_g^2(1) - \mu_{f_A}(2) + \mu_{f_A}^2(1) = \kappa_g(2) - \kappa_{f_A}(2)\\ &= \sum_{j\notin A}(\kappa_{g_j}(2) - \kappa_{f_j}(2)) = \sum_{j\notin A}(\pi_j\mu_{h_j}(2) - \kappa_{f_j}(2))\\ &= \sum_{j\notin A}(\mu_{f_j}(2) - \kappa_{f_j}(2)) = \sum_{j\notin A}\mu_{f_j}^2(1),\end{aligned}$$

and insertion in (7.14) gives

$$e_A = \frac{\sum_{j\notin A}\mu_{f_j}^2(1)}{\sum_{j=1}^{M}\mu_{f_j}^2(1)}.$$

We see that under our present efficiency criterion, it is the f_js with the largest mean that we ought to treat individually.

Let us now compare our approximations with exact evaluation in a numerical example.

Example 7.1 We consider De Pril's individual model on a dataset that has often been applied in the literature. We have $I = 5$ and $J = 4$. For $i = 1, 2, \dots, I$, the severity distribution is concentrated in i. The π_js and the M_{ij}s are given in Table 7.1.

In Fig. 7.1, we display the efficiency e_A of the semi-collective approximation as a function of n when treating the n policies with largest mean individually. We clearly see that the effect of removing one more policy from collective treatment decreases

Table 7.1 Dataset

i		1	2	3	4	5
j	π_j			M_{ij}		
1	0.03	2	3	1	2	0
2	0.04	0	1	2	2	1
3	0.05	0	2	4	2	2
4	0.06	0	2	2	2	1

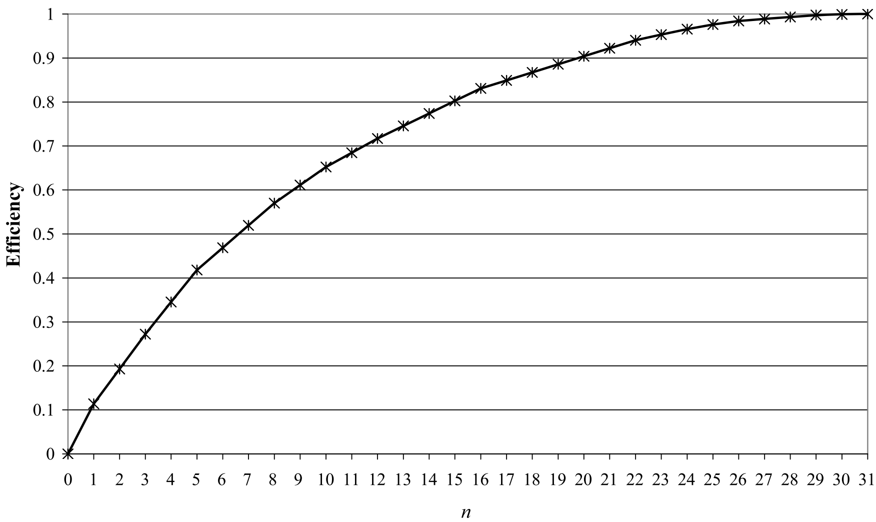

Fig. 7.1 Efficiency of the semi-collective approximation when treating the n policies with largest mean individually

in the number of policies already removed; the policy with the highest mean (0.3) gives a contribution of 0.114 to the efficiency whereas the one with the lowest mean (0.03) contributes with 0.00114.

In Table 7.2, we compare the exact aggregate claims distribution f of the portfolio with the natural approximation and the collective compound Poisson approximation, as well as the semi-collective approximation with the five policies with largest mean treated individually; that gives an efficiency of 0.418. Table 7.3 gives an analogous comparison in terms of stop loss transforms. □

Further Remarks and References

De Pril's second method for evaluation of the convolution of M distributions in $\mathcal{P}_{\underline{10}}$ was first presented by White and Greville (1959) for evaluation of the distribution

Table 7.2 The exact distribution and some approximations

x	Exact	Semi-collective	Collective	Natural
0	2.38195E-01	2.44580E-01	2.46597E-01	2.38688E-01
1	1.47337E-02	1.46748E-02	1.47958E-02	1.49986E-02
2	8.77342E-02	8.60432E-02	8.67528E-02	8.79481E-02
3	1.13183E-01	1.10314E-01	1.11224E-01	1.12820E-01
4	1.10709E-01	1.11367E-01	1.10397E-01	1.12203E-01
5	9.63274E-02	9.44358E-02	9.28589E-02	9.47052E-02
6	6.15487E-02	6.13015E-02	6.10080E-02	6.25913E-02
7	6.90221E-02	6.65192E-02	6.54270E-02	6.70024E-02
8	5.48171E-02	5.49829E-02	5.45768E-02	5.56748E-02
9	4.31471E-02	4.26416E-02	4.13208E-02	4.18689E-02
10	3.01073E-02	3.01176E-02	3.05794E-02	3.06936E-02
11	2.35291E-02	2.36543E-02	2.33078E-02	2.31499E-02
12	1.82824E-02	1.83879E-02	1.83438E-02	1.80376E-02
13	1.25093E-02	1.29504E-02	1.31494E-02	1.27325E-02
14	8.71076E-03	8.99017E-03	9.21800E-03	8.75461E-03
15	5.91165E-03	6.25302E-03	6.50426E-03	6.05269E-03
16	4.15190E-03	4.41542E-03	4.59553E-03	4.19105E-03
17	2.71505E-03	2.96194E-03	3.17641E-03	2.83267E-03
18	1.74094E-03	1.93839E-03	2.12340E-03	1.84149E-03
19	1.11736E-03	1.26792E-03	1.41386E-03	1.18991E-03
20	7.11015E-04	8.28029E-04	9.39530E-04	7.67248E-04
30	3.09434E-06	5.31779E-06	8.63294E-06	4.57655E-06
40	3.53514E-09	1.28654E-08	3.64155E-08	9.89289E-09

of the number of policies with claims. Within the framework of De Pril's individual model, De Pril's first and second method were deduced for the individual life model by De Pril (1986b) and extended to severity distributions in $\mathcal{P}_{11}$ by De Pril (1989). Dufresne (1996) considered the case where each severity distribution is distributed on two points.

The backward recursion in Sect. 7.1 was presented by De Pril (1986b) within the context of the individual life model.

Dhaene–Vandebroek's method was introduced for the individual life model by Waldmann (1994) and extended to general distributions in $\mathcal{P}_{1\underline{0}}$ by Dhaene and Vandebroek (1995).

Sundt and Vernic (2006) introduced the last two methods we presented for evaluation of the aggregate claims distribution in De Pril's individual model. They also compared these methods with other methods by counting the number of dot operations. Such comparisons of De Pril's methods with other methods have also been presented by Kuon et al. (1987), Waldmann (1994), Dhaene and Vandebroek (1995),

Table 7.3 The exact stop loss transform and some approximations

x	Exact	Semi-collective	Collective	Natural
0	4.49000E+00	4.49000E+00	4.49000E+00	4.49000E+00
1	3.72819E+00	3.73458E+00	3.73660E+00	3.72869E+00
2	2.98112E+00	2.99383E+00	2.99799E+00	2.98237E+00
3	2.32179E+00	2.33913E+00	2.34614E+00	2.32401E+00
4	1.77563E+00	1.79475E+00	1.80551E+00	1.77846E+00
5	1.34019E+00	1.36172E+00	1.37527E+00	1.34512E+00
6	1.00107E+00	1.02314E+00	1.03790E+00	1.00648E+00
7	7.23501E-01	7.45857E-01	7.61530E-01	7.30437E-01
8	5.14954E-01	5.35093E-01	5.50590E-01	5.21393E-01
9	3.61224E-01	3.79312E-01	3.94228E-01	3.68024E-01
10	2.50642E-01	2.66173E-01	2.79186E-01	2.56524E-01
11	1.70166E-01	1.83151E-01	1.94723E-01	1.75717E-01
12	1.13220E-01	1.23783E-01	1.33568E-01	1.18061E-01
13	7.45566E-02	8.28035E-02	9.07573E-02	7.84415E-02
14	4.84022E-02	5.47743E-02	6.10958E-02	5.15549E-02
15	3.09585E-02	3.57352E-02	4.06522E-02	3.34229E-02
16	1.94265E-02	2.29492E-02	2.67130E-02	2.13437E-02
17	1.20464E-02	1.45786E-02	1.73693E-02	1.34554E-02
18	7.38134E-03	9.16997E-03	1.12019E-02	8.39986E-03
19	4.45721E-03	5.69970E-03	7.15801E-03	5.18578E-03
20	2.65044E-03	3.49735E-03	4.52794E-03	3.16162E-03
30	7.25353E-06	1.53233E-05	2.97954E-05	1.27278E-05
40	5.72551E-09	2.86340E-08	1.01131E-07	2.10815E-08

Sundt and Dickson (2000), Dickson and Sundt (2001), Dhaene et al. (2006), and, in a bivariate setting, Walhin and Paris (2001c), some of them also counting bar operations.

Ribas et al. (2003) extended the individual life model by allowing for pairwise dependence of policies within the portfolio and deduced Dhaene–Vandebroek's method and De Pril's second method within that setting. Cossette et al. (2002) discussed recursive evaluation of the aggregate claims distribution in another extension of the individual model giving dependence between policies.

The natural approximation to the individual model was discussed by Sundt (1985). Jewell and Sundt (1981) approximated the individual model by compound mixed binomial distributions and compound zero-modified binomial distributions.

Lemma 7.1 and Theorem 7.2 are based on Bühlmann et al. (1977). Error bounds for compound Poisson approximations to an individual model of independent policies are also studied by Gerber (1984), Hipp (1985, 1986), Michel (1987), Hipp and Michel (1990), and De Pril and Dhaene (1992). Goovaerts and Dhaene (1996) extended an error bound of Michel (1987) to a collective Poisson approximation of an individual model with dependencies between policies. In the non-actuarial literature, compound Poisson approximations have a long history; an overview is given by Barbour and Chryssaphinou (2001).

The discussion on treating some policies collectively and others individually is based on Kaas et al. (1988a), who restricted to the individual life model. This subject is also studied by Kaas et al. (1988b, 1989).

The dataset that we applied in Example 7.1, was introduced by Gerber (1979) and has later been applied by Jewell and Sundt (1981), Chan (1984), Sundt (1985), Hipp (1986), Vandebroek and De Pril (1988), Kuon et al. (1993), and Dhaene and Goovaerts (1997).

Most of the remaining results in this chapter are based on Sundt (1995).

Chapter 8
Cumulative Functions and Tails

Summary

The present chapter is devoted to recursions for $\Gamma^t p$ and $\Lambda^t p$ with $p \in \mathcal{P}_{10}$ and $t = 0, 1, 2, \dots$. In some cases, we present the recursions more generally for $p \in \mathcal{F}_{10}$ when the proof does not depend on properties specific for distributions.

In Sect. 8.1, we deduce recursions when p satisfies the recursion (5.6). A short Sect. 8.2 treats recursive evaluation of $\Gamma^t p^{M*}$, and Sect. 8.3 is devoted to compound distributions. In Sect. 8.4, we study De Pril transforms, that is, the special case of (5.6) with $q = a \equiv 0$ and $b = \varphi_p$. Finally, in Sect. 8.5, we turn to the special case of (5.6) with $b \equiv 0$.

8.1 General Results

In the preceding chapters, we have discussed algorithms for recursive evaluation of probability functions of distributions in $\mathcal{P}_{10}$ under various assumptions. However, there are situations where we would be interested in other representations of the distribution than the probability function, for instance, the cumulative distribution function or the tail. In the present chapter, we shall consider recursions for $\Gamma^t p$ and $\Lambda^t p$ for $p \in \mathcal{P}_{10}$ and $t = 0, 1, 2, \dots$.

Any recursion for p can be used for recursive evaluation of $\Gamma^t p$ and $\Lambda^t p$ by adding the recursions

$$\Gamma^t p(n) = \Gamma^t p(n-1) + \Gamma^{t-1} p(n) \tag{8.1}$$

$$\Lambda^t p(n) = \Lambda^t p(n-1) - \Lambda^{t-1} p(n) \tag{8.2}$$

for $t = 1, 2, \dots$, starting with $n = 0$ or $n = -1$. However, in the following, we shall deduce direct recursions from recursions we have for p, and similar to these. One advantage with these recursions compared to the corresponding recursions for p is that for $p > 0$, $\Gamma^t p$ and $\Lambda^t p$ will be monotonic. That is an advantage in connection with numerical stability. As the proofs of the recursions hold more generally than for distributions in $\mathcal{P}_{10}$, we shall state the recursions for functions in $\mathcal{F}_{10}$.

For the recursion (8.1), we have the initial value $\Gamma^t p(0) = p(0)$ or $\Gamma^t p(-1) = 0$. For (8.2), Theorem 1.3 gives $\Lambda^t p(-1) = \nu_p(t-1)/(t-1)!$ for $t = 1, 2, \dots$.

It seems that the most interesting applications of recursions for $\Gamma^t p$ and $\Lambda^t p$ are evaluation of the cumulative distribution function, tail, stop loss transform, and retention transform of p, so that we are interested in $\Gamma^t p$ and $\Lambda^t p$ mainly for $t = 0, 1, 2$.

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_8, © Springer-Verlag Berlin Heidelberg 2009

We shall now deduce recursions for $\Gamma^t p$ and $\Lambda^t p$ with $p \in \mathcal{F}_{10}$ satisfying the recursion (5.6) and $t = 0, 1, 2, \ldots$. We start with $\Gamma^t p(n)$.

Theorem 8.1 *If $p \in \mathcal{F}_{10}$ satisfies the recursion (5.6) where k is a positive integer or infinity, then, for $t = 0, 1, 2, \ldots$, $\Gamma^t p$ satisfies the recursion*

$$\Gamma^t p(n) = q(n) + \sum_{i=1}^{k+t} \left(\Delta^t a(i) + \frac{\Delta^t b(i)}{n} \right) \Gamma^t p(n-i) \quad (n = 1, 2, \ldots) \tag{8.3}$$

with $a(0) = -1$.

Proof The recursion (8.3) trivially holds for $t = 0$ as in that case it reduces to (5.6). Let us now assume that it holds for $t = s - 1$ for some positive integer s. For $n = 1, 2, \ldots$, application of (8.1), (8.3), (1.35) gives

$$\begin{aligned}
\Gamma^s p(n) &= \Gamma^s p(n-1) + \Gamma^{s-1} p(n) \\
&= \Gamma^s p(n-1) + q(n) + \sum_{i=1}^{k+s-1} \left(\Delta^{s-1} a(i) + \frac{\Delta^{s-1} b(i)}{n} \right) \Gamma^{s-1} p(n-i) \\
&= \Gamma^s p(n-1) + q(n) + \sum_{i=1}^{k+s-1} \left(\Delta^{s-1} a(i) + \frac{\Delta^{s-1} b(i)}{n} \right) \Delta \Gamma^s p(n-i) \\
&= q(n) + \sum_{i=1}^{k+s} \left(\Delta^s a(i) + \frac{\Delta^s b(i)}{n} \right) \Gamma^s p(n-i),
\end{aligned}$$

so that (8.3) holds also for $t = s$. Induction gives that it holds for all non-negative integers t. □

The following theorem gives an analogous recursion for $\Lambda^t p$.

Theorem 8.2 *If $p \in \mathcal{F}_{10}$ satisfies the recursion (5.6) for some positive integer k, then, for all non-negative integers t for which $\Lambda^t p$ exists, $\Lambda^t p$ satisfies the recursion*

$$\Lambda^t p(n) = (-1)^t q(n) + \sum_{i=1}^{k+t} \left(\Delta^t a(i) + \frac{\Delta^t b(i)}{n} \right) \Lambda^t p(n-i) \qquad (n = 1, 2, \ldots) \tag{8.4}$$

with $a(0) = -1$. For $t > 0$, the initial values are given by

$$\Lambda^t p(n) = \sum_{i=t+n}^{\infty} \binom{i-n-1}{t-1} p(i). \quad (n = 0, -1, -2, \ldots, 1-k-t) \tag{8.5}$$

Proof The deduction of (8.4) is analogous to the deduction of (8.3), the initial values (8.5) are obtained from Corollary 1.1. This completes the proof of Theorem 8.2. □

Note that as in (8.4) we normally have $\Lambda^t p(n-i) \neq 0$ when $i > n$, we cannot just replace $k+t$ with n as upper limit in the summation in (8.4). This is the reason that we cannot apply this recursion with $k = \infty$.

We shall now deduce an alternative recursion for $\Gamma^t p$.

Theorem 8.3 *If $p \in \mathcal{F}_{10}$ satisfies the recursion (5.6) where k is a positive integer or infinity, then, for $t = 0, 1, 2, \ldots$, $\Gamma^t p$ satisfies the recursion*

$$\Gamma^t p(n) = \frac{\Gamma^t \Phi q(n)}{n} + \sum_{i=1}^{n} \left(a(i) + \frac{b(i) + t(1 - \Gamma a(i-1))}{n} \right) \Gamma^t p(n-i). \quad (n = 1, 2, \ldots) \tag{8.6}$$

Proof We shall prove the theorem by induction on two levels. We first note that (8.6) trivially holds for $t = 0$ as in that case it reduces to (5.6). Then we assume that it holds for $t = s - 1$ for some positive integer s. Under this assumption, we first show that (8.6) holds for $t = s$ and $n = 1$. Then we prove that if it holds for $t = s$ and $n = j - 1$ for some integer $j > 1$, then it must hold also for $n = j$. It then follows by induction that it holds for all positive integers n when $t = s$, and from this it follows by induction that this result must hold for all non-negative integers t.

Let us now assume that (8.6) holds for $t = s - 1$ for some positive integer s. From (8.6), using that $\Gamma^t p(0) = p(0)$ for all t, we obtain

$$\begin{aligned}
\Gamma^s p(1) &= \Gamma^{s-1} p(0) + \Gamma^{s-1} p(1) \\
&= \Gamma^{s-1} p(0) + \Gamma^{s-1} \Phi q(1) + (a(1) + b(1) + s - 1)\Gamma^{s-1} p(0) \\
&= \Gamma^s \Phi q(1) + (a(1) + b(1) + s)\Gamma^s p(0),
\end{aligned}$$

so that (8.6) holds for $n = 1$ when $t = s$. Let us now assume that it holds for $n = j - 1$ for some integer $j > 1$ when $t = s$. Then, application of (8.6) gives

$$\begin{aligned}
j\Gamma^s p(j) &= j(\Gamma^s p(j-1) + \Gamma^{s-1} p(j)) \\
&= \Gamma^s p(j-1) + (j-1)\Gamma^s p(j-1) + j\Gamma^{s-1} p(j) \\
&= \Gamma^s p(j-1) + \Gamma^s \Phi q(j-1) + \sum_{i=1}^{j-1} ((j-1)a(i) + b(i) \\
&\quad + s(1 - \Gamma a(i-1)))\Gamma^s p(j-1-i) + \Gamma^{s-1} \Phi q(j) \\
&\quad + \sum_{i=1}^{j} (ja(i) + b(i) + (s-1)(1 - \Gamma a(i-1)))\Gamma^{s-1} p(j-i) \\
&= \Gamma^s \Phi q(j) + \sum_{i=1}^{j} ((ja(i) + b(i) + s(1 - \Gamma a(i-1)))\Gamma^s p(j-i) \\
&\quad + \Gamma a(i-1)(\Gamma^s p(j-i) - \Gamma^s p(j-1-i)) - a(i)\Gamma^s p(j-1-i))
\end{aligned}$$

$$= \Gamma^s \Phi q(j) + \sum_{i=1}^{j} ((ja(i) + b(i) + s(1 - \Gamma a(i-1)))\Gamma^s p(j-i)$$
$$+ \Gamma a(i-1)\Gamma^s p(j-i) - \Gamma a(i)\Gamma^s p(j-1-i))$$
$$= \Gamma^s \Phi q(j) + \sum_{i=1}^{j} (ja(i) + b(i) + s(1 - \Gamma a(i-1)))\Gamma^s p(j-i),$$

from which we see that (8.6) holds also for $n = j$ when $t = s$. By induction, it then holds for all positive integers n when $t = s$. Thus, we have shown that if (8.6) holds for $t = s - 1$, then it also holds for $t = s$, and it then follows by induction that it holds for all non-negative integers t. □

In the following corollary, we modify the recursion (8.6) in a way that will usually be more convenient when k is finite.

Corollary 8.1 *If $p \in \mathcal{F}_{10}$ satisfies the recursion (5.6) where k is a positive integer or infinity, then, for $t = 1, 2, \ldots$, we have the recursion*

$$\Gamma^t p(n) = \frac{\Gamma^{t-1}\Phi q(n)}{n} + \sum_{i=1}^{k+1} \left(\Delta a(i) + \frac{\Delta b(i) - (t-1)a(i-1)}{n} \right) \Gamma^t p(n-i)$$
$$(n = 1, 2, \ldots) \tag{8.7}$$

with $a(0) = -1$.

Proof For the moment, we let $a(0) = 0$ as usual.

For $t, n = 1, 2, \ldots$, application of (8.1) and (8.6) gives

$$\Gamma^t p(n) = \Gamma^t p(n-1) + \Gamma^{t-1} p(n)$$
$$= \Gamma^t p(n-1) + \frac{\Gamma^{t-1}\Phi q(n)}{n}$$
$$+ \sum_{i=1}^{n} \left(a(i) + \frac{b(i) + (t-1)(1 - \Gamma a(i-1))}{n} \right) \Gamma^{t-1} p(n-i)$$
$$= \frac{\Gamma^{t-1}\Phi q(n)}{n} + \Gamma^t p(n-1)$$
$$+ \sum_{i=1}^{n} \left(a(i) + \frac{b(i) + (t-1)(1 - \Gamma a(i-1))}{n} \right)$$
$$\times (\Gamma^t p(n-i) - \Gamma^t p(n-i-1))$$
$$= \frac{\Gamma^{t-1}\Phi q(n)}{n} + \Gamma^t p(n-1)$$

$$+\sum_{i=1}^{n}\left(a(i)+\frac{b(i)+(t-1)(1-\Gamma a(i-1))}{n}\right)\Gamma^{t}p(n-i)$$

$$-\sum_{i=2}^{n}\left(a(i-1)+\frac{b(i-1)+(t-1)(1-\Gamma a(i-2))}{n}\right)\Gamma^{t}p(n-i)$$

$$=\frac{\Gamma^{t-1}\Phi q(n)}{n}+\left(1+a(1)+\frac{b(1)+t-1}{n}\right)\Gamma^{t}p(n-1)$$

$$+\sum_{i=2}^{n}\left(\Delta a(i)+\frac{\Delta b(i)-(t-1)a(i-1)}{n}\right)\Gamma^{t}p(n-i),$$

which gives (8.7) when letting $a(0)=-1$. □

When k is finite, (8.7) will normally be more convenient than (8.6) as in (8.7), the summation term vanishes for all $i>k+1$, whereas in (8.6) this is not the case unless $\mu_a(0)=1$. From (8.6), we obtain

$$\Gamma^{t}p(n)=\frac{\Gamma^{t}\Phi q(n)}{n}+\sum_{i=1}^{k}\left(a(i)+\frac{b(i)+t(1-\Gamma a(i-1))}{n}\right)\Gamma^{t}p(n-i)$$
$$+\frac{t}{n}(1-\mu_a(0))\Gamma^{t+1}p(n-k-1). \qquad (n=1,2,\dots;\ t=0,1,2,\dots) \tag{8.8}$$

8.2 Convolutions

By combining Corollary 5.6 and Theorem 8.3, we obtain the following corollary.

Corollary 8.2 *If $p\in\mathcal{P}_{1\underline{0}}$, then for $t=0,1,2,\dots$ and $M=1,2,\dots$, we have the recursion*

$$\Gamma^{t}p^{M*}(n)=\frac{1}{p(0)}\sum_{i=1}^{n}\left(\left((M+1)\frac{i}{n}-1\right)p(i)+\frac{t}{n}\Gamma p(i-1)\right)\Gamma^{t}p^{M*}(n-i).$$
$$(n=1,2,\dots)$$

In particular, for $M=1$, this gives

$$\Gamma^{t}p(n)=\frac{1}{p(0)}\sum_{i=1}^{n}\left(\left(2\frac{i}{n}-1\right)p(i)+\frac{t}{n}\Gamma p(i-1)\right)\Gamma^{t}p(n-i). \quad (n=1,2,\dots)$$

With this recursion, we can evaluate $\Gamma^{t}p$ directly from p instead of having to evaluate $\Gamma^{s}p$ by (8.1) for $s=1,2,\dots,t$.

Example 8.1 Let $p \in \mathcal{P}_{1\underline{0}}$ be the discrete uniform distribution given by (2.24). We want to evaluate $\Gamma^t p^{M*}$ by application of (5.22) with one of the recursions we have deduced now. In the present case, $\mu_a(0) = 1$, so that the last term in (8.8) vanishes, and this recursion becomes simpler than (8.7). We obtain

$$\begin{aligned}\Gamma^t p^{M*}(n) &= \left(1 + \frac{M+t-1}{n}\right)\Gamma^t p^{M*}(n-1)\\ &\quad + \left(1 - \frac{(M+1)(k+1)}{n}\right)\Gamma^t p^{M*}(n-k-1)\\ &\quad + \left(\frac{(M+1)k-t+2}{n} - 1\right)\Gamma^t p^{M*}(n-k-2). \qquad (n = 1, 2, \dots)\end{aligned}$$

□

8.3 Compound Distributions

If $f = p \vee h$ with $h \in \mathcal{P}_{10}$ and $p \in \mathcal{P}_{10}$ satisfying the recursion (5.6), then we can obtain recursions for $\Gamma^t f$ and $\Lambda^t f$ by replacing p, k, q, a, and b with f, ∞, $\tilde{q}$ given by (5.9), and $\tilde{a}$ and $\tilde{b}$ given by (5.10). In particular, Theorem 8.3 gives the following corollary.

Corollary 8.3 *If $h \in \mathcal{P}_{10}$ and $p \in \mathcal{P}_{10}$ satisfies the recursion* (5.6), *then for $t = 0, 1, 2, \dots$, $f = p \vee h$ satisfies the recursion*

$$\begin{aligned}\Gamma^t f(x) &= \frac{1}{1-\tau_a(h(0))}\left(\frac{\Gamma^t \Phi(q \vee h)(x)}{x} + \sum_{y=1}^{x}\left((a \vee h)(y)\right.\right.\\ &\quad \left.\left. + \frac{\Phi(\Psi b \vee h)(y) + t(1 - \Gamma(a \vee h)(y-1))}{x}\right)\Gamma^t f(x-y)\right).\\ &\quad (x = 1, 2, \dots)\end{aligned} \tag{8.9}$$

In the following corollary, we consider the special case when the counting distribution p belongs to the Panjer class.

Corollary 8.4 *If p is $R_1[a, b]$ and $h \in \mathcal{P}_{10}$, then for $t = 0, 1, 2, \dots$, $f = p \vee h$ satisfies the recursion*

$$\Gamma^t f(x) = \frac{1}{1-ah(0)}\sum_{y=1}^{x}\left(\left(a + b\frac{y}{x}\right)h(y) + \frac{t}{x}(1 - a\Gamma h(y-1))\right)\Gamma^t f(x-y).$$
$$(x = 1, 2, \dots)$$

Application of Theorem 5.5 in (8.6) and (8.7) gives the following corollary.

Corollary 8.5 *If* $f = p \vee h$ *where* p *is the Poisson distribution* $\mathrm{Po}(\lambda)$ *and* $h \in \mathcal{P}_{10}$ *satisfies the conditions of Theorem 2.3, then*

$$\Gamma^t f(x) = \sum_{y=1}^{x} \left(\frac{\lambda}{x}\eta(y) + \left(1 - \frac{y}{x}\right)\chi(y) + \frac{t}{x}(1 - \Gamma\chi(y-1)) \right)\Gamma^t f(x-y)$$
$$(x = 1, 2, \dots;\ t = 0, 1, 2, \dots)$$

and, with $\chi(0) = -1$,

$$\Gamma^t f(x) = \sum_{y=1}^{r+1} \left(\Delta\chi(y) + \frac{\Delta(\lambda\eta - \Phi\chi)(y) - (t-1)\chi(y-1)}{x} \right)\Gamma^t f(x-y).$$
$$(x, t = 1, 2, \dots) \tag{8.10}$$

Example 8.2 Let $h \in \mathcal{P}_{11}$ be the shifted geometric distribution given by (2.21). Then h satisfies the conditions of Theorem 2.3 with r, η, and χ given by (2.23). Insertion in (8.10) gives

$$\begin{aligned}\Gamma^t f(x) &= \left(\frac{\lambda(1-\pi) - 2\pi + t - 1}{x} + 2\pi + 1 \right)\Gamma^t f(x-1) \\ &\quad + \left(\frac{\lambda(\pi - 1) + 2\pi(\pi - t + 2)}{x} - \pi(\pi + 2) \right)\Gamma^t f(x-2) \\ &\quad + \left(\frac{t-3}{x} + 1 \right)\pi^2\Gamma^t f(x-3). \qquad (x, t = 1, 2, \dots)\end{aligned} \tag{8.11}$$

□

8.4 De Pril Transforms

By letting $k = \infty$, $q = a \equiv 0$, and $b = \varphi_p$ in Theorem 8.3, we obtain the following corollary.

Corollary 8.6 *For all* $f \in \mathcal{P}_{1\underline{0}}$ *and* $t = 0, 1, 2, \dots$, *we have the recursion*

$$\Gamma^t f(x) = \frac{1}{x}\sum_{y=1}^{x} (\varphi_f(y) + t)\Gamma^t f(x-y). \qquad (x = 1, 2, \dots) \tag{8.12}$$

This recursion is thought-provoking. By comparison with (6.2), we see that the first factor in the summand takes the role of a De Pril transform of $\Gamma^t f$, that is,

$$\varphi_{\Gamma^t f}(x) = \varphi_f(x) + t. \qquad (x = 1, 2, \dots) \tag{8.13}$$

However, till now we have defined De Pril transforms only for distributions in $\mathcal{P}_{1\underline{0}}$, and $\Gamma^t f$ is not a distribution, so if we should introduce this De Pril transform, then we would have to extend the definition of the De Pril transform to a wider class of functions. Here, $\mathcal{F}_{1\underline{0}}$ seems to be a natural candidate, and for functions in this class, we still define the De Pril transform by (6.1).

In the proof of Theorem 1.2, we argued that $\Gamma^t f = \gamma^{t*} * f$ with $\gamma \in \mathcal{F}_{1\underline{0}}$ given by (1.37). If (6.1) should hold also for functions in $\mathcal{F}_{1\underline{0}}$, then we should have $\varphi_{\Gamma^t f} = \varphi_f + t\varphi_\gamma$. This is consistent with (8.13), and then we would have

$$\varphi_\gamma(x) = 1. \qquad (x = 1, 2, \dots) \tag{8.14}$$

Application of (6.1) also gives this expression.

Example 8.3 For $j = 1, 2, \dots, M$, let f_j be the discrete uniform distribution on $0, 1, 2, \dots, k_j$. Then insertion of (6.8) in (8.12) gives

$$\Gamma^t(*_{j=1}^M f_j)(x) = \frac{1}{x}\Biggl((M+t)\Gamma^{t+1}(*_{j=1}^M f_j)(x-1) - \sum_{j=1}^M (k_j+1) \sum_{z=1}^{[x/(k_j+1)]} \Gamma^t(*_{j=1}^M f_j)(x - z(k_j+1))\Biggr).$$

$(x = 1, 2, \dots)$ □

We shall discuss De Pril transforms of functions in $\mathcal{F}_{1\underline{0}}$ more thoroughly in Chap. 10.

8.5 The Special Case $b \equiv 0$

In this section, we consider functions that satisfy (5.6) with $b \equiv 0$, that is,

$$p(n) = q(n) + \sum_{i=1}^k a(i)p(n-i) \qquad (n = 1, 2, \dots) \tag{8.15}$$

with k being a positive integer or infinity.

Theorem 8.4 *If $p \in \mathcal{F}_{10}$ satisfies the recursion* (8.15), *then, for $t = 0, 1, 2, \dots$, we have the recursion*

$$\Gamma^t p(n) = \Gamma^t q(n) + \sum_{i=1}^k a(i)\Gamma^t p(n-i) \qquad (n = 1, 2, \dots) \tag{8.16}$$

with

$$q(0) = p(0). \tag{8.17}$$

Proof The recursion (8.16) trivially holds for $t = 0$ as it then reduces to (8.15). Let us now assume that it holds for $t = s - 1$ for some positive integer s. For $n = 1, 2, \ldots$, we then have

$$\Gamma^s p(n) = \sum_{i=0}^{n} \Gamma^{s-1} p(i) = p(0) + \sum_{i=1}^{n} \left(\Gamma^{s-1} q(i) + \sum_{j=1}^{k} a(j) \Gamma^{s-1} p(i-j) \right)$$

$$= \Gamma^s q(n) + \sum_{j=1}^{k} a(j) \sum_{i=1}^{n} \Gamma^{s-1} p(i-j) = \Gamma^s q(n) + \sum_{j=1}^{k} a(j) \Gamma^s p(n-j)$$

so that (8.16) holds also for $t = s$. By induction follows that it holds for all non-negative integers t. $\square$

Dropping the convention (8.17) and now letting $q(0) = 0$ as usual, we can rewrite (8.16) as

$$\Gamma^t p(n) = p(0) \Gamma^t r(n) + \Gamma^t q(n) + \sum_{i=1}^{k} a(i) \Gamma^t p(n-i), \quad (n = 1, 2, \ldots) \tag{8.18}$$

with $r(n) = I(n = 0)$ for $n = 0, 1, 2, \ldots$. Application of Theorem 1.2 gives

$$\Gamma^t r(n) = \binom{n+t-1}{t-1}. \qquad (n = 0, 1, 2, \ldots; t = 1, 2, \ldots)$$

Insertion in (8.18) gives

$$\Gamma^t p(n) = p(0) \binom{n+t-1}{t-1} + \Gamma^t q(n) + \sum_{i=1}^{k} a(i) \Gamma^t p(n-i). \tag{8.19}$$
$$(n, t = 1, 2, \ldots)$$

Theorem 8.5 *If $p \in \mathcal{F}_{10}$ satisfies the recursion*

$$p(n) = q(n) + \sum_{i=1}^{n} a(i) p(n-i), \qquad (n = 1, 2, \ldots) \tag{8.20}$$

then, for $t = 0, 1, 2, \ldots$, we have the recursion

$$\Lambda^t p(n) = \Lambda^t q(n) + \sum_{u=1}^{t} \Lambda^u a(n) \Lambda^{t-u+1} p(-1) + \sum_{i=1}^{n} a(i) \Lambda^t p(n-i).$$
$$(n = 1, 2, \ldots) \tag{8.21}$$

Proof The recursion (8.21) trivially holds for $t = 0$ as it then reduces to (8.20). Let us now assume that it holds for $t = s - 1$ for some positive integer s. For $n =$

1, 2, ..., we then have

$$\begin{aligned}
\Lambda^s p(n) &= \sum_{i=n+1}^{\infty} \Lambda^{s-1} p(i) \\
&= \sum_{i=n+1}^{\infty} \left(\Lambda^{s-1} q(i) + \sum_{u=1}^{s-1} \Lambda^u a(i) \Lambda^{s-u} p(-1) + \sum_{j=1}^{i} a(j) \Lambda^{s-1} p(i-j) \right) \\
&= \Lambda^s q(n) + \sum_{u=1}^{s-1} \Lambda^{u+1} a(n) \Lambda^{s-u} p(-1) \\
&\quad + \sum_{j=1}^{\infty} a(j) \sum_{i=\max(n+1,j)}^{\infty} \Lambda^{s-1} p(i-j) \\
&= \Lambda^s q(n) + \sum_{u=2}^{s} \Lambda^u a(n) \Lambda^{s-u+1} p(-1) + \sum_{j=1}^{\infty} a(j) \Lambda^s p((n+1-j)_+ - 1) \\
&= \Lambda^s q(n) + \sum_{u=1}^{s} \Lambda^u a(n) \Lambda^{s-u+1} p(-1) + \sum_{j=1}^{n} a(j) \Lambda^s p(n-j),
\end{aligned}$$

that is, (8.21) holds also for $t = s$. By induction follows that it holds for all non-negative integers t. □

By combining Theorems 8.4 and 8.5 with (5.40), we obtain the following corollary.

Corollary 8.7 *If $p \in \mathcal{P}_{1\underline{0}}$ has rational generating function given by (5.38), then p satisfies the recursions*

$$\Gamma^t p(n) = \Gamma^t \alpha(n) + \sum_{i=1}^{m} \beta(i) \Gamma^t p(n-i)$$

$$\Lambda^t p(n) = \Lambda^t \alpha(n) + \sum_{u=1}^{t} \Lambda^u \beta(n) \Lambda^{t-u+1} p(-1) + \sum_{i=1}^{n} \beta(i) \Lambda^t p(n-i) \qquad (8.22)$$

for $n = 1, 2, \ldots$ and $t = 0, 1, 2, \ldots$.

In this corollary, we did not need to state a convention like (8.17) as (5.40) implies that $\alpha(0) = p(0)$.

If $f = p \vee h$ with $p, h \in \mathcal{P}_{10}$ and p satisfying the recursion (8.15), then, by (5.8), we obtain

$$f(x) = \frac{1}{1 - \tau_a(h(0))} \left((q \vee h)(x) + \sum_{y=1}^{x} (a \vee h)(y) f(x-y) \right), \qquad (x = 1, 2, \ldots)$$

that is, a recursion in the same form. Hence, (8.19) gives

$$\Gamma^t f(x) = \binom{x+t-1}{t-1} f(0) + \frac{1}{1-\tau_a(h(0))}\Bigg(\Gamma^t (q \vee h)(x) - (q \vee h)(0) + \sum_{y=1}^{x} (a \vee h)(y)\Gamma^t f(x-y)\Bigg), \qquad (x, t = 1, 2, \ldots) \tag{8.23}$$

where we have adjusted for the possibility of $(q \vee h)(0) \neq 0$.

From Theorem 8.5 we get

$$\Lambda^t f(x) = \frac{1}{1-\tau_a(h(0))}\Bigg(\Lambda^t (q \vee h)(x) + \sum_{j=1}^{t} \Lambda^j (a \vee h)(x)\Lambda^{t-j+1} f(-1) + \sum_{y=1}^{x} (a \vee h)(y)\Lambda^t f(x-y)\Bigg). \tag{8.24}$$

$$(x = 1, 2, \ldots;\ t = 0, 1, 2, \ldots)$$

Example 8.4 Let p be the geometric distribution geo(π). Then the recursion (8.15) is satisfied as

$$p(n) = \pi p(n-1), \qquad (n = 1, 2, \ldots)$$

and insertion in (8.23) and (8.24) gives that for $x = 1, 2, \ldots,$

$$\Gamma^t f(x) = \frac{1}{1-\pi h(0)}\left((1-\pi)\binom{x+t-1}{t-1} + \pi \sum_{y=1}^{x} h(y)\Gamma^t f(x-y)\right)$$

$$\Lambda^t f(x) = \frac{\pi}{1-\pi h(0)}\left(\sum_{j=1}^{t} \Lambda^j h(x)\Lambda^{t-j+1} f(-1) + \sum_{y=1}^{x} h(y)\Lambda^t f(x-y)\right).$$

For $t = 1$, we obtain

$$\Gamma f(x) = \frac{1}{1-\pi h(0)}\left(1 - \pi + \pi \sum_{y=1}^{x} h(y)\Gamma f(x-y)\right)$$

$$\Lambda f(x) = \frac{\pi}{1-\pi h(0)}\left(\Lambda h(x) + \sum_{y=1}^{x} h(y)\Lambda f(x-y)\right).$$

The former recursion is in the form of a discrete renewal equation with defective distribution. □

Further Remarks and References

Sundt (1982) deduced recursions for the cumulative distribution function, tail, and stop loss transform of a compound geometric distribution. Compound geometric distributions with continuous severity distribution often appear in connection with infinite time ruin theory. By discretising the severity distribution, one can obtain upper and lower bounds for the ruin probability, see Dickson (1995).

The first recursion in Corollary 8.5 was presented by Chadjiconstantinidis and Pitselis (2008), who also gave recursions for $\Lambda^t f$ under the same assumptions. They also deduced the recursion (8.11).

Examples 8.1 and 8.3 are based on Sundt (1999a).

Most of the other results in the present chapter are based on Dhaene et al. (1999).

Waldmann (1996) discussed recursions for the cumulative distribution function, stop loss transform, and retention transform of a compound distribution with counting distribution in $\bigcup_{l=0}^{\infty} \mathcal{S}_l$ and severity distribution in $\mathcal{P}_{11}$.

Antzoulakos and Chadjiconstantinidis (2004) presented recursions for cumulative functions and tails of mixed Poisson distributions and compound mixed Poisson distributions with mixing distribution in the Willmot class and severity distribution in $\mathcal{P}_{10}$.

Waldmann (1995) deduced an algorithm for recursive evaluation of the retention transform of the aggregate claims distribution in the individual life model.

Chapter 9
Moments

Summary

Recursions for moments is the topic of the present chapter. It consists of two sections; Sect. 9.1 is devoted to evaluation of moments of convolutions of a distribution in terms of moments of that distribution, whereas in Sect. 9.2, we consider evaluation of moments of a compound distribution in terms of moments of its severity distribution.

In Sect. 9.1.1, we develop recursions for ordinary moments of convolutions of a distribution. We apply one of these recursions to prove a characterisation of normal distributions in Sect. 9.1.2. By comparing a differential equation for generating functions with a differential equation for moment generating functions in Sect. 9.1.3, we conclude that a recursion for factorial moments has the same shape as the recursion for central moments.

In Sect. 9.2.1, we deduce recursions for moments of compound distributions with counting distribution satisfying (5.6). The special case with counting distribution in the Panjer class is treated in Sect. 9.2.2. Section 9.2.3 is devoted to compound Poisson distributions with severity distribution satisfying the conditions of Theorem 2.3.

9.1 Convolutions of a Distribution

9.1.1 Ordinary Moments

In Sect. 9.1, we shall discuss recursions for moments of the M-fold convolution of univariate distributions. As we do not need to restrict to distributions on integers, we shall identify the distributions by their cumulative distribution function.

We shall need the following lemma.

Lemma 9.1 *Let X and Y be independent random variables such that the distribution of X is the M-fold convolution of the distribution of Y. Then*

$$\mathrm{E}(X - MY)r(X + Y) = 0$$

for any function r.

Proof For any number x, we have

$$0 = r(x)\,\mathrm{E}[X - MY|X + Y = x] = \mathrm{E}[(X - MY)r(X + Y)|X + Y = x].$$

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_9, © Springer-Verlag Berlin Heidelberg 2009

From this, we obtain

$$\mathrm{E}(X - MY)r(X+Y) = \mathrm{E}\,\mathrm{E}[(X-MY)r(X+Y)|X+Y] = 0. \qquad \square$$

In particular, by letting $r(z) = I(z = x)$ for some number x, we obtain

$$0 = \mathrm{E}(X - MY)I(X+Y = x) = \mathrm{E}(x - (M+1)Y)I(X+Y = x),$$

and by division by x, we obtain (2.52), that is, for this choice of r, we have already applied Lemma 9.1 to deduce a recursion for an M-fold convolution. We shall now apply another choice of r to deduce a recursion for moments of an M-fold convolution.

Theorem 9.1 *Let G be a univariate distribution and $F = G^{M*}$ for some positive integer M. Then*

$$\mu_F(j;c) = \sum_{u=1}^{j} \binom{j-1}{u-1} \left(\left(M + 1 - \frac{j}{u} \right) \mu_G(u) - c\mu_G(u-1) \right) \mu_F(j-u;c). \quad (j = 1, 2, \dots) \tag{9.1}$$

Proof Let j be a positive integer and X and Y independent random variables with distribution F and G respectively. By letting $r(x) = (x-c)^{j-1}$ in Lemma 9.1, we obtain

$$\begin{aligned}
0 &= \mathrm{E}(X - MY)(X+Y-c)^{j-1} \\
&= \sum_{u=0}^{j-1} \binom{j-1}{u} \mathrm{E}((X-c) - MY + c)Y^u (X-c)^{j-1-u} \\
&= \sum_{u=0}^{j-1} \binom{j-1}{u} \big(\mu_G(u)\mu_F(j-u;c) - M\mu_G(u+1)\mu_F(j-u-1;c) \\
&\quad + c\mu_G(u)\mu_F(j-u-1;c) \big) \\
&= \mu_F(j;c) \\
&\quad - \sum_{u=1}^{j} \binom{j-1}{u-1} \left(M\mu_G(u) - c\mu_G(u-1) - \frac{j-u}{u}\mu_G(u) \right) \mu_F(j-u;c),
\end{aligned}$$

which gives (9.1). $\square$

For all the recursions that we deduce in this chapter, the initial condition is that $\mu_F(0;c) = 1$.

By replacing Y with $Y-b$, X with $X-Mb$, and c with $c-Mb$, we obtain

$$\mu_F(j;c)=\sum_{u=1}^{j}\binom{j-1}{u-1}\left(\left(M+1-\frac{j}{u}\right)\mu_G(u;b)+(Mb-c)\mu_G(u-1;b)\right)$$
$$\times\,\mu_F(j-u;c). \qquad (j=1,2,\dots) \tag{9.2}$$

In particular, letting $c=0$ gives

$$\mu_F(j)=\sum_{u=1}^{j}\binom{j-1}{u-1}\left(\left(M+1-\frac{j}{u}\right)\mu_G(u;b)+Mb\mu_G(u-1;b)\right)\mu_F(j-u).$$
$$(j=1,2,\dots)$$

Let us consider some special cases.

1. For ordinary non-central moments, that is, $b=c=0$, we obtain

$$\mu_F(j)=\sum_{u=1}^{j}\binom{j-1}{u-1}\left(M+1-\frac{j}{u}\right)\mu_G(u)\mu_F(j-u). \qquad (j=1,2,\dots) \tag{9.3}$$

2. To evaluate the central moments of F, we let $c=\mu_F(1)=M\mu_G(1)$. By, in addition, letting $b=\mu_G(1)$, we can evaluate the central moments of F from the central moments of G. We obtain

$$\mu_F(j;\mu_F(1))=M\mu_G(j;\mu_G(1))$$
$$+\sum_{u=2}^{j-2}\binom{j-1}{u-1}\left(M+1-\frac{j}{u}\right)\mu_G(u;\mu_G(1))\mu_F(j-u;\mu_F(1)).$$
$$(j=1,2,\dots)$$

In particular, we see that $\mu_F(j;\mu_F(1))=M\mu_G(j;\mu_G(1))$ for $j=2,3$. If G is symmetric, then that is the case with F too, and then $\mu_F(j;\mu_F(1))=\mu_G(j;\mu_G(1))=0$ when j is odd, so that

$$\mu_F(2j;\mu_F(1))=\sum_{u=1}^{j}\binom{2j-1}{2u-1}\left(M+1-\frac{j}{u}\right)\mu_G(2u;\mu_G(1))$$
$$\times\,\mu_F(2(j-u);\mu_F(1)). \qquad (j=1,2,\dots)$$

3. For $M=1$, (9.2) reduces to

$$\mu_F(j;c)=\sum_{u=1}^{j}\binom{j-1}{u-1}\left(\left(2-\frac{j}{u}\right)\mu_F(u;b)+(b-c)\mu_F(u-1;b)\right)$$
$$\times\,\mu_F(j-u;c), \qquad (j=1,2,\dots)$$

which can be applied to obtain moments of F around c from moments of F around b. In particular, for $j = 1, 2, \ldots$, we have

$$\mu_F(j;c) = \sum_{u=1}^{j} \binom{j-1}{u-1} \left(\left(2 - \frac{j}{u}\right) \mu_F(u) - c\mu_F(u-1) \right) \mu_F(j-u;c)$$

$$\mu_F(j) = \sum_{u=1}^{j} \binom{j-1}{u-1} \left(\left(2 - \frac{j}{u}\right) \mu_F(u;c) + c\mu_F(u-1;c) \right) \mu_F(j-u). \tag{9.4}$$

Example 9.1 In the special case when g is the Bernoulli distribution $\mathrm{Bern}(\pi)$, $f = g^{M*}$ is the binomial distribution $\mathrm{bin}(M, \pi)$. In this case, we have $\mu_g(j) = \pi$ for $j = 1, 2, \ldots$. Insertion in (9.1) gives

$$\mu_f(j;c) = ((M+1-j)\pi - c)\mu_f(j-1;c) + \pi \sum_{u=2}^{j} \binom{j-1}{u-1} \left(M + 1 - \frac{j}{u} - c\right) \mu_f(j-u;c); \qquad (j = 1, 2, \ldots)$$

in particular,

$$\mu_f(j) = \pi \sum_{u=1}^{j} \binom{j-1}{u-1} \left(M + 1 - \frac{j}{u}\right) \mu_f(j-u). \qquad (j = 1, 2, \ldots)$$

□

Let X and Y be independent random variables with distribution F and G respectively. Sometimes it can be of interest to evaluate the moments of a rescaled convolution variable, that is, the moments of the distribution $\tilde{F}$ of X/d instead of F. If we were interested in the moments of the empirical mean, then we would have $d = M$, and it could also be of interest to consider the case $d = \sqrt{M}$ such that $\tilde{F}$ and G have the same variance. With $\tilde{G}$ denoting the distribution of Y/d, we can of course immediately replace F and G with $\tilde{F}$ and $\tilde{G}$ in (9.1), and, as $\mu_{\tilde{G}}(u) = \mu_G(u)/d^u$, we obtain

$$\mu_{\tilde{F}}(j;c) = \sum_{u=1}^{j} \binom{j-1}{u-1} d^{-u} \left(\left(M + 1 - \frac{j}{u}\right) \mu_G(u) - cd\mu_G(u-1) \right) \times \mu_{\tilde{F}}(j-u;c). \qquad (j = 1, 2, \ldots)$$

In particular, with $d = \sqrt{M}$, $c = \mu_F(1) = 0$, and hence $\mu_{\tilde{F}}(1) = 0$, we obtain

$$\mu_{\tilde{F}}(j) = \sum_{u=2}^{j} \binom{j-1}{u-1} \left(M + 1 - \frac{j}{u}\right) \frac{\mu_G(u)}{M^{u/2}} \mu_{\tilde{F}}(j-u). \quad (j = 2, 3, \ldots) \tag{9.5}$$

9.1.2 The Normal Distribution

In Sect. 9.1.2, we shall apply the recursion (9.5) to deduce a characterisation of normal distributions with mean zero. However, we first recapitulate some properties of normal distributions.

The *standard normal distribution* N(0, 1) is a continuous distribution H with density

$$h(y) = \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{y^2}{2}} \qquad (-\infty < y < \infty)$$

with π denoting the Ludolph number. If a random variable Y has the standard normal distribution, then the distribution G of

$$X = \sigma Y + \xi \tag{9.6}$$

is the *normal distribution* $\mathrm{N}(\xi, \sigma)$ with density

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma}\,\mathrm{e}^{-\frac{(x-\xi)^2}{2\sigma^2}}. \qquad (-\infty < x < \infty)$$

We have

$$\omega_H(s) = \mathrm{E}\,\mathrm{e}^{sY} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{y^2-2sy}{2}}\,\mathrm{d}y = \mathrm{e}^{\frac{s^2}{2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{(y-s)^2}{2}}\,\mathrm{d}y = \mathrm{e}^{\frac{s^2}{2}},$$

which gives

$$\omega_G(s) = \mathrm{E}\,\mathrm{e}^{sX} = \mathrm{E}\,\mathrm{e}^{s(\xi+\sigma Y)} = \mathrm{e}^{\xi s}\,\omega_H(\sigma s),$$

that is,

$$\omega_G(s) = \mathrm{e}^{\xi s + \frac{(\sigma s)^2}{2}}. \tag{9.7}$$

This function has obviously finite derivatives of all orders, and, hence, $\mathrm{N}(\xi, \sigma)$ has moments of all orders. In particular, we obtain that it has mean ξ and variance σ^2.

From (9.7) and (1.19), we obtain that for all non-negative integers M

$$\omega_{G^{M*}}(s) = \mathrm{e}^{M\xi s + \frac{(\sqrt{M}\sigma s)^2}{2}}, \tag{9.8}$$

that is, G^{M*} is $\mathrm{N}(M\xi, \sqrt{M}\sigma)$.

We are now ready to give a characterisation of the normal distribution with mean zero.

Theorem 9.2 *A univariate distribution G with finite moments of first and second order is normal with mean zero iff*

$$G^{M*}(x) = G\big(x/\sqrt{M}\big) \qquad (-\infty < x < \infty) \tag{9.9}$$

for some integer $M > 1$.

Proof We have shown above that a normal distribution has finite moments of all orders. Furthermore, we have shown that for any positive integer M, the M-fold convolution of the normal distribution $\mathrm{N}(\xi, \sigma)$ is $\mathrm{N}(M\xi, \sqrt{M}\sigma)$, and this implies that (9.9) holds for all positive integers M when $\xi = 0$.

Let us now assume that (9.9) holds for some integer $M > 1$. This means that if $Y_1, Y_2, \ldots, Y_M$ are independent and identically distributed random variables with distribution G, then $Y_{\bullet M}/\sqrt{M}$ also has distribution G. Then they must have the same mean, which must be zero. From (9.5), we obtain

$$\mu_G(j) = \sum_{u=2}^{j} \binom{j-1}{u-1} \left(M + 1 - \frac{j}{u}\right) \frac{\mu_G(u)}{M^{u/2}} \mu_G(j-u). \qquad (j = 2, 3, \ldots)$$

Solving for $\mu_G(j)$ gives

$$\mu_G(j) = \frac{1}{1 - M^{1-j/2}} \sum_{u=2}^{j-1} \binom{j-1}{u-1} \left(M + 1 - \frac{j}{u}\right) \frac{\mu_G(u)}{M^{u/2}} \mu_G(j-u).$$
$$(j = 3, 4, \ldots)$$

This recursion determines all moments of G from $\mu_G(1) = 0$ and a given value of $\mu_G(2)$. As there exists a normal distribution with the same first and second order moments as G and this normal distribution satisfies (9.9) for all positive integers M, this normal distribution must have the same moments as G, and as the moment generating function of a normal distribution exists for all real numbers, this implies that G must be equal to that normal distribution.

This completes the proof of Theorem 9.2. □

Letting $M \uparrow \infty$ in (9.5), we obtain that, in the limit,

$$\mu_{\tilde{F}}(j) = (j-1)\mu_G(2)\mu_{\tilde{F}}(j-2), \qquad (j = 2, 3, \ldots)$$

which gives that for $j = 1, 2, \ldots$, we have $\mu_{\tilde{F}}(2j-1) = 0$ and $\mu_{\tilde{F}}(2j) = \mu_G(2)^j \prod_{i=1}^{j}(2i-1)$. As

$$\prod_{i=1}^{j}(2i-1) = \frac{(2j-1)!}{\prod_{i=1}^{j-1}(2i)} = \frac{(2j-1)!}{(j-1)!2^{j-1}} = \frac{(2j-1)^{(j)}}{2^{j-1}},$$

we obtain

$$\mu_{\tilde{F}}(j) = \begin{cases} \frac{(j-1)^{(\frac{j}{2})}}{2^{\frac{j}{2}-1}} \mu_G(2)^{\frac{j}{2}} & (j = 2, 4, 6, \ldots) \\ 0. & (j = 1, 3, 5, \ldots) \end{cases} \qquad (9.10)$$

From this, we see that the asymptotic moments of $\tilde{F}$ are determined by $\mu_G(2)$. As there exists a normal distribution with the same first and second order moments as G, and $\tilde{F}$ is normal when G is normal, the recursion (9.10) must generate the moments of a normal distribution, as expected. In particular, if G is normal, then Theorem 9.2 gives that $\tilde{F} = G$. Hence, if G is the normal distribution $\mathrm{N}(0, \sigma)$, then we obtain from (9.10) that

$$\mu_G(j) = \begin{cases} \frac{(j-1)^{(\frac{j}{2})}}{2^{\frac{j}{2}-1}}\sigma^j & (j = 2, 4, 6, \dots) \\ 0. & (j = 1, 3, 5, \dots) \end{cases} \tag{9.11}$$

More generally, if F is the normal distribution $\mathrm{N}(\xi, \sigma)$, then the central moments of F are the moments of G given by (9.11). For the non-central moments of F, insertion of (9.11) in (9.4) gives that for $j = 1, 2, \dots,$

$$\begin{aligned} \mu_F(j) &= \sum_{u=1}^{j} \binom{j-1}{u-1}\left(\left(2 - \frac{j}{u}\right)\mu_F(u; \xi) + \xi\mu_F(u-1; \xi)\right)\mu_F(j-u) \\ &= \sum_{u=1}^{[j/2]} \binom{j-1}{2u-1}\left(2 - \frac{j}{2u}\right)\mu_F(2u; \xi)\mu_F(j-2u) \\ &\quad + \xi \sum_{u=0}^{[(j-1)/2]} \binom{j-1}{2u}\mu_F(2u; \xi)\mu_F(j-2u-1) \\ &= \sum_{u=1}^{[j/2]} \binom{j-1}{2u-1}\left(2 - \frac{j}{2u}\right)\frac{(2u-1)^{(u)}}{2^{u-1}}\sigma^{2u}\mu_F(j-2u) \\ &\quad + \xi \sum_{u=0}^{[(j-1)/2]} \binom{j-1}{2u}\frac{(2u-1)^{(u)}}{2^{u-1}}\sigma^{2u}\mu_F(j-2u-1). \end{aligned}$$

9.1.3 Factorial Moments

Let G be a univariate distribution and $F = G^{M*}$ for some positive integer M. Then, letting $r(x)$ equal to s^x and e^{sx} in Lemma 9.1 gives respectively

$$\tau'_F(s)\tau_G(s) - M\tau'_G(s)\tau_F(s) = 0 \tag{9.12}$$

$$\omega'_F(s)\omega_G(s) - M\omega'_G(s)\omega_F(s) = 0. \tag{9.13}$$

We can obviously deduce (9.3) by taking the jth order derivative of (9.13) at zero, using (1.12), and solving for $\mu_F(j)$. Analogously, we can find a recursion for the factorial moments of F by taking the jth order derivative of (9.12) at one, using (1.14), and solving for $\nu_F(j)$. Because of the similarity of the formulae involved in these two deductions, the latter recursion is obtained by simply replacing the moments in (9.3) with the corresponding factorial moments, that is,

$$\nu_F(j) = \sum_{u=1}^{j} \binom{j-1}{u-1} \left(M + 1 - \frac{j}{u} \right) \nu_G(u) \nu_F(j-u). \qquad (j = 1, 2, \dots)$$

9.2 Compound Distributions

9.2.1 General Results

In Sect. 9.2.1, we consider recursions for moments of a compound distribution $F = p \vee H$ with univariate severity distribution H and counting distribution $p \in \mathcal{P}_{10}$ satisfying the recursion (5.6).

In this situation, we have the following analogue to Lemma 9.1.

Lemma 9.2 *Let $X, Y_1, Y_2, \dots$ be independent random variables, the Y_is with distribution H and X with distribution $p \vee H$ with $p \in \mathcal{P}_{10}$ satisfying the recursion* (5.6) *where k is a positive number or infinity. Then*

$$\begin{aligned} \mathrm{E}\, X r(X) &= \sum_{n=1}^{\infty} q(n)\, \mathrm{E}\, Y_{\bullet n} r(Y_{\bullet n}) \\ &\quad + \sum_{i=1}^{k} \mathrm{E}(a(i)(X + Y_{\bullet i}) + \Psi b(i) Y_{\bullet i}) r(X + Y_{\bullet i}) \end{aligned} \qquad (9.14)$$

for any function r for which these expectations exist.

Proof Let $N, Z_1, Z_2, \dots$ be mutually independent random variables independent of the Y_is; N with distribution p and the Z_is with distribution H. Then

$$\begin{aligned} \mathrm{E}\, X r(X) = \mathrm{E}\, Z_{\bullet N} r(Z_{\bullet N}) &= \sum_{n=1}^{\infty} p(n)\, \mathrm{E}\, Z_{\bullet n} r(Z_{\bullet n}) \\ &= \sum_{n=1}^{\infty} \left(q(n) + \sum_{i=1}^{k} \left(a(i) + \frac{b(i)}{n} \right) p(n-i) \right) \mathrm{E}\, Z_{\bullet n} r(Z_{\bullet n}), \end{aligned}$$

that is,

$$\mathrm{E}\, X r(X) = \sum_{n=1}^{\infty} q(n)\, \mathrm{E}\, Y_{\bullet n} r(Y_{\bullet n}) + \sum_{i=1}^{k} \sum_{n=1}^{\infty} p(n-i) \left(a(i) + \frac{b(i)}{n} \right) \mathrm{E}\, Z_{\bullet n} r(Z_{\bullet n}). \qquad (9.15)$$

For $i = 1, 2, \ldots, k$, we have

$$\begin{aligned}
&\sum_{n=1}^{\infty} p(n-i)\left(a(i) + \frac{b(i)}{n}\right) \mathrm{E}\, Z_{\bullet n} r(Z_{\bullet n}) \\
&\quad = \sum_{n=0}^{\infty} p(n)\left(a(i) + \frac{b(i)}{n+i}\right) \mathrm{E}\, Z_{\bullet(n+i)} r(Z_{\bullet(n+i)}) \\
&\quad = \sum_{n=0}^{\infty} p(n)\left(a(i) + \frac{b(i)}{n+i}\right) \mathrm{E}(Z_{\bullet n} + Y_{\bullet i}) r(Z_{\bullet n} + Y_{\bullet i}) \\
&\quad = \sum_{n=0}^{\infty} p(n)\, \mathrm{E}\left(a(i)(Z_{\bullet n} + Y_{\bullet i}) + \frac{b(i)}{i} Y_{\bullet i}\right) r(Z_{\bullet n} + Y_{\bullet i}) \\
&\quad = \mathrm{E}(a(i)(Z_{\bullet N} + Y_{\bullet i}) + \Psi b(i) Y_{\bullet i}) r(Z_{\bullet N} + Y_{\bullet i}) \\
&\quad = \mathrm{E}(a(i)(X + Y_{\bullet i}) + \Psi b(i) Y_{\bullet i}) r(X + Y_{\bullet i}).
\end{aligned}$$

Insertion in (9.15) gives (9.14). □

If $H \in \mathcal{P}_1$, then letting $r(z) = I(z = x)$ in (9.14) for some integer x gives

$$xf(x) = \sum_{n=0}^{\infty} q(n) x h^{n*}(x) + \sum_{i=1}^{k} \sum_{y=-\infty}^{\infty} (a(i)x + \Psi b(i) y) h^{i*}(y) f(x-y), \quad (9.16)$$

which is a trivial extension of (5.7).

For deducing a recursion for moments of F, we shall also need the following lemma.

Lemma 9.3 *For constants A, B, v, and w, we have*

$$A(v+w)^j + Bw(v+w)^{j-1} = Av^j + \sum_{u=1}^{j} \binom{j-1}{u-1}\left(\frac{j}{u}A + B\right) w^u v^{j-u}.$$

$$(j = 1, 2, \ldots)$$

Proof We have

$$\begin{aligned}
A(v+w)^j + Bw(v+w)^{j-1} &= A\sum_{u=0}^{j} \binom{j}{u} w^u v^{j-u} + B \sum_{u=0}^{j-1} \binom{j-1}{u} w^{u+1} v^{j-1-u} \\
&= Av^j + \sum_{u=1}^{j} \binom{j-1}{u-1}\left(\frac{j}{u}A + B\right) w^u v^{j-u}.
\end{aligned}$$
□

We are now ready to deduce a recursion for moments of F.

Theorem 9.3 *If $F = p \vee H$ with univariate severity distribution $H \in \mathcal{P}_1$ and counting distribution $p \in \mathcal{P}_{10}$ satisfying* (5.6) *with $\mu_a(0) \neq 1$, then*

$$\mu_F(j;c) = \frac{1}{1-\mu_a(0)}\Bigg(\mu_{q\vee H}(j;c) + c\mu_{q\vee H}(j-1;c) - c\mu_F(j-1;c) + \sum_{u=1}^{j}\binom{j-1}{u-1}\left(\frac{j}{u}\mu_{a\vee H}(u) + \mu_{\Psi b\vee H}(u) + c\mu_{a\vee H}(u-1)\right) \times \mu_F(j-u;c)\Bigg) \qquad (j=1,2,\dots) \tag{9.17}$$

for any constant c.

Proof Let $X, Y_1, Y_2, \dots$ be independent random variables, the Y_is with distribution H and X with distribution F. For any positive integer j, we then have

$$\mu_F(j;c) = \mathrm{E}(X-c)^j = \mathrm{E}\,X(X-c)^{j-1} - c\mu_F(j-1;c).$$

Insertion of (9.14) with $r(x) = (x-c)^{j-1}$ gives

$$\mu_F(j;c) = \sum_{n=1}^{\infty} q(n)\,\mathrm{E}\,Y_{\bullet n}(Y_{\bullet n}-c)^{j-1} + \sum_{i=1}^{k}\mathrm{E}(a(i)(X+Y_{\bullet i}) + \Psi b(i)Y_{\bullet i})(X+Y_{\bullet i}-c)^{j-1} - c\mu_F(j-1;c). \tag{9.18}$$

For $i = 1,2,\dots,k$, application of Lemma 9.3 gives

$$\begin{aligned}
&\mathrm{E}(a(i)(X+Y_{\bullet i}) + \Psi b(i)Y_{\bullet i})(X+Y_{\bullet i}-c)^{j-1}\\
&\quad= \mathrm{E}\big(a(i)((X-c)+Y_{\bullet i})^j + \Psi b(i)Y_{\bullet i}((X-c)+Y_{\bullet i})^{j-1}\\
&\qquad + a(i)c((X-c)+Y_{\bullet i})^{j-1}\big)\\
&\quad= a(i)\mu_F(j;c) + \sum_{u=1}^{j}\binom{j-1}{u-1}\left(\left(\frac{j}{u}a(i) + \Psi b(i)\right)\mu_{H^{i*}}(u)\right.\\
&\qquad \left. + ca(i)\mu_{H^{i*}}(u-1)\right)\mu_F(j-u;c).
\end{aligned}$$

By insertion in (9.18), we obtain

$$\mu_F(j;c) = \mu_{q\vee H}(j;c) + c\mu_{q\vee H}(j-1;c) - c\mu_F(j-1;c) + \mu_a(0)\mu_F(j;c)$$
$$+ \sum_{u=1}^{j}\binom{j-1}{u-1}\left(\frac{j}{u}\mu_{a\vee H}(u) + \mu_{\Psi b\vee H}(u) + c\mu_{a\vee H}(u-1)\right)$$
$$\times \mu_F(j-u;c). \tag{9.19}$$

Solving for $\mu_F(j;c)$ gives (9.23). □

Theorem 5.11 is a special case of this theorem.

For evaluation of $\mu_{a\vee H}$ and $\mu_{\Psi b\vee H}$, we need the non-central moments of convolutions of H. These can be evaluated recursively by (9.3). However, if we need $\mu_{H^{i*}}(u)$ for several values of i, then it might be more efficient to use that

$$\mu_{H^{(i+j)*}}(u) = \sum_{v=0}^{u}\binom{u}{v}\mu_{H^{i*}}(v)\mu_{H^{j*}}(u-v), \qquad (i,j,u = 1,2,\dots)$$

which follows from (1.22). When $i = j$, we have the more efficient formula

$$\mu_{H^{2j*}}(u) = 2\sum_{v=0}^{[(u-1)/2]}\binom{u}{v}\mu_{H^{j*}}(v)\mu_{H^{j*}}(u-v) + I(u \text{ even})\binom{u}{u/2}\mu_{H^{j*}}(u/2)^2.$$
$$(j,u = 1,2,\dots)$$

For $c = 0$, (9.17) reduces to

$$\mu_F(j) = \frac{1}{1-\mu_a(0)}\Bigg(\mu_{q\vee H}(j)$$
$$+ \sum_{u=1}^{j}\binom{j-1}{u-1}\left(\frac{j}{u}\mu_{a\vee H}(u) + \mu_{\Psi b\vee H}(u)\right)\mu_F(j-u)\Bigg). \tag{9.20}$$
$$(j = 1,2,\dots)$$

In particular, this gives

$$\mu_F(1) = \frac{\mu_{q\vee H}(1) + \mu_{a\vee H}(1) + \mu_{\Psi b\vee H}(1)}{1-\mu_a(0)}. \tag{9.21}$$

By grouping the terms with factor $\mu_F(j-1;c)$ in (9.17), we obtain

$$\mu_F(j;c) = \frac{1}{1-\mu_a(0)}\Bigg(\mu_{q\vee H}(j;c) + c\mu_{q\vee H}(j-1;c)$$
$$+ (j\mu_{a\vee H}(1) + \mu_{\Psi b\vee H}(1) - c(1-\mu_a(0)))\mu_F(j-1;c)$$

$$+\sum_{u=2}^{j}\binom{j-1}{u-1}\left(\frac{j}{u}\mu_{a\vee H}(u)+\mu_{\Psi b\vee H}(u)+c\mu_{a\vee H}(u-1)\right)$$

$$\left.\times\,\mu_F(j-u;c)\right). \qquad (j=1,2,\dots) \tag{9.22}$$

Insertion of $c=\mu_F(1)$ and application of (9.21) give the following recursion for the central moments:

$$\mu_F(j;\mu_F(1))=\frac{1}{1-\mu_a(0)}\Bigg(\mu_{q\vee H}(j;\mu_F(1))+\mu_F(1)\mu_{q\vee H}(j-1;\mu_F(1))$$

$$+((j-1)\mu_{a\vee H}(1)-\mu_{q\vee H}(1))\mu_F(j-1;\mu_F(1))$$

$$+\sum_{u=2}^{j}\binom{j-1}{u-1}\left(\frac{j}{u}\mu_{a\vee H}(u)+\mu_{\Psi b\vee H}(u)+\mu_F(1)\mu_{a\vee H}(u-1)\right)$$

$$\left.\times\,\mu_F(j-u;\mu_F(1))\right). \qquad (j=1,2,\dots) \tag{9.23}$$

In the special case when H is symmetric around zero, we have

$$\mu_{H^{n*}}(j)=\mu_F(j)=\mu_{a\vee H}(j)=\mu_{\Psi b\vee H}(j)=0$$

for all integers n when j is odd so that in this case (9.20) gives

$$\mu_F(2j)=\frac{1}{1-\mu_a(0)}\Bigg(\mu_{q\vee H}(2j)$$

$$+\sum_{u=1}^{j}\binom{2j-1}{2u-1}\left(\frac{j}{u}\mu_{a\vee H}(2u)+\mu_{\Psi b\vee H}(2u)\right)\mu_F(2(j-u))\Bigg).$$

$$(j=1,2,\dots)$$

If p is $R_k[a,b]$, then (9.17) reduces to

$$\mu_F(j;c)=\frac{1}{1-\mu_a(0)}\Bigg(\sum_{u=1}^{j}\binom{j-1}{u-1}\left(\frac{j}{u}\mu_{a\vee H}(u)+\mu_{\Psi b\vee H}(u)+c\mu_{a\vee H}(u-1)\right)$$

$$\left.\times\,\mu_F(j-u;c)-c\mu_F(j-1;c)\right). \qquad (j=1,2,\dots) \tag{9.24}$$

With $k = \infty$, $a \equiv 0$, and $b = \varphi_p$, we obtain

$$\mu_F(j; c) = \sum_{u=1}^{j} \binom{j-1}{u-1} \mu_{\Psi\varphi_p \vee H}(u)\mu_F(j-u; c) - c\mu_F(j-1; c).$$
$$(j = 1, 2, \dots)$$

Let us now look at the case when H is concentrated in one, so that F reduces to p. In this case, (9.17) reduces to

$$\mu_p(j; c) = \frac{1}{1-\mu_a(0)} \Bigg(\mu_q(j; c) + c\mu_q(j-1; c) + \sum_{u=1}^{j} \binom{j-1}{u-1} \left(\frac{j}{u}\mu_a(u) + \mu_{b+ca}(u-1) \right) \mu_p(j-u; c) - c\mu_p(j-1; c) \Bigg). \qquad (j = 1, 2, \dots)$$

When p is $R_k[a, b]$, we obtain

$$\mu_p(j; c) = \frac{1}{1-\mu_a(0)} \Bigg(\sum_{u=1}^{j} \binom{j-1}{u-1} \left(\frac{j}{u}\mu_a(u) + \mu_{b+ca}(u-1) \right) \mu_p(j-u; c) - c\mu_p(j-1; c) \Bigg), \qquad (j = 1, 2, \dots) \tag{9.25}$$

and with $k = \infty$, $a \equiv 0$, and $b = \varphi_p$, we get

$$\mu_p(j; c) = \sum_{u=1}^{j} \binom{j-1}{u-1} \mu_{\varphi_p}(u-1)\mu_p(j-u; c) - c\mu_p(j-1; c). \qquad (j = 1, 2, \dots) \tag{9.26}$$

By application of Corollary 5.6 in (9.25) and some manipulation, we get back to (9.1).

Let us look at some examples.

Example 9.2 Let p be the logarithmic distribution Log(π). Then p satisfies (5.6) with $k = 1$, a and b given by (2.64), and

$$q(n) = p(1)I(n = 1) = \frac{\pi}{-\ln(1-\pi)} I(n = 1). \qquad (n = 1, 2, \dots)$$

Insertion in (9.17) gives that for $j = 1, 2, \ldots,$

$$\mu_F(j;c) = \frac{1}{1-\pi}\left(\frac{\pi}{-\ln(1-\pi)}(\mu_H(j;c) + c\mu_H(j-1;c)) - c\mu_F(j-1;c)\right.$$
$$\left. + \pi\sum_{u=1}^{j}\binom{j-1}{u-1}\left(\left(\frac{j}{u}-1\right)\mu_H(u) + c\mu_H(u-1)\right)\mu_F(j-u;c)\right).$$

In particular, we obtain

$$\mu_p(j;c) = \frac{1}{1-\pi}\left(\frac{(1+c)\pi}{-\ln(1-\pi)} - c\mu_p(j-1;c)\right.$$
$$\left. + \pi\sum_{u=1}^{j}\binom{j-1}{u-1}\left(\frac{j}{u}+c-1\right)\mu_p(j-u;c)\right).$$

□

9.2.2 Compound Panjer Distributions

In Sect. 9.2.2, we consider the special case when p is $R_1[a,b]$, that is, $q \equiv 0$ and $k = 1$.

Under this assumption, (9.17), (9.20), and (9.23) reduce to

$$\mu_F(j;c) = \frac{1}{1-a}\left(\sum_{u=1}^{j}\binom{j-1}{u-1}\left(\left(\frac{j}{u}a+b\right)\mu_H(u) + ca\mu_H(u-1)\right)\mu_F(j-u;c)\right.$$
$$\left. - c\mu_F(j-1;c)\right) \tag{9.27}$$

$$\mu_F(j) = \frac{1}{1-a}\sum_{u=1}^{j}\binom{j-1}{u-1}\left(\frac{j}{u}a+b\right)\mu_H(u)\mu_F(j-u) \tag{9.28}$$

$$\mu_F(j;\mu_F(1)) = \frac{1}{1-a}\left((j-1)a\mu_H(1)\mu_F(j-1;\mu_F(1))\right.$$
$$+ \sum_{u=2}^{j}\binom{j-1}{u-1}\left(\left(\frac{j}{u}a+b\right)\mu_H(u) + a\mu_F(1)\mu_H(u-1)\right)$$
$$\left. \times \mu_F(j-u;\mu_F(1))\right) \tag{9.29}$$

for $j = 1, 2, \ldots.$

When H is concentrated in one, F reduces to p and $\mu_H(u) = 1$ for all u. Thus, these recursions reduce to

$$\mu_p(j;c) = \frac{1}{1-a}\left(\sum_{u=1}^{j}\binom{j-1}{u-1}\left(\frac{j}{u}a+b+ca\right)\mu_p(j-u;c) - c\mu_p(j-1;c)\right) \tag{9.30}$$

$$\mu_p(j) = \frac{1}{1-a}\sum_{u=1}^{j}\binom{j-1}{u-1}\left(\frac{j}{u}a+b\right)\mu_p(j-u) \tag{9.31}$$

$$\begin{aligned}\mu_p(j;\mu_p(1)) = \frac{1}{1-a}\Bigg((j-1)a\mu_p(j-1;\mu_p(1)) \\ + \sum_{u=2}^{j}\binom{j-1}{u-1}\left(\frac{j}{u}a+b+a\mu_p(1)\right)\mu_p(j-u;\mu_p(1))\Bigg)\end{aligned} \tag{9.32}$$

for $j = 1, 2, \ldots$.

In our present case, (9.14) reduces to

$$\mathrm{E}\, Xr(X) = \mathrm{E}(a(X+Y_1)+bY_1)r(X+Y_1). \tag{9.33}$$

Letting $r(x)$ equal to s^x and e^{sx} gives respectively

$$\tau'_F(s) = a\tau_G(s)\tau'_F(s) + (a+b)\tau'_G(s)\tau_F(s) \tag{9.34}$$

$$\omega'_F(s) = a\omega_G(s)\omega'_F(s) + (a+b)\omega'_G(s)\omega_F(s). \tag{9.35}$$

We can obviously deduce (9.28) by taking the jth order derivative of (9.35) at zero and using (1.12). Analogously, we can find a recursion for the factorial moments of F by taking the jth order derivative of (9.34) at one and using (1.14). Because of the similarity of the formulae involved in these two deductions, the latter recursion is obtained by simply replacing the moments in (9.28) with the corresponding factorial moments, that is,

$$\nu_F(j) = \frac{1}{1-a}\sum_{u=1}^{j}\binom{j-1}{u-1}\left(\frac{j}{u}a+b\right)\nu_H(u)\nu_F(j-u). \tag{9.36}$$

When H is concentrated in one, we have $\nu_H(u) = I(u=1)$ for $u = 1, 2, \ldots$. Hence,

$$\nu_p(j) = \frac{ja+b}{1-a}\nu_p(j-1),$$

that is,

$$\nu_p(j) = \frac{\prod_{i=1}^{j}(ia+b)}{(1-a)^j}. \tag{9.37}$$

More generally, when p is $R_k[a,b]$, by applying (9.24), reasoning like we did for setting up (9.36), and assuming that $\nu_a(0) \neq 0$, we obtain

$$\nu_F(j) = \frac{1}{1-\nu_a(0)} \sum_{u=1}^{j} \binom{j-1}{u-1} \left(\frac{j}{u} \nu_{a \vee H}(u) + \nu_{\Psi b \vee H}(u) \right) \nu_F(j-u).$$
$$(j = 1, 2, \dots)$$

When H is concentrated in one, this reduces to

$$\nu_p(j) = \frac{1}{1-\nu_a(0)} \sum_{u=1}^{j} \binom{j-1}{u-1} \left(\frac{j}{u} \nu_a(u) + \nu_{\Psi b}(u) \right) \nu_p(j-u).$$
$$(j = 1, 2, \dots)$$

As $i^{(u)} = 0$ when $u > i$, we have $\nu_a(u) = \nu_{\Psi b}(u) = 0$ when $u > k$. Hence, in this case, we obtain

$$\nu_p(j) = \frac{1}{1-\nu_a(0)} \sum_{u=1}^{\min(j,k)} \binom{j-1}{u-1} \left(\frac{j}{u} \nu_a(u) + \nu_{\Psi b}(u) \right) \nu_p(j-u).$$
$$(j = 1, 2, \dots) \tag{9.38}$$

Let us apply the recursions (9.27)–(9.32) and (9.37) to the three main classes in $\mathcal{R}_1$ for $j = 1, 2, \dots$. We do not bother to display the recursion (9.36) as that is obtained from (9.28) by simply replacing the moments by factorial moments. The values of a and b are found in Table 2.1.

1. *Poisson distribution* $\text{Po}(\lambda)$.

$$\mu_F(j; c) = \lambda \sum_{u=1}^{j} \binom{j-1}{u-1} \mu_H(u) \mu_F(j-u; c) - c\mu_F(j-1; c) \tag{9.39}$$

$$\mu_F(j) = \lambda \sum_{u=1}^{j} \binom{j-1}{u-1} \mu_H(u) \mu_F(j-u) \tag{9.40}$$

$$\mu_F(j; \mu_F(1)) = \lambda \sum_{u=2}^{j} \binom{j-1}{u-1} \mu_H(u) \mu_F(j-u; \mu_F(1)) \tag{9.41}$$

$$\mu_p(j; c) = \lambda \sum_{u=1}^{j} \binom{j-1}{u-1} \mu_p(j-u; c) - c\mu_p(j-1; c)$$

$$\mu_p(j) = \lambda \sum_{u=1}^{j} \binom{j-1}{u-1} \mu_p(j-u) \tag{9.42}$$

$$\mu_p(j;\mu_p(1)) = \lambda \sum_{u=2}^{j} \binom{j-1}{u-1} \mu_p(j-u;\mu_p(1))$$

$$\nu_p(j) = \lambda^j.$$

The last formula was also given in (3.4).

2. *Binomial distribution* $\mathrm{bin}(M, \pi)$.

$$\mu_F(j;c) = \pi \sum_{u=1}^{j} \binom{j-1}{u-1} \left(\left(M + 1 - \frac{j}{u} \right) \mu_H(u) - c\mu_H(u-1) \right) \mu_F(j-u;c) - (1-\pi)c\mu_F(j-1;c) \tag{9.43}$$

$$\mu_F(j) = \pi \sum_{u=1}^{j} \binom{j-1}{u-1} \left(M + 1 - \frac{j}{u} \right) \mu_H(u)\mu_F(j-u)$$

$$\mu_F(j;\mu_F(1)) = \pi \left(\sum_{u=2}^{j} \binom{j-1}{u-1} \left(\left(M + 1 - \frac{j}{u} \right) \mu_H(u) - \mu_F(1)\mu_H(u-1) \right) \times \mu_F(j-u;\mu_F(1)) - (j-1)\mu_H(1)\mu_F(j-1;\mu_F(1)) \right)$$

$$\mu_p(j;c) = \pi \sum_{u=1}^{j} \binom{j-1}{u-1} \left(M + 1 - \frac{j}{u} - c \right) \mu_p(j-u;c) - (1-\pi)c\mu_p(j-1;c)$$

$$\mu_p(j) = \pi \sum_{u=1}^{j} \binom{j-1}{u-1} \left(M + 1 - \frac{j}{u} \right) \mu_p(j-u)$$

$$\mu_p(j;\mu_p(1)) = \pi \left(\sum_{u=2}^{j} \binom{j-1}{u-1} \left(M + 1 - \frac{j}{u} - \mu_p(1) \right) \mu_p(j-u;\mu_p(1)) - (j-1)\mu_p(j-1;\mu_p(1)) \right)$$

$$\nu_p(j) = M^{(j)}\pi^j.$$

3. *Negative binomial distribution* $\mathrm{NB}(\alpha, \pi)$.

$$\mu_F(j;c) = \frac{1}{1-\pi} \left(\pi \sum_{u=1}^{j} \binom{j-1}{u-1} \left(\left(\frac{j}{u} + \alpha - 1 \right) \mu_H(u) + c\mu_H(u-1) \right) \mu_F(j-u;c) - c\mu_F(j-1;c) \right)$$

$$\mu_F(j) = \frac{\pi}{1-\pi} \sum_{u=1}^{j} \binom{j-1}{u-1} \left(\frac{j}{u} + \alpha - 1\right) \mu_H(u) \mu_F(j-u)$$

$$\mu_F(j;\mu_F(1)) = \frac{\pi}{1-\pi} \Bigg((j-1)\mu_H(1)\mu_F(j-1;\mu_F(1)) \\ + \sum_{u=2}^{j} \binom{j-1}{u-1} \left(\left(\frac{j}{u} + \alpha - 1\right) \mu_H(u) + \mu_F(1)\mu_H(u-1) \right) \\ \times \mu_F(j-u;\mu_F(1)) \Bigg)$$

$$\mu_p(j;c) = \frac{1}{1-\pi} \Bigg(\pi \sum_{u=1}^{j} \binom{j-1}{u-1} \left(\frac{j}{u} + \alpha + c - 1\right) \mu_p(j-u;c) \\ - c\mu_p(j-1;c) \Bigg)$$

$$\mu_p(j) = \frac{\pi}{1-\pi} \sum_{u=1}^{j} \binom{j-1}{u-1} \left(\frac{j}{u} + \alpha - 1\right) \mu_p(j-u)$$

$$\mu_p(j;\mu_p(1)) = \frac{\pi}{1-\pi} \Bigg((j-1)\mu_p(j-1;\mu_p(1)) \\ + \sum_{u=2}^{j} \binom{j-1}{u-1} \left(\frac{j}{u} + \alpha - 1 + \mu_p(1)\right) \mu_p(j-u;\mu_p(1)) \Bigg)$$

$$\nu_p(j) = (\alpha + j - 1)^{(j)} \left(\frac{\pi}{1-\pi}\right)^{j}.$$

9.2.3 Compound Poisson Distributions

Let $f = p \vee h$ where p is the Poisson distribution $\text{Po}(\lambda)$ and $h \in \mathcal{P}_{10}$ satisfies the conditions of Theorem 2.3. Application of Theorem 5.5 in (9.25) and (9.38) gives that for $j = 1, 2, \ldots,$

$$\mu_f(j) = \frac{1}{1-\mu_\chi(0)} \sum_{u=1}^{j} \binom{j-1}{u-1} \left(\lambda\mu_\eta(u-1) + \frac{j-u}{u}\mu_\chi(u) \right) \mu_f(j-u) \tag{9.44}$$

$$\nu_f(j) = \frac{1}{1-\nu_\chi(0)} \sum_{u=1}^{\min(j,r)} \binom{j-1}{u-1} \left(\lambda \nu_{\Psi\eta}(u) + \frac{j-u}{u} \nu_\chi(u) \right) \nu_f(j-u). \tag{9.45}$$

In (9.45) we can stop the summation at r. Unfortunately, this is not the case with (9.44); there we have to sum to j. Hence, if the moments of h are known, then that recursion seems more complicated than the general recursion (9.40), and we therefore discard it. This was also the reason that we did not bother to state that recursion more generally for $\mu_f(j;c)$.

As pointed out after the proof of Theorem 2.3, the conditions of that theorem are always satisfied with $r=\infty$, $\eta = \Phi h$, and $\chi \equiv 0$. Insertion in (9.45) gives the special case of (9.36) corresponding to (9.42).

Example 9.3 Let $h \in \mathcal{P}_{11}$ be the shifted geometric distribution given by (2.21). Then h satisfies the conditions of Theorem 2.3 with r, η, and χ given by (2.23), so that

$$\nu_{\Psi\eta}(1) = 1-\pi; \quad \nu_\chi(0) = \pi(2-\pi); \quad \nu_\chi(1) = 2\pi(1-\pi); \quad \nu_\chi(2) = -2\pi^2,$$

and $\nu_\eta(u)$ and $\nu_\chi(u)$ equal to zero for all other values of u. Insertion in (9.45) gives

$$\begin{aligned}\nu_f(j) = \frac{1}{1-\pi}\bigg(&(\lambda + 2(j-1)\pi)\nu_f(j-1) \\ &- (j-1)^{(2)} \frac{\pi^2}{1-\pi} \nu_f(j-2) I(j \geq 2)\bigg). \qquad (j = 1, 2, \ldots)\end{aligned}$$

□

Further Remarks and References

Section 9.1 is to a large extent based on Sundt (2003a). A more general version of Theorem 9.2 was proved by Shimizu (1962) by using cumulant generating functions. For more information on the normal distribution, see e.g. Johnson et al. (1994).

Section 9.2 is primarily based on Sundt (2003b).

The recursion (9.24) was deduced by Murat and Szynal (1998). With $c=0$ and $H \in \mathcal{P}_{11}$, Sundt et al. (1998) deduced the recursions (9.24)–(9.26).

The recursions (9.27)–(9.29) were deduced by De Pril (1986b), using moment generating functions. Murat and Szynal (2000a) presented a more general version of (9.27). The special case of moments of the counting distribution was discussed by Szynal and Teugels (1995), who also deduced another recursion for these moments. Goovaerts et al. (1984) deduced the recursion (9.41), and Shiu (1977) gave an explicit expression for central moments of a compound Poisson distribution. Recursions for moments of compound distributions are also discussed by Gerhold et al. (2008).

For the Poisson case, the moment relation (9.33) was used by Ross (1996).

For $H \in \mathcal{P}_{11}$, deduction of the recursion (9.36) was given as a question in Exercise 11.13 of Sundt (1999b).

Recursions for moments of counting distributions belonging to other classes have been presented by Murat and Szynal (1998, 2000a) and for the counting distribution by Murat and Szynal (2000b).

Kaas and Goovaerts (1985) presented a recursion for the non-central moments of compound distributions in terms of the factorial moments of the counting distribution and the non-central moments of the severity distribution. This recursion is applicable for any type of counting and severity distribution.

The recursions (9.44) and (9.45) and Example 9.3 were presented by Chadjiconstantinidis and Pitselis (2008), who also gave other recursions for more general classes of moments within the same setting.

Chadjiconstantinidis and Antzoulakos (2002) presented recursions for moments of mixed Poisson distributions and compound mixed Poisson distributions with mixing distribution in the Willmot class.

Chapter 10
Approximations Based on De Pril Transforms

Summary

In Chap. 7, we pointed out that because of the need for all the convolutions of the h_is in (7.11), De Pril's second method can be rather inefficient for exact results in practice, but that it has been used as basis for developing approximations where we replace each p_j with a function $p_j^{(r)} \in \mathcal{F}_{1\underline{0}}$ with $\varphi_{p_j^{(r)}}(y) = 0$ for all integers y greater than some positive integer r. The present chapter is devoted to such approximations.

As $p_j^{(r)}$ is not necessarily a distribution, we need to extend the definition of the De Pril transform to functions in $\mathcal{F}_{1\underline{0}}$. This is done in Sect. 10.1, where we also introduce an error measure for the quality of approximations in $\mathcal{F}_{1\underline{0}}$ to distributions in $\mathcal{P}_{1\underline{0}}$. Furthermore, we briefly define the approximations of De Pril, Kornya, and Hipp.

When studying approximations in $\mathcal{F}_{1\underline{0}}$ of distributions in $\mathcal{P}_{1\underline{0}}$, we need to extend results that we have proved for distributions in $\mathcal{P}_{1\underline{0}}$ to functions in $\mathcal{F}_{1\underline{0}}$. Such extension is performed in Sect. 10.2.

In Sect. 10.3, we deduce upper bounds for our error measure for approximations.

Section 10.4 is devoted to a generalisation of De Pril's individual model, extending the class of the p_js from the Bernoulli class to $\mathcal{P}_{1\underline{0}}$. Within that model, we study the approximations of De Pril, Kornya, and Hipp in Sects. 10.5, 10.6, and 10.7 respectively. We deduce expressions for the approximations and give upper bounds for the error measure.

Finally, in Sect. 10.8, we present a numerical example.

10.1 Introduction

We now return to the setting of Sect. 7.1. To avoid evaluating all the convolutions of the h_is in (7.11), we want to apply an approximation where we replace each p_j with a function $p_j^{(r)} \in \mathcal{F}_{1\underline{0}}^{(r)}$, the class of functions $p \in \mathcal{F}_{1\underline{0}}$ for which $\varphi_p(y) = 0$ for all integers y greater than some positive integer r. We have $\mathcal{F}_{1\underline{0}}^{(r-1)} \subset \mathcal{F}_{1\underline{0}}^{(r)}$ for $r = 2, 3, \dots$.

Although p_j is a distribution, that does not need to be the case with $p_j^{(r)}$. As indicated above and in Chap. 7, we need the De Pril transform of $p_j^{(r)}$, but we have defined the De Pril transform only for distributions in $\mathcal{P}_{1\underline{0}}$. Hence, we need to extend the definition of the De Pril transform to a more general class of functions. We also touched the desirability of such an extension in the discussion to Corollary 8.6.

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_10, © Springer-Verlag Berlin Heidelberg 2009

There we suggested to extend the definition (6.1) of the De Pril transform to functions in $\mathcal{F}_{1\underline{0}}$, and that is what we are going to do, that is, we define the De Pril transform φ_f of a function $f \in \mathcal{F}_{1\underline{0}}$ by

$$\varphi_f(x) = \frac{1}{f(0)}\left(xf(x) - \sum_{y=1}^{x-1} \varphi_f(y) f(x-y)\right). \qquad (x = 1, 2, \dots) \tag{10.1}$$

Then each of the relations (6.2) and (6.3) still holds and determines φ_f uniquely.

We have earlier deduced several results for De Pril transforms of distributions in $\mathcal{P}_{1\underline{0}}$. Now we have to check to what extent these results still hold for, or could be extended to, functions in $\mathcal{F}_{1\underline{0}}$. Many of the results were presented as special cases of results that we had deduced for distributions in the form $R_k[a, b]$ by letting $k = \infty$ and $a \equiv 0$. For distributions in $\mathcal{P}_{1\underline{0}}$, the functions a and b determined the distribution uniquely. This was because a distribution sums to one. This is not necessarily the case with functions in $\mathcal{F}_{1\underline{0}}$, so such a function is determined by a and b only up to a multiplicative constant. We could say that a distribution in $\mathcal{P}_{1\underline{0}}$ was $R_k[a, b]$, but for a function in $\mathcal{F}_{1\underline{0}}$, we can only say that it is in the form $R_k[a, b]$, by which we mean that it satisfies the recursion (5.12); if a function in $\mathcal{F}_{1\underline{0}}$ is in the form $R_k[a, b]$, then that is the case for all functions in $\mathcal{F}_{1\underline{0}}$ that are proportional to it. Unfortunately, that also goes for De Pril transforms of functions in $\mathcal{F}_{1\underline{0}}$; each function in $\mathcal{F}_{1\underline{0}}$ has a unique De Pril transform, but all functions in $\mathcal{F}_{1\underline{0}}$ proportional to it have that De Pril transform too. The additional information we need for determining the function uniquely, will typically be its value at zero, that is, the initial value for the recursion (6.2).

Of course, to any function from the set of non-negative integers to the set of real numbers, any other function between the same two sets can be considered as an approximation, but it would not necessarily be a good approximation. But what do we mean by a good approximation? What do we mean when we say that one approximation is better than another approximation? To say that, we need a measure of the quality of approximations; we shall call such a measure an *error measure*. Let ε denote such a measure. Relative to this measure, we say that an approximation $\hat{f}$ of a function f is better than another approximation $\hat{\hat{f}}$ if $\varepsilon(f, \hat{f}) < \varepsilon(f, \hat{\hat{f}})$. Our choice of ε is

$$\varepsilon(f, \hat{f}) = \mu_{|f-\hat{f}|}(0) = \sum_{x=0}^{\infty} |f - \hat{f}|(x). \qquad (f, \hat{f} \in \mathcal{F}_{10}) \tag{10.2}$$

This measure satisfies the desirable property that $\varepsilon(f, f) = 0$.

The three most well-known classes of approximations based on approximating De Pril transforms are the approximations of De Pril, Kornya, and Hipp. The De Pril approximation is presumably the simplest and the most intuitive of these. There we simply replace $\varphi_{p_j}(x)$ with zero for all integers x greater than r and let $p_j^{(r)}(0) = p_j(0)$, or, equivalently, we determine $p_j^{(r)} \in \mathcal{F}_{1\underline{0}}^{(r)}$ such that $p_j^{(r)}(n) = p_j(n)$ for $n = 0, 1, 2, \dots, r$. The Kornya approximation has the same

De Pril transform, but now $p_j^{(r)}(0)$ is determined such that $p_j^{(r)}$ sums to one like a probability distribution. In that sense, there is a tiny touch of moment-matching in the Kornya approximation; the zeroth order moment of the approximation should match the zeroth order moment of the exact distribution. In the Hipp approximation, the element of moment-matching is more pronounced; here the approximation is determined such that the moments of order $0, 1, 2, \dots, r$ of the approximation match the corresponding moments of the exact distribution.

10.2 Extension of Results for Distributions

10.2.1 Key Result

When extending the results for distributions to more general functions, it is natural to pose the following questions:

1. What properties of distributions that do not hold for functions, do we use?
2. Where do we use them?
3. How do we use them?

We sometimes use that a distribution is non-negative and sums to one, but the relation (5.4) is more essential; if we can extend that relation from distributions in $\mathcal{P}_{10}$ to functions in $\mathcal{F}_{10}$, then the proofs of most of the results will still hold. Obviously, for functions in $\mathcal{F}_{10}$, we cannot express the relation in terms of conditional expectations, so let us now express it in terms of distributions. Let $Y_1, Y_2, \dots$ be independent and identical distributed random variables with distribution $h \in \mathcal{P}_{10}$. Then we can express (5.4) as

$$\frac{\sum_{y=0}^{x}(a + \frac{b}{i}\frac{y}{x})h^{i*}(y)h^{(n-i)*}(x-y)}{h^{n*}(x)} = a + \frac{b}{n}$$

for $x = 1, 2, \dots; n = i, i+1, i+2, \dots,$ and $i = 1, 2, \dots$. Strictly speaking, we have to be a bit careful here as the denominator $h^{n*}(x)$ could be equal to zero. However, in the deductions we made in earlier chapters, that did not do any harm as the relation always appeared in the context of $a + b/n$ being multiplied with $h^{n*}(x)$. To be safe, we multiply the relation with $h^{n*}(x)$ and obtain

$$\sum_{y=0}^{x}\left(a + \frac{b}{i}\frac{y}{x}\right)h^{i*}(y)h^{(n-i)*}(x-y) = \left(a + \frac{b}{n}\right)h^{n*}(x).$$

This relation can be split into the two relations

$$\sum_{y=0}^{x} h^{i*}(y)h^{(n-i)*}(x-y) = h^{n*}(x)$$

$$\frac{n}{i}\sum_{y=0}^{x} yh^{i*}(y)h^{(n-i)*}(x-y) = xh^{n*}(x).$$

The first one follows immediately from the definition of the convolution between h^{i*} and $h^{(n-i)*}$. It remains to prove the second one, which we reformulate as

$$\Phi h^{n*} = \frac{n}{i}\Phi h^{i*} * h^{(n-i)*}. \tag{10.3}$$

To prove that this relation holds more generally for all $h \in \mathcal{F}_{10}$, we shall need the following lemma.

Lemma 10.1 *If $f, g \in \mathcal{F}_{10}$, then $\Phi(f * g) = \Phi f * g + f * \Phi g$.*

Proof For any non-negative integer x, we have

$$\begin{aligned}
\Phi(f * g)(x) = x(f * g)(x) &= x\sum_{y=0}^{x} f(y)g(x-y) \\
&= \sum_{y=0}^{x} yf(y)g(x-y) + \sum_{y=0}^{x} f(y)(x-y)g(x-y) \\
&= \sum_{y=0}^{x} \Phi f(y)g(x-y) + \sum_{y=0}^{x} f(y)\Phi g(x-y) \\
&= (\Phi f * g)(x) + (f * \Phi g)(x) = (\Phi f * g + f * \Phi g)(x),
\end{aligned}$$

which proves the lemma. □

We are now ready to prove our key result.

Theorem 10.1 *For $h \in \mathcal{F}_{10}$, the relation* (10.3) *holds for $n = 1, 2, \dots$ and $i = 1, 2, \dots, n$.*

Proof For $i = 1$, (10.3) reduces to

$$\Phi h^{n*} = n\Phi h * h^{(n-1)*}. \tag{10.4}$$

We shall prove this relation by induction on n. It obviously holds for $n = 1$. Let us now assume that it holds for $n = j - 1$ for some integer $j > 1$. Then, by using Lemma 10.1 and the induction hypothesis, we obtain

$$\begin{aligned}
\Phi h^{j*} &= \Phi(h * h^{(j-1)*}) = \Phi h * h^{(j-1)*} + h * \Phi h^{(j-1)*} \\
&= \Phi h * h^{(j-1)*} + h * ((j-1)\Phi h * h^{(j-2)*}) = j\Phi h * h^{(j-1)*},
\end{aligned}$$

that is, (10.4) holds also for $n = j$, and by induction it holds for all positive integers n.

Let us now prove (10.3) for $i = 2, 3, \ldots, n$. By using (10.4) twice, we obtain

$$\Phi h^{n*} = n\Phi h * h^{(n-1)*} = \frac{n}{i}(i\Phi h * h^{(i-1)*}) * h^{(n-i)*} = \frac{n}{i}\Phi h^{i*} * h^{(n-i)*},$$

which completes the proof of Theorem 10.1. □

10.2.2 Applications

Theorem 10.1 gives us the key to extend to functions in $\mathcal{F}_{10}$ many of the results that we have deduced for distributions in $\mathcal{P}_{10}$. We shall not go through all of these. However, as we shall need the extension of Corollaries 6.1 and 6.3 in the following, we shall now prove these results.

Theorem 10.2 *The convolution of a finite number of functions in $\mathcal{F}_{1\underline{0}}$ is a function in $\mathcal{F}_{1\underline{0}}$, and its De Pril transform is the sum of the De Pril transforms of these functions.*

Proof We prove the theorem for the convolution of two functions $f, g \in \mathcal{F}_{1\underline{0}}$; the general case follows by induction.

As $f, g \in \mathcal{F}_{1\underline{0}}$, $f(0) > 0$ and $g(0) > 0$. Hence, $(f * g)(0) = f(0)g(0) > 0$, so that $f * g \in \mathcal{F}_{1\underline{0}}$.

By using Lemma 10.1 and (6.3), we obtain

$$\Phi(f * g) = \Phi f * g + f * \Phi g = (\varphi_f * f) * g + f * (\varphi_g * g) = (\varphi_f + \varphi_g) * (f * g).$$

As the relation (6.3) determines the De Pril transform of a function in $\mathcal{F}_{1\underline{0}}$ uniquely, we must have $\varphi_{f*g} = \varphi_f + \varphi_g$. This proves that the De Pril transform of the convolution of two functions in $\mathcal{F}_{1\underline{0}}$ is the sum of the De Pril transforms of these two functions.

This completes the proof of Theorem 10.2. □

Theorem 10.3 *If $p \in \mathcal{F}_{1\underline{0}}$ and $h \in \mathcal{F}_{11}$, then*

$$\varphi_{p \vee h} = \Phi(\Psi\varphi_p \vee h). \tag{10.5}$$

Proof For $x = 1, 2, \ldots$, application of (1.8), (6.2), and (10.3) gives

$$\Phi(p \vee h)(x) = \sum_{n=1}^{x} p(n)\Phi h^{n*}(x) = \sum_{n=1}^{x} \frac{1}{n}\left(\sum_{i=1}^{n} \varphi_p(i)p(n-i)\right)\Phi h^{n*}(x)$$

$$= \sum_{n=1}^{x}\sum_{i=1}^{n} \frac{\varphi_p(i)}{i} p(n-i)(\Phi h^{i*} * h^{(n-i)*})(x)$$

$$= \sum_{n=1}^{x} \sum_{i=1}^{n} \Psi\varphi_p(i) p(n-i) \sum_{y=1}^{x} \Phi h^{i*}(y) h^{(n-i)*}(x-y)$$

$$= \sum_{y=1}^{x} \sum_{i=1}^{x} \Psi\varphi_p(i) \Phi h^{i*}(y) \sum_{n=i}^{x} p(n-i) h^{(n-i)*}(x-y)$$

$$= \sum_{y=1}^{x} \Phi(\Psi\varphi_p \vee h)(y)(p \vee h)(x-y) = (\Phi(\Psi\varphi_p \vee h) * (p \vee h))(x).$$

Comparison with (6.3) gives (10.5). □

10.3 Error Bounds

10.3.1 Main Result

The error measure ε can be considered as a measure for the distance between two functions in $\mathcal{F}_{1\underline{0}}$. We shall also need the distance measure δ given by

$$\delta(f, \hat{f}) = \left|\ln \frac{\hat{f}(0)}{f(0)}\right| + \sum_{x=1}^{\infty} \frac{|\varphi_f - \varphi_{\hat{f}}|(x)}{x} = \left|\ln \frac{\hat{f}(0)}{f(0)}\right| + \mu_{|\varphi_f - \varphi_{\hat{f}}|}(-1).$$
$$(f, \hat{f} \in \mathcal{F}_{10}) \tag{10.6}$$

Like ε, this measure satisfies the desirable property that $\delta(f, \hat{f}) = 0$. The following theorem gives a relation between these two distance measures.

Theorem 10.4 *If $f \in \mathcal{P}_{1\underline{0}}$ and $\hat{f} \in \mathcal{F}_{1\underline{0}}$, then*

$$\varepsilon(f, \hat{f}) \leq \mathrm{e}^{\delta(f, \hat{f})} - 1. \tag{10.7}$$

Proof As the theorem trivially holds when $\delta(f, \hat{f}) = \infty$, we assume in the following that $\delta(f, \hat{f}) < \infty$.

Let $g \in \mathcal{F}_{1\underline{0}}$ be defined by

$$\varphi_g = \varphi_{\hat{f}} - \varphi_f; \qquad g(0) = \frac{\hat{f}(0)}{f(0)}. \tag{10.8}$$

Then Theorem 10.2 gives that $\hat{f} = f * g$, and we obtain

$$(f - \hat{f})(x) = (1 - g(0)) f(x) - \sum_{y=1}^{x} g(y) f(x-y). \quad (x = 0, 1, 2, \dots) \tag{10.9}$$

Thus,

$$\begin{aligned}\varepsilon(f,\hat{f}) &= \sum_{x=0}^{\infty}\left|(1-g(0))f(x)-\sum_{y=1}^{x}g(y)f(x-y)\right| \\ &\le \sum_{x=0}^{\infty}\left(|1-g(0)|f(x)+\sum_{y=1}^{x}|g(y)|f(x-y)\right) \\ &= |1-g(0)|+\sum_{y=1}^{\infty}|g(y)|\sum_{x=y}^{\infty}f(x-y),\end{aligned}$$

from which we obtain

$$\varepsilon(f,\hat{f}) \le |1-g(0)|+\sum_{y=1}^{\infty}|g(y)|. \tag{10.10}$$

Let k be a compound Poisson distribution with Poisson parameter

$$\lambda = \delta(f,\hat{f}) - |\ln g(0)| = \sum_{y=1}^{\infty}\frac{|\varphi_g(y)|}{y}$$

and severity distribution $h \in \mathcal{P}_{11}$ given by

$$h(y) = \frac{|\varphi_g(y)|}{\lambda y}. \qquad (y = 1, 2, \dots)$$

From (4.13), we see that $\varphi_k = |\varphi_g|$.

We shall prove by induction that

$$|g(x)| \le k(x)\mathrm{e}^{\delta(f,\hat{f})}. \qquad (x = 0, 1, 2, \dots) \tag{10.11}$$

We have

$$|g(0)| = g(0) = \mathrm{e}^{\ln g(0)} \le \mathrm{e}^{|\ln g(0)|} = \mathrm{e}^{-\lambda+\delta(f,\hat{f})} = k(0)\mathrm{e}^{\delta(f,\hat{f})},$$

so (10.11) holds for $x = 0$. Let us now assume that it holds for $x = 0, 1, 2, \dots, z-1$ for some positive integer z. Then, by application of (6.2), we obtain

$$\begin{aligned}|g(z)| &= \left|\frac{1}{z}\sum_{y=1}^{z}\varphi_g(y)g(z-y)\right| \le \frac{1}{z}\sum_{y=1}^{z}\varphi_k(y)|g(z-y)| \\ &\le \frac{1}{z}\sum_{y=1}^{z}\varphi_k(y)k(z-y)\mathrm{e}^{\delta(f,\hat{f})} = k(z)\mathrm{e}^{\delta(f,\hat{f})}.\end{aligned}$$

Hence, (10.11) holds also for $x = z$, and by induction it holds for all non-negative integers z.

From (10.11), we obtain

$$\begin{aligned}|1 - g(0)| + \sum_{y=1}^{\infty} |g(y)| &\le |1 - e^{\ln g(0)}| + \sum_{x=1}^{\infty} k(x) e^{\delta(f,\hat{f})} \\ &\le e^{|\ln g(0)|} - 1 + (1 - k(0)) e^{\delta(f,\hat{f})} \\ &= e^{\delta(f,\hat{f})-\lambda} - 1 + (1 - e^{-\lambda}) e^{\delta(f,\hat{f})},\end{aligned}$$

that is,

$$|1 - g(0)| + \sum_{y=1}^{\infty} |g(y)| \le e^{\delta(f,\hat{f})} - 1, \tag{10.12}$$

and together with (10.10), this gives (10.7).

This completes the proof of Theorem 10.4. □

As proportional functions in $\mathcal{F}_{1\underline{0}}$ have the same De Pril transform, (10.6) gives that

$$\delta(f, c\hat{f}) = \left| \ln \frac{c\hat{f}(0)}{f(0)} \right| + \mu_{|\varphi_f - \varphi_{\hat{f}}|}(-1)$$

for any positive constant c. From this, we see that $\delta(f, c\hat{f})$ and, hence, the upper bound in (10.7) is minimised when $c = f(0)/\hat{f}(0)$, that is, when we scale the approximation $\hat{f}$ such that $\hat{f}(0) = f(0)$. This does not appeal to our intuition; it does not seem reasonable that whatever approximation $\hat{f}$ we apply for f, it will always be best to scale it such that $\hat{f}(0) = f(0)$. This property is a deficiency of the error bound of Theorem 10.4. In particular, Theorem 10.4 will always give a lower value for the bound for the De Pril approximation than for the Kornya approximation. However, as it is only an upper bound for $\varepsilon(f, \hat{f})$, that does not necessarily imply that the De Pril approximation is always better than the Kornya approximation relative to the error measure ε. Let us look at a simple example.

Example 10.1 Let $f \in \mathcal{P}_{1\underline{0}}$ with $f(0) < 1/2$. We define $\hat{f} \in \mathcal{F}_{1\underline{0}}$ by

$$\hat{f}(x) = \begin{cases} cf(0) & (x = 0) \\ f(x) & (x = 1, 2, \dots) \end{cases}$$

with $1 < c < 1/f(0) - 1$ and let $\hat{\hat{f}} = \hat{f}/c$. Then

$$\delta(\varphi_f, \varphi_{\hat{f}}) = |\ln c| + \delta(\varphi_f, \varphi_{\hat{\hat{f}}}) > \delta(\varphi_f, \varphi_{\hat{\hat{f}}}).$$

Hence, Theorem 10.4 will give a higher value of the upper bound for $\varepsilon(f, \hat{f})$ than for $\varepsilon(f, \hat{\hat{f}})$. However,

$$\frac{\varepsilon(f, \hat{f})}{\varepsilon(f, \hat{\hat{f}})} = \frac{(c-1)f(0)}{(1-1/c)(1-f(0))} = c\frac{f(0)}{1-f(0)} < 1,$$

that is, $\varepsilon(f, \hat{f}) < \varepsilon(f, \hat{\hat{f}})$, so relative to the error measure ε, f is better approximated by $\hat{f}$ than by $\hat{\hat{f}}$ although $\delta(\varphi_f, \varphi_{\hat{f}}) > \delta(\varphi_f, \varphi_{\hat{\hat{f}}})$. □

10.3.2 The Dhaene–De Pril Transform

For any $f \in \mathcal{F}_{1\underline{0}}$, the *Dhaene–De Pril transform* $\psi_f \in \mathcal{F}_{10}$ is defined by

$$\psi_f(x) = \begin{cases} \ln f(0) & (x = 0) \\ \varphi_f(x)/x. & (x = 1, 2, \dots) \end{cases}$$

This gives $\delta(f, \hat{f}) = \varepsilon(\psi_f, \psi_{\hat{f}})$ so that (10.7) can be written as

$$\varepsilon(f, \hat{f}) \leq \mathrm{e}^{\varepsilon(\psi_f, \psi_{\hat{f}})} - 1.$$

Furthermore, (6.6) and (6.11) can be merged into the more pleasant shape

$$\kappa_p(j) = \mu_{\psi_p}(j) \qquad (j = 0, 1, 2, \dots) \tag{10.13}$$

in terms of the Dhaene–De Pril transform. Hence, we see that in connection with approximations, the Dhaene–De Pril transform has some advantages compared to the De Pril transform:

1. We do not need to introduce the distance measure δ; it suffices to use ε.
2. The connection to cumulants becomes simpler.
3. A function in $\mathcal{F}_{1\underline{0}}$ is uniquely determined by its Dhaene–De Pril transform.

Re Advantage 1: Construction of error bounds was not our main reason for introducing the De Pril transform. Our primary reason was to deduce a simple algorithm for recursive evaluation of the convolution of a finite number of functions in $\mathcal{F}_{1\underline{0}}$ by applying the recursions (6.2) and (6.1) as described in Sect. 4.4.1, and these recursions are more complicated in terms of the Dhaene–De Pril transform.

Re Advantage 2: Our main application of (10.13) is in connection with the Hipp approximation. However, as we shall see in Sect. 10.7, the relation for $j = 0$ gets a special treatment anyhow, and in that setting it is more convenient to have it in its old shape.

Re Advantage 3: As we have not defined $\varphi_f(0)$, we could also have defined that as e.g. $\ln f(0)$. However, it does not seem that urgent to have a function in $\mathcal{F}_{1\underline{0}}$

uniquely determined by its De Pril transform. In Chap. 11, we shall see that when extending the definition of the De Pril transform to functions in $\mathcal{F}_{1\_}$, then it will be convenient to define $\varphi_f(0) = 0$ for functions in $\mathcal{F}_{1\underline{0}}$.

Let us also point out that in terms of Dhaene–De Pril transforms, the relation (10.5) for compound distributions reduces to $\psi_{p\vee h} = \psi_p \vee h$.

10.3.3 Corollaries to the Main Result

We shall prove some corollaries to Theorem 10.4.

Corollary 10.1 *If* $f \in \mathcal{P}_{1\underline{0}}$ *and* $\hat{f} \in \mathcal{F}_{1\underline{0}}$, *then*

$$|\Gamma f - \Gamma \hat{f}| \leq (e^{\delta(f,\hat{f})} - 1)\Gamma f \leq e^{\delta(f,\hat{f})} - 1. \tag{10.14}$$

Proof With g defined by (10.8), we obtain that for $x = 0, 1, 2, \dots,$

$$\begin{aligned}(\Gamma f - \Gamma \hat{f})(x) &= \sum_{y=0}^{x}(f - \hat{f})(y) = \sum_{y=0}^{x}\left(f(y) - \sum_{z=0}^{y} g(z) f(y-z)\right) \\ &= \Gamma f(x) - \sum_{z=0}^{x} g(z) \sum_{y=z}^{x} f(y-z) = \Gamma f(x) - \sum_{z=0}^{x} g(z)\Gamma f(x-z) \\ &= (1 - g(0))\Gamma f(x) - \sum_{z=1}^{x} g(z)\Gamma f(x-z),\end{aligned}$$

so that

$$\begin{aligned}|\Gamma f - \Gamma \hat{f}|(x) &\leq |1 - g(0)|\Gamma f(x) + \sum_{z=1}^{x} |g(z)|\Gamma f(x-z) \\ &\leq \left(|1 - g(0)| + \sum_{z=1}^{\infty} |g(z)|\right)\Gamma f(x).\end{aligned}$$

Application of (10.12) gives the first inequality in (10.14).

As Γf is a cumulative distribution function, $\Gamma f \leq 1$, and this gives the second inequality in (10.14).

This completes the proof of Corollary 10.1. □

The following corollary gives bounds for Γf.

Corollary 10.2 *If* $f \in \mathcal{P}_{1\underline{0}}$ *and* $\hat{f} \in \mathcal{F}_{1\underline{0}}$, *then*

$$\Gamma f \geq e^{-\delta(f,\hat{f})}\Gamma \hat{f}. \tag{10.15}$$

If we also have $\delta(f,\hat{f}) < \ln 2$, *then*

$$\Gamma f \le \frac{\Gamma \hat{f}}{2 - \mathrm{e}^{\delta(f,\hat{f})}}. \tag{10.16}$$

Proof By application of the first inequality in (10.14), we get

$$\Gamma \hat{f} - \Gamma f \le (\mathrm{e}^{\delta(f,\hat{f})} - 1)\Gamma f,$$

that is,

$$\mathrm{e}^{\delta(f,\hat{f})}\Gamma f \ge \Gamma \hat{f}, \tag{10.17}$$

from which we obtain (10.15). Furthermore, we have

$$\Gamma f - \Gamma \hat{f} \le (\mathrm{e}^{\delta(f,\hat{f})} - 1)\Gamma f,$$

that is,

$$(2 - \mathrm{e}^{\delta(f,\hat{f})})\Gamma f \le \Gamma \hat{f},$$

from which we obtain (10.16) when $\delta(f,\hat{f}) < \ln 2$.

This completes the proof of Corollary 10.2. □

As $\lim_{x\uparrow\infty} \Gamma f(x) = 1$, (10.17) gives that $\overline{\lim}_{x\uparrow\infty} \Gamma \hat{f}(x) \le \mathrm{e}^{\delta(f,\hat{f})}$.
As $\Gamma f \ge 0$, (10.16) gives that $\Gamma \hat{f} \ge 0$ when $\delta(f,\hat{f}) < \ln 2$.
When $\delta(f,\hat{f}) < \ln 2$ and $\Gamma \hat{f}(x) > 0$, Corollary 10.2 gives that

$$\mathrm{e}^{-\delta(f,\hat{f})} \le \frac{\Gamma f(x)}{\Gamma \hat{f}(x)} \le \frac{1}{2 - \mathrm{e}^{\delta(f,\hat{f})}}. \tag{10.18}$$

Unfortunately, to apply the first inequality in (10.14), we need to know Γf, that is, the quantity that we want to approximate. The following corollary avoids that problem.

Corollary 10.3 *If* $f \in \mathcal{P}_{1\underline{0}}$ *and* $\hat{f} \in \mathcal{F}_{1\underline{0}}$ *with* $\delta(f,\hat{f}) < \ln 2$, *then*

$$|\Gamma f - \Gamma \hat{f}| \le \frac{\mathrm{e}^{\delta(f,\hat{f})} - 1}{2 - \mathrm{e}^{\delta(f,\hat{f})}} \Gamma \hat{f}. \tag{10.19}$$

Proof By Corollary 10.1, we obtain that

$$\begin{aligned} |\Gamma f - \Gamma \hat{f}| &\le (\mathrm{e}^{\delta(f,\hat{f})} - 1)(\Gamma f - \Gamma \hat{f} + \Gamma \hat{f}) \\ &\le (\mathrm{e}^{\delta(f,\hat{f})} - 1)(|\Gamma f - \Gamma \hat{f}| + \Gamma \hat{f}), \end{aligned}$$

from which the corollary easily follows. □

We see that when $\delta(f, \hat{f})$ goes to zero, the upper bounds in (10.7), (10.14) and (10.19) go to zero whereas the bounds in (10.18) go to one. This is very satisfactory. Furthermore, the upper bounds in (10.7), (10.14), (10.18), and (10.19) are increasing in $\delta(f, \hat{f})$ whereas the lower bound in (10.18) is decreasing. Hence, we get weaker bounds by replacing $\delta(f, \hat{f})$ with a higher value.

10.3.4 Convolutions and Compound Distributions

Our interest for approximations originated in De Pril's individual model . There we wanted to approximate a convolution of a finite number of compound distributions by replacing each of the counting distributions with an approximation. Therefore, we are interested in the properties of the distance measure δ in connection with convolutions as well as with compound functions where both the original distribution and the approximation have the same severity distribution.

We start with convolutions.

Theorem 10.5 *If* $f_j, \hat{f}_j \in \mathcal{F}_{1\underline{0}}$ *for* $j = 1, 2, \ldots, M$, *then*

$$\delta(*_{j=1}^M f_j, *_{j=1}^M \hat{f}_j) \le \sum_{j=1}^M \delta(f_j, \hat{f}_j).$$

Proof Application of Theorem 10.2 gives

$$\begin{aligned}\delta(*_{j=1}^M f_j, *_{j=1}^M \hat{f}_j) &= \left|\ln \frac{(*_{j=1}^M \hat{f}_j)(0)}{(*_{j=1}^M f_j)(0)}\right| + \sum_{x=1}^\infty \frac{|\varphi_{*_{j=1}^M f_j} - \varphi_{*_{j=1}^M \hat{f}_j}|(x)}{x} \\ &= \left|\ln \frac{\prod_{j=1}^M \hat{f}_j(0)}{\prod_{j=1}^M f_j(0)}\right| + \sum_{x=1}^\infty \frac{|\sum_{j=1}^M (\varphi_{f_j} - \varphi_{\hat{f}_j})(x)|}{x} \\ &\le \left|\sum_{j=1}^M \ln \frac{\hat{f}_j(0)}{f_j(0)}\right| + \sum_{x=1}^\infty \frac{\sum_{j=1}^M |\varphi_{f_j} - \varphi_{\hat{f}_j}|(x)}{x} \\ &\le \sum_{j=1}^M \left(\left|\ln \frac{\hat{f}_j(0)}{f_j(0)}\right| + \sum_{x=1}^\infty \frac{|\varphi_{f_j} - \varphi_{\hat{f}_j}|(x)}{x}\right) = \sum_{j=1}^M \delta(f_j, \hat{f}_j). \qquad \square\end{aligned}$$

Let us now turn to compound functions.

Theorem 10.6 *If* $p, \hat{p} \in \mathcal{F}_{1\underline{0}}$ *and* $h \in \mathcal{P}_{11}$, *then*

$$\delta(p \vee h, \hat{p} \vee h) \le \delta(p, \hat{p}).$$

Proof By application of Theorem 10.3, we obtain

$$\begin{aligned}
\delta(p \vee h, \hat{p} \vee h) &= \left|\ln \frac{(\hat{p} \vee h)(0)}{(p \vee h)(0)}\right| + \sum_{x=1}^{\infty} \frac{|\varphi_{p\vee h} - \varphi_{\hat{p}\vee h}|(x)}{x} \\
&= \left|\ln \frac{\hat{p}(0)}{p(0)}\right| + \sum_{x=1}^{\infty} \frac{|\Phi(\Psi\varphi_p \vee h) - \Phi(\Psi\varphi_{\hat{p}} \vee h)|(x)}{x} \\
&= \left|\ln \frac{\hat{p}(0)}{p(0)}\right| + \sum_{x=1}^{\infty} \left|\sum_{n=1}^{x} \frac{(\varphi_p - \varphi_{\hat{p}})(n) h^{n*}(x)}{n}\right| \\
&\leq \left|\ln \frac{\hat{p}(0)}{p(0)}\right| + \sum_{x=1}^{\infty} \sum_{n=1}^{\infty} \frac{|\varphi_p - \varphi_{\hat{p}}|(n) h^{n*}(x)}{n} \\
&= \left|\ln \frac{\hat{p}(0)}{p(0)}\right| + \sum_{n=1}^{\infty} \frac{|\varphi_p - \varphi_{\hat{p}}|(n) \sum_{x=1}^{\infty} h^{n*}(x)}{n} \\
&= \left|\ln \frac{\hat{p}(0)}{p(0)}\right| + \sum_{n=1}^{\infty} \frac{|\varphi_p - \varphi_{\hat{p}}|(n)}{n} = \delta(p, \hat{p}).
\end{aligned}$$

□

Now let N be a random variable with distribution p, and let p_m denote the distribution of mN for some positive integer m. As discussed in connection with Corollary 5.4, then $p_m = p \vee h_m$ with h_m concentrated in m. It then seems natural to approximate p_m with $\hat{p}_m = \hat{p} \vee h_m$. Application of Corollary 6.5 on (10.6) gives that $\delta(p_m, \hat{p}_m) = \delta(p, \hat{p})$. As

$$p_m(n) = \begin{cases} p(n/m) & (n = 0, m, 2m, \dots) \\ 0 & (\text{otherwise}) \end{cases}$$

$$\hat{p}_m(n) = \begin{cases} \hat{p}(n/m) & (n = 0, m, 2m, \dots) \\ 0, & (\text{otherwise}) \end{cases}$$

the definition (10.2) of ε immediately gives that $\varepsilon(p_m, \hat{p}_m) = \varepsilon(p, \hat{p})$. The invariance of the distance measures ε and δ against scaling seems very logical and satisfactory. Note that if $\hat{p} \in \mathcal{F}_{1\underline{0}}^{(r)}$, then $\hat{p}_m \in \mathcal{F}_{1\underline{0}}^{(mr)}$.

For ε, we have the following parallel to Theorem 10.6.

Theorem 10.7 *If* $p, \hat{p} \in \mathcal{F}_{1\underline{0}}$ *and* $h \in \mathcal{P}_{10}$, *then*

$$\varepsilon(p \vee h, \hat{p} \vee h) \leq \varepsilon(p, \hat{p}).$$

Proof We have

$$\varepsilon(p \vee h, \hat{p} \vee h) = \sum_{x=0}^{\infty} |p \vee h - \hat{p} \vee h|(x)$$

$$= \sum_{x=0}^{\infty}\left|\sum_{n=0}^{\infty}(p-\hat{p})(n)h^{n*}(x)\right| \le \sum_{x=0}^{\infty}\sum_{n=0}^{\infty}|p-\hat{p}|(n)h^{n*}(x)$$

$$= \sum_{n=0}^{\infty}|p-\hat{p}|(n)\sum_{x=0}^{\infty}h^{n*}(x) = \sum_{n=0}^{\infty}|p-\hat{p}|(n) = \varepsilon(p,\hat{p}). \qquad \square$$

By successive application of Theorems 10.7 and 10.4, we obtain

$$\varepsilon(p \vee h, \hat{p} \vee h) \le \varepsilon(p,\hat{p}) \le \mathrm{e}^{\delta(p,\hat{p})} - 1. \qquad (p, \hat{p} \in \mathcal{F}_{1\underline{0}}, h \in \mathcal{P}_{11})$$

We see that compounding the exact counting distribution and its approximation with the same severity distribution reduces the error measure.

With respect to convolution, we have the same situation. For $f, \hat{f} \in \mathcal{F}_{1\underline{0}}$ and $g \in \mathcal{P}_{11}$, we obtain

$$\varepsilon(f*g, \hat{f}*g) = \sum_{x=0}^{\infty}|f*g-\hat{f}*g|(x) = \sum_{x=0}^{\infty}\left|\sum_{y=0}^{x}(f-\hat{f})(y)g(x-y)\right|$$

$$\le \sum_{x=0}^{\infty}\sum_{y=0}^{x}|f-\hat{f}|(y)g(x-y) = \sum_{y=0}^{\infty}|f-\hat{f}|(y)\sum_{x=y}^{\infty}g(x-y)$$

$$= \sum_{y=0}^{\infty}|f-\hat{f}|(y) = \varepsilon(f,\hat{f}),$$

and application of Theorem 10.4 gives

$$\varepsilon(f*g, \hat{f}*g) \le \varepsilon(f,\hat{f}) \le \mathrm{e}^{\delta(f,\hat{f})} - 1.$$

From Theorem 10.5, we immediately obtain that $\delta(f*g, \hat{f}*g) \le \delta(f,\hat{f})$.

10.4 The Generalised De Pril Individual Model

Let us now return to De Pril's individual model as presented in Sect. 7.3. However, we assume that the p_js are in $\mathcal{P}_{1\underline{0}}$ without restricting them to the Bernoulli class. We shall refer to this model as the *generalised De Pril individual model*. We want to approximate f with

$$f^{(r)} = *_{i=1}^{I} *_{j=1}^{J} (p_j^{(r)} \vee h_i)^{M_{ij}*}, \tag{10.20}$$

where each $p_j^{(r)} \in \mathcal{F}_{1\underline{0}}^{(r)}$ for some fixed positive integer r called the *order of the approximation*. Then

$$\varphi_f(x) = x\sum_{n=1}^{x}\frac{1}{n}\sum_{j=1}^{J}\varphi_{p_j}(n)\sum_{i=1}^{I}M_{ij}h_i^{n*}(x) \tag{10.21}$$

$$\varphi_{f^{(r)}}(x) = x \sum_{n=1}^{r} \frac{1}{n} \sum_{j=1}^{J} \varphi_{p_j^{(r)}}(n) \sum_{i=1}^{I} M_{ij} h_i^{n*}(x) \tag{10.22}$$

for $x = 1, 2, \ldots$ and

$$f(0) = \prod_{j=1}^{J} p_j(0)^{M_{\bullet j}}$$

$$f^{(r)}(0) = \prod_{j=1}^{J} p_j^{(r)}(0)^{M_{\bullet j}} \tag{10.23}$$

with $M_{\bullet j} = \sum_{i=1}^{I} M_{ij}$ for $j = 1, 2, \ldots, J$. By application of Theorems 10.5 and 10.6, we obtain

$$\delta(f, f^{(r)}) \leq \sum_{i=1}^{I} \sum_{j=1}^{J} M_{ij} \delta(p_j \vee h_i, p_j^{(r)} \vee h_i) \leq \sum_{i=1}^{I} \sum_{j=1}^{J} M_{ij} \delta(p_j, p_j^{(r)}),$$

that is,

$$\delta(f, f^{(r)}) \leq \sum_{j=1}^{J} M_{\bullet j} \delta(p_j, p_j^{(r)}). \tag{10.24}$$

Combining this inequality with Theorem 10.4 gives

$$\varepsilon(f, f^{(r)}) \leq \exp\left(\sum_{j=1}^{J} M_{\bullet j} \delta(p_j, p_j^{(r)}) \right) - 1. \tag{10.25}$$

This error bound depends on only the counting distributions and their approximations, not the severity distributions, and it depends on them only through the $\delta(p_j, p_j^{(r)})$s. Hence, when applying this inequality for this kind of approximations in the following sections, we should focus on $\delta(p, p^{(r)})$ for some representative counting distribution $p \in \mathcal{P}_{1\underline{0}}$ and its rth order approximation $p^{(r)} \in \mathcal{F}_{1\underline{0}}^{(r)}$.

When each h_i is concentrated in a positive integer s_i, (10.21) and (10.22) give that for $x = 1, 2, \ldots,$

$$\varphi_f(x) = \sum_{i=1}^{I} s_i I(x/s_i = [x/s_i] > 0) \sum_{j=1}^{J} M_{ij} \varphi_{p_j}(x/s_i)$$

$$\varphi_{f^{(r)}}(x) = \sum_{i=1}^{I} s_i I(0 < x/s_i = [x/s_i] \leq r) \sum_{j=1}^{J} M_{ij} \varphi_{p_j^{(r)}}(x/s_i),$$

and by insertion in (6.2) we obtain that for $x = 1, 2, \ldots,$

$$f(x) = \frac{1}{x}\sum_{i=1}^{I} s_i \sum_{n=1}^{[x/s_i]} f(x - ns_i)\sum_{j=1}^{J} M_{ij}\varphi_{p_j}(n)$$

$$f^{(r)}(x) = \frac{1}{x}\sum_{i=1}^{I} s_i \sum_{n=1}^{[r/s_i]} f^{(r)}(x - ns_i)\sum_{j=1}^{J} M_{ij}\varphi_{p_j^{(r)}}(n). \tag{10.26}$$

We have assumed that the aggregate claims distribution of each policy is approximated with an approximation of order r. In the semi-collective approximation of Sect. 7.4, we suggested using a collective compound Poisson approximation for the normal policies and the exact distribution for some special policies. Analogously, in our present setting, we can use the rth order approximation for the normal policies and the exact distribution for the special policies. More generally, we can use different order of the approximation for different policies; the more normal policy, the lower order of the approximation. Exact evaluation corresponds to infinite order. One way to achieve this is to approximate f_{ij} with $f_{ij}^{(r_{ij})}$, where the order r_{ij} is chosen such that $\varepsilon(f_{ij}, f_{ij}^{(r_{ij})})$ or $\delta(f, \hat{f})$ does not exceed some fixed positive number.

In Sects. 10.5–10.7, we discuss the approximations of De Pril, Kornya, and Hipp respectively. In each case, we start with the general assumption that $p \in \mathcal{P}_{1\underline{0}}$ and define an rth order approximation $p^{(r)}$ of p. For the De Pril and Kornya approximations, we then restrict to $p \in \mathcal{R}_1$ and after that further to p being a Bernoulli distribution; we can exclude the case where p is a Poisson distribution as then $p \in \mathcal{F}_{1\underline{0}}^{(1)}$. For the Hipp approximation, we go straight from the general assumption $p \in \mathcal{P}_{1\underline{0}}$ to Bernoulli p. When we have found $p^{(r)}(0)$, the De Pril transform $\varphi_{p^{(r)}}$, and the distance $\delta(p, p^{(r)})$, or an upper bound for it, in the Bernoulli case, we turn to De Pril's individual model and insert these quantities for each p_j in (10.22)–(10.24) and (10.26). In Table 10.1, we display the expressions for $f^{(r)}(0)$ and $\varphi_{f^{(r)}}$ for these three types of approximation, and Table 10.2 gives upper bounds for $\delta(f, f^{(r)})$ for the case when all the π_js are less than 1/2. For $0 < \pi < 1/2$, we have

$$2\frac{\pi(1-\pi)}{1-2\pi} = \frac{\pi}{1-2\pi} + \pi > \frac{\pi}{1-2\pi},$$

that is, as expected, the error bound for the De Pril approximation is lower than for the Kornya approximation. Furthermore, for any positive integer r, we have

$$2\frac{\pi(1-\pi)}{1-2\pi}\left(\frac{\pi}{1-\pi}\right)^r < 2\frac{\pi(1-\pi)}{1-2\pi}(2\pi)^r = \frac{1-\pi}{1-2\pi}(2\pi)^{r+1} \le \frac{(2\pi)^{r+1}}{1-2\pi},$$

so that the error bound for the Kornya approximation is lower than the error bound for the Hipp approximation. However, these are only upper bounds so one should not necessarily conclude that they give a correct ranking between the three classes of approximations, cf. the discussion after the proof of Theorem 10.4.

Table 10.1 Approximations of order r

Approximation	$f^{(r)}(0)$	$\varphi_{f^{(r)}}(x)$
De Pril	$\sum_{j=1}^{J}(1-\pi_j)^{M_{\bullet j}}$	$-x\sum_{n=1}^{r}\frac{1}{n}\sum_{j=1}^{J}(\frac{\pi_j}{\pi_j-1})^n\sum_{i=1}^{I}M_{ij}h_i^{n*}(x)$
Kornya	$\exp(\sum_{n=1}^{r}\frac{1}{n}\sum_{j=1}^{J}M_{\bullet j}(\frac{\pi_j}{\pi_j-1})^n)$	$-x\sum_{n=1}^{r}\frac{1}{n}\sum_{j=1}^{J}(\frac{\pi_j}{\pi_j-1})^n$ $\times\sum_{i=1}^{I}M_{ij}h_i^{n*}(x)$
Hipp	$\exp(-\sum_{l=1}^{r}\sum_{j=1}^{J}M_{\bullet j}\frac{\pi_j^l}{l})$	$x\sum_{l=1}^{r}\frac{1}{l}\sum_{n=1}^{l}(-1)^{n+1}\binom{l}{n}\sum_{j=1}^{J}\pi_j^l$ $\times\sum_{i=1}^{I}M_{ij}h_i^{n*}(x)$

Table 10.2 Upper bounds for $\delta(f, f^{(r)})$

Approximation	Bound
De Pril	$\frac{1}{r+1}\sum_{j=1}^{J}M_{\bullet j}\frac{\pi_j}{1-2\pi_j}(\frac{\pi_j}{1-\pi_j})^r$
Kornya	$\frac{2}{r+1}\sum_{j=1}^{J}M_{\bullet j}\frac{\pi_j(1-\pi_j)}{1-2\pi_j}(\frac{\pi_j}{1-\pi_j})^r$
Hipp	$\frac{1}{r+1}\sum_{j=1}^{J}M_{\bullet j}\frac{(2\pi_j)^{r+1}}{1-2\pi_j}$

We see that for each of the three classes of approximations, the bound in Table 10.2 goes to zero when r goes to infinity. Hence, this is also the case with $\delta(f, f^{(r)})$, and, by (10.7), $\varepsilon(f, f^{(r)})$. For any positive integer x and any non-negative integer r, we have

$$|f(x)-f^{(r)}(x)|\le\sum_{y=0}^{\infty}|f(y)-f^{(r)}(y)|=\varepsilon(f, f^{(r)})$$

$$|\Gamma f(x)-\Gamma f^{(r)}(x)|=\left|\sum_{y=0}^{x}(f(y)-f^{(r)}(y))\right|\le\sum_{y=0}^{\infty}|f(y)-f^{(r)}(y)|=\varepsilon(f, f^{(r)}),$$

from which we obtain that $f^{(r)}$ and $\Gamma f^{(r)}$ converge uniformly to f and Γf.

10.5 The De Pril Approximation

10.5.1 General Counting Distribution

Let us look more closely at the distance measure δ when approximating a distribution $p\in\mathcal{P}_{1\underline{0}}$ with a function $p^{(r)}\in\mathcal{F}_{1\underline{0}}^{(r)}$. From (10.6), we obtain

$$\delta(p, p^{(r)})=\left|\ln\frac{p^{(r)}(0)}{p(0)}\right|+\sum_{n=1}^{r}\frac{|\varphi_p(n)-\varphi_{p^{(r)}}(n)|}{n}+\Lambda\Psi|\varphi_p|(r). \qquad (10.27)$$

We have already pointed out that among proportional approximations, $\delta(p, p^{(r)})$ is minimised by the approximation that approximates $p(0)$ with $p(0)$. Let us therefore now restrict to approximations $p^{(r)}$ for which

$$p^{(r)}(0) = p(0). \tag{10.28}$$

Then (10.27) reduces to

$$\delta(p, p^{(r)}) = \sum_{n=1}^{r} \frac{|\varphi_p(n) - \varphi_{p^{(r)}}(n)|}{n} + \Lambda\Psi|\varphi_p|(r).$$

In this class, we see that $\delta(p, p^{(r)})$ is minimised when

$$\varphi_{p^{(r)}}(n) = \varphi_p(n). \qquad (n = 1, 2, \dots, r) \tag{10.29}$$

This is the rth order De Pril approximation and gives

$$\delta(p, p^{(r)}) = \Lambda\Psi|\varphi_p|(r) \le \frac{\Lambda|\varphi_p|(r)}{r+1}. \tag{10.30}$$

Hence,

$$\delta(p, p^{(r)}) = \delta(p, p^{(r-1)}) - \Psi|\varphi_p(r)|. \qquad (r = 2, 3, \dots) \tag{10.31}$$

Suppose that we require the approximation $p^{(r)}$ to be so good that $\varepsilon(p, p^{(r)})$ is not greater than some given positive number ε_0. By Theorem 10.4, this condition is satisfied if $\delta(p, p^{(r)}) \le \ln(\varepsilon_0 + 1)$. If $\mu_{|\varphi_p|}(-1) < \infty$, then we can evaluate $\delta(p, p^{(r)})$ recursively by (10.31) until we get to the first r for which this inequality is satisfied, and evaluate $p^{(r)}$ with this r. In this case, we have the initial value $\delta(p, p^{(1)}) = \mu_{|\varphi_p|}(-1) - |\varphi_p(1)|$ for the recursion (10.31).

Once more, we stress that Theorem 10.4 only gives an upper bound for $\varepsilon(p, p^{(r)})$, so, even if the rth order De Pril approximation minimises $\delta(p, p^{(r)})$ among all rth order approximations, that does not necessarily imply that it also minimises $\varepsilon(p, p^{(r)})$.

From (10.29), we see that if φ_p is non-negative (that is, p is infinitely divisible, cf. Theorem 4.5), then

$$\varphi_{p^{(1)}} \le \varphi_{p^{(2)}} \le \varphi_{p^{(3)}} \le \cdots \le \varphi_p$$

and application of (10.28) and (6.2) gives that

$$p^{(1)} \le p^{(2)} \le p^{(3)} \le \cdots \le p.$$

10.5.2 Counting Distribution in $\mathcal{R}_1$

Let us now assume that p is $R_1[a,b]$ with $a \neq 0$. Then, insertion of (6.17) in (10.30) gives

$$\delta(p, p^{(r)}) = (a+b) \sum_{n=r+1}^{\infty} \frac{|a|^{n-1}}{n} < \frac{a+b}{r+1} \sum_{n=r+1}^{\infty} |a|^{n-1}. \tag{10.32}$$

We see that $\delta(p, p^{(r)}) = \infty$ when $a \leq -1$. For $0 < |a| < 1$, (10.32) gives

$$\delta(p, p^{(r)}) = -\frac{a+b}{|a|}\left(\ln(1-|a|) + \sum_{n=1}^{r} \frac{|a|^n}{n}\right) < \frac{a+b}{1-|a|}\frac{|a|^r}{r+1}, \tag{10.33}$$

so that we have the recursion

$$\delta(p, p^{(r)}) = \begin{cases} \delta(p, p^{(r-1)}) - (a+b)\frac{|a|^{r-1}}{r} & (r = 2, 3, \dots) \\ -(a+b)(\frac{\ln(1-|a|)}{|a|} + 1). & (r = 1) \end{cases} \tag{10.34}$$

When p is the Bernoulli distribution Bern(π), then

$$a = -\frac{\pi}{1-\pi}; \qquad b = 2\frac{\pi}{1-\pi}. \tag{10.35}$$

In this case, the condition $|a| < 1$ means that $\pi < 1/2$. Under this condition, insertion of (10.35) in (10.33) and (10.34) gives

$$\delta(p, p^{(r)}) = \ln\frac{1-\pi}{1-2\pi} - \sum_{n=1}^{r} \frac{1}{n}\left(\frac{\pi}{1-\pi}\right)^n < \frac{1}{r+1}\frac{\pi}{1-2\pi}\left(\frac{\pi}{1-\pi}\right)^r \tag{10.36}$$

$$\delta(p, p^{(r)}) = \begin{cases} \delta(p, p^{(r-1)}) - \frac{1}{r}(\frac{\pi}{1-\pi})^r & (r = 2, 3, \dots) \\ \ln\frac{1-\pi}{1-2\pi} - \frac{\pi}{1-\pi}. & (r = 1) \end{cases}$$

10.5.3 De Pril's Individual Model

Let us now return to De Pril's individual model. By application of (10.22), we obtain that the De Pril transform of the rth order De Pril approximation is given by

$$\varphi_{f^{(r)}}(x) = -x \sum_{n=1}^{r} \frac{1}{n} \sum_{j=1}^{J} \left(\frac{\pi_j}{\pi_j - 1}\right)^n \sum_{i=1}^{I} M_{ij} h_i^{n*}(x), \quad (x = 1, 2, \dots) \tag{10.37}$$

and the initial value for the recursion (6.2) for $f^{(r)}$ is

$$f^{(r)}(0) = f(0) = \prod_{j=1}^{J} (1-\pi_j)^{M_{\bullet j}}. \tag{10.38}$$

If $\pi_j < 1/2$ for $j = 1, 2, \ldots, J$, then insertion of (10.36) in (10.24) gives

$$\delta(f, f^{(r)}) \le \tilde{\delta}(f, f^{(r)}) < \frac{1}{r+1} \sum_{j=1}^{J} M_{\bullet j} \frac{\pi_j}{1 - 2\pi_j} \left(\frac{\pi_j}{1 - \pi_j} \right)^r, \tag{10.39}$$

with

$$\tilde{\delta}(f, f^{(r)}) = \sum_{j=1}^{J} M_{\bullet j} \left(\ln \frac{1 - \pi_j}{1 - 2\pi_j} - \sum_{n=1}^{r} \frac{1}{n} \left(\frac{\pi_j}{1 - \pi_j} \right)^n \right),$$

for which we have the recursion

$$\tilde{\delta}(f, f^{(r)}) = \begin{cases} \tilde{\delta}(f, f^{(r-1)}) - \frac{1}{r} \sum_{j=1}^{J} M_{\bullet j} (\frac{\pi_j}{1-\pi_j})^r & (r = 2, 3, \ldots) \\ \sum_{j=1}^{J} M_{\bullet j} (\ln \frac{1-\pi_j}{1-2\pi_j} - \frac{\pi_j}{1-\pi_j}). & (r = 1) \end{cases}$$

For the individual life model, application of (10.26) gives the recursion

$$f^{(r)}(x) = -\frac{1}{x} \sum_{i=1}^{I} s_i \sum_{n=1}^{[r/s_i]} f^{(r)}(x - ns_i) \sum_{j=1}^{J} M_{ij} \left(\frac{\pi_j}{\pi_j - 1} \right)^n. \qquad (x = 1, 2, \ldots) \tag{10.40}$$

10.6 The Kornya Approximation

10.6.1 General Counting Distribution

For $p \in \mathcal{P}_{1\underline{0}}$, we let $p^{(r)} \in \mathcal{F}_{1\underline{0}}^{(r)}$ be defined by (10.29) and $\mu_{p^{(r)}}(0) = 1$. Thus, the Kornya approximation and the De Pril approximation of p are proportional, and that also goes for these approximations of f. As

$$\mu_{|\varphi_{p^{(r)}}|}(-1) = \Gamma|\Psi\varphi_p|(r) < \infty,$$

and Theorem 6.1 trivially extends to functions in $\mathcal{F}_{1\underline{0}}^{(r)}$, we obtain

$$p^{(r)}(0) = \exp(-\Gamma\Psi\varphi_p(r)). \tag{10.41}$$

Hence,

$$\left| \ln \frac{p^{(r)}(0)}{p(0)} \right| = |\ln p(0) + \Gamma\Psi\varphi_p(r)|.$$

By insertion of this and (10.29) in (10.27), we obtain

$$\delta(p, p^{(r)}) = |\ln p(0) + \Gamma\Psi\varphi_p(r)| + \Lambda|\Psi\varphi_p|(r), \tag{10.42}$$

and when $\mu_{|\varphi_p|}(-1) < \infty$, application of Theorem 6.1 gives that

$$\delta(p, p^{(r)}) = |\Lambda\Psi\varphi_p(r)| + \Lambda|\Psi\varphi_p|(r). \tag{10.43}$$

10.6.2 Counting Distribution in $\mathcal{R}_1$

Let us now assume that p is $R_1[a, b]$ with $a \neq 0$. Then insertion of (6.17) in (10.41) and (10.42) gives

$$p^{(r)}(0) = \exp\left(-(a+b)\sum_{n=1}^{r}\frac{a^{n-1}}{n}\right) \tag{10.44}$$

$$\delta(p, p^{(r)}) = (a+b)\left(\left|\frac{\ln p(0)}{a+b} + \sum_{n=1}^{r}\frac{a^{n-1}}{n}\right| + \sum_{n=r+1}^{\infty}\frac{|a|^{n-1}}{n}\right). \tag{10.45}$$

If $a < -1$, then $\delta(p, p^{(r)}) = \infty$.
When $0 < |a| < 1$, (10.45) gives

$$\begin{aligned}\delta(p, p^{(r)}) &= \frac{a+b}{|a|}\left(\left|\frac{a}{a+b}\ln p(0) + \sum_{n=1}^{r}\frac{a^n}{n}\right| - \ln(1-|a|) - \sum_{n=1}^{r}\frac{|a|^n}{n}\right)\\ &= (a+b)\left(\left|\sum_{n=r+1}^{\infty}\frac{a^{n-1}}{n}\right| + \sum_{n=r+1}^{\infty}\frac{|a|^{n-1}}{n}\right).\end{aligned} \tag{10.46}$$

For $n = 0, 1, 2, \dots$, we have

$$\frac{a^{2n}}{2n+r+1} + \frac{a^{2n+1}}{2n+r+2} = \frac{a^{2n}}{2n+r+1}\left(1 + \frac{2n+r+1}{2n+r+2}a\right) > 0.$$

Hence,

$$\begin{aligned}\left|\sum_{n=r+1}^{\infty}\frac{a^{n-1}}{n}\right| &= |a|^r\left|\sum_{n=0}^{\infty}\left(\frac{a^{2n}}{2n+r+1} + \frac{a^{2n+1}}{2n+r+2}\right)\right|\\ &= |a|^r\sum_{n=0}^{\infty}\left(\frac{a^{2n}}{2n+r+1} + \frac{a^{2n+1}}{2n+r+2}\right) = |a|^r\sum_{n=0}^{\infty}\frac{a^n}{n+r+1}.\end{aligned}$$

Insertion in (10.46) gives

$$\delta(p, p^{(r)}) = (a+b)|a|^r\sum_{n=0}^{\infty}\frac{a^n + |a|^n}{n+r+1} \le (a+b)\frac{|a|^r}{r+1}\sum_{n=0}^{\infty}(a^n + |a|^n)$$

$$= (a+b)\left(\frac{1}{1-a} + \frac{1}{1-|a|}\right)\frac{|a|^r}{r+1},$$

that is,

$$\delta(p, p^{(r)}) \leq \begin{cases} 2\frac{a+b}{1-a}\frac{a^r}{r+1} & (a > 0) \\ 2\frac{a+b}{1-a^2}\frac{|a|^r}{r+1}. & (a < 0) \end{cases} \tag{10.47}$$

When $0 < a < 1$, (10.46) and (10.47) give

$$\delta(p, p^{(r)}) = 2(a+b)\sum_{n=r+1}^{\infty}\frac{a^{n-1}}{n} \leq 2\frac{a+b}{1-a}\frac{|a|^r}{r+1},$$

that is, twice of what we got for the De Pril approximation.

Let us now assume that p is the Bernoulli distribution Bern(π). Then insertion of (10.35) in (10.44) gives

$$p^{(r)}(0) = \exp\left(\sum_{n=1}^{r}\frac{1}{n}\left(\frac{\pi}{\pi-1}\right)^n\right). \tag{10.48}$$

If $\pi < 1/2$, then insertion of (10.35) in (10.46) and (10.47) gives

$$\begin{aligned}\delta(p, p^{(r)}) &= \left|\ln(1-\pi) - \sum_{n=1}^{r}\frac{1}{n}\left(\frac{\pi}{\pi-1}\right)^n\right| + \ln\frac{1-\pi}{1-2\pi} - \sum_{n=1}^{r}\frac{1}{n}\left(\frac{\pi}{1-\pi}\right)^n \\ &\leq \frac{2}{r+1}\frac{\pi(1-\pi)}{1-2\pi}\left(\frac{\pi}{1-\pi}\right)^r. \end{aligned} \tag{10.49}$$

10.6.3 De Pril's Individual Model

Let us now return to De Pril's individual model. Under the rth order Kornya approximation, $\varphi_{f^{(r)}}$ is given by (10.37), and by insertion of (10.48) in (10.23) we obtain

$$f^{(r)}(0) = \exp\left(\sum_{n=1}^{r}\frac{1}{n}\sum_{j=1}^{J}M_{\bullet j}\left(\frac{\pi_j}{\pi_j-1}\right)^n\right). \tag{10.50}$$

If $\pi_j < 1/2$ for $j = 1, 2, \ldots, J$, then insertion of (10.49) in (10.24) gives

$$\begin{aligned}\delta(f, f^{(r)}) \leq \sum_{j=1}^{J}M_{\bullet j}\Bigg(&\left|\ln(1-\pi_j) - \sum_{n=1}^{r}\frac{1}{n}\left(\frac{\pi_j}{\pi_j-1}\right)^n\right| \\ &+ \ln\frac{1-\pi_j}{1-2\pi_j} - \sum_{n=1}^{r}\frac{1}{n}\left(\frac{\pi_j}{1-\pi_j}\right)^n\Bigg)\end{aligned}$$

$$\leq \frac{2}{r+1}\sum_{j=1}^{J} M_{\bullet j}\frac{\pi_j(1-\pi_j)}{1-2\pi_j}\left(\frac{\pi_j}{1-\pi_j}\right)^r. \tag{10.51}$$

Let us consider the special case $r = 1$. Then, by (10.1) we obtain that for $j = 1, 2, \ldots, J$,

$$\varphi_{p_j^{(1)}}(1) = \varphi_{p_j}(1) = \frac{p_j(1)}{p_j(0)} = \frac{\pi_j}{1-\pi_j},$$

and from (4.17) we see that $p_j^{(1)}$ is the Poisson distribution $\mathrm{Po}(\pi_j/(1-\pi_j))$. Theorem 4.4 gives that the approximation $f^{(1)}$ is a compound Poisson distribution with Poisson parameter

$$\lambda = \sum_{j=1}^{J} M_{\bullet j}\frac{\pi_j}{1-\pi_j}$$

and severity distribution

$$h = \frac{1}{\lambda}\sum_{j=1}^{J}\frac{\pi_j}{1-\pi_j}\sum_{i=1}^{I} M_{ij}h_i,$$

that is, a collective compound Poisson approximation similar to the one of Sect. 7.4. For our present approximation, (10.51) gives

$$\delta(f, f^{(1)}) \leq \sum_{j=1}^{J} M_{\bullet j}\left(\left|\ln(1-\pi_j) + \frac{\pi_j}{1-\pi_j}\right| + \ln\frac{1-\pi_j}{1-2\pi_j} - \frac{\pi_j}{1-\pi_j}\right)$$

$$\leq \sum_{j=1}^{J} M_{\bullet j}\frac{\pi_j^2}{1-2\pi_j}. \tag{10.52}$$

Analogous to what we did in Sect. 7.4, we could apply this collective approximation for normal policies and exact evaluation for some special policies. We see that it is the policies with the highest value of π_j that give the highest contribution to the weaker bound in (10.52), so that it seems reasonable to choose those policies for exact evaluation.

Let us now compare the first order Kornya approximation with the collective compound Poisson approximation g of Sect. 7.4 where each p_j is approximated with the Poisson distribution $\mathrm{Po}(\pi_j)$. As

$$\frac{\pi_j}{1-\pi_j} - \pi_j = \frac{\pi_j^2}{1-\pi_j} > 0,$$

$\mathrm{Po}(\pi_j/(1-\pi_j))$ is the convolution of $\mathrm{Po}(\pi_j)$ and $\mathrm{Po}(\pi_j^2/(1-\pi_j))$. Hence, Theorem 4.4 gives that $f^{(1)} = g * k$ with $k = *_{i=1}^{I} *_{j=1}^{J} (q_j \vee h_i)^{Mij*}$ where q_j is

$\mathrm{Po}(\pi_j^2/(1-\pi_j))$. As $k \in \mathcal{P}_{1\underline{0}}$, we must have $\Gamma f^{(1)} \leq \Gamma g$. Application of Theorem 1.8 gives that $\Pi_{f^{(1)}} \geq \Pi_g$, and by combining this with Theorem 7.2, we obtain that $\Pi_f \leq \Pi_g \leq \Pi_{f^{(1)}}$. Hence, for the stop loss transform, the collective compound Poisson approximation of Sect. 7.4 gives a closer approximation than the first order Kornya approximation.

For the individual life model, (10.40) still holds.

10.7 The Hipp Approximation

10.7.1 General Counting Distribution

For $p \in \mathcal{P}_{1\underline{0}}$, we let $p^{(r)} \in \mathcal{F}_{1\underline{0}}^{(r)}$ be defined by the constraints

$$\kappa_{p^{(r)}}(j) = \kappa_p(j), \qquad (j = 0, 1, 2, \ldots, r)$$

assuming that these cumulants exist and that $\mu_{|\varphi_p|}(-1) < \infty$. The advantage of expressing these conditions in terms of cumulants instead of ordinary moments is the simple relation (6.6) between cumulants of a distribution and moments of its De Pril transform. This relation gives that the last r constraints can be expressed as

$$\mu_{\varphi_{p^{(r)}}}(j-1) = \kappa_p(j), \qquad (j = 1, 2, \ldots, r)$$

that is,

$$\sum_{n=1}^{r} n^{j-1} \varphi_{p^{(r)}}(n) = \kappa_p(j). \qquad (j = 1, 2, \ldots, r) \tag{10.53}$$

Hence, determining $\varphi_{p^{(r)}}$ boils down to solving a system of r linear equations with r unknowns. By application of Theorem 6.1, the first constraint gives

$$p^{(r)}(0) = \exp\left(-\sum_{n=1}^{r} \frac{\varphi_{p^{(r)}}(n)}{n}\right). \tag{10.54}$$

The following theorem gives an algorithm for solving the equations (10.53).

Theorem 10.8 *Let*

$$\eta_j^{(i)} = \eta_j^{(i-1)} - (i-1)\eta_{j-1}^{(i-1)} \qquad (j = i, i+1, i+2, \ldots, r;\ i = 2, 3, \ldots, r)$$

$$\eta_j^{(1)} = \kappa_p(j). \qquad (j = 1, 2, \ldots, r)$$

Then $\varphi_{p^{(r)}}(r), \varphi_{p^{(r)}}(r-1), \ldots, \varphi_{p^{(r)}}(1)$ can be evaluated by the recursion

$$\varphi_{p^{(r)}}(i) = \frac{\eta_i^{(i)}}{(i-1)!} - \sum_{n=i+1}^{r} \binom{n-1}{i-1} \varphi_{p^{(r)}}(n). \quad (i = r, r-1, \ldots, 1) \tag{10.55}$$

Proof We shall first prove by induction in i that

$$\eta_j^{(i)} = \sum_{n=i}^{r} n^{j-i}(n-1)^{(i-1)}\varphi_{p^{(r)}}(n). \tag{10.56}$$

$$(j = i, i+1, i+2, \dots, r;\ i = 1, 2, \dots, r)$$

From (10.53), (10.56) holds for $i = 1$. Let us now assume that it holds for $i = k-1$ for some $k \in \{2, 3, \dots, r\}$. Then, for $j = k, k+1, k+2, \dots, r$, we have

$$\begin{aligned}\eta_j^{(k)} &= \eta_j^{(k-1)} - (k-1)\eta_{j-1}^{(k-1)} \\ &= \sum_{n=k-1}^{r} n^{j-k+1}(n-1)^{(k-2)}\varphi_{p^{(r)}}(n) \\ &\quad - (k-1)\sum_{n=k-1}^{r} n^{j-k}(n-1)^{(k-2)}\varphi_{p^{(r)}}(n) \\ &= \sum_{n=k}^{r} n^{j-k}(n-1)^{(k-1)}\varphi_{p^{(r)}}(n),\end{aligned}$$

that is, the induction hypothesis holds also for $i = k$, and, hence, it holds for $i = 1, 2, \dots, r$.

For $j = i$, (10.56) gives

$$\eta_i^{(i)} = \sum_{n=i}^{r}(n-1)^{(i-1)}\varphi_{p^{(r)}}(n), \qquad (i = 1, 2, \dots, r)$$

from which we obtain (10.55). □

10.7.2 Bernoulli Counting Distribution

We now assume that p is the Bernoulli distribution Bern(π). From (5.36), we obtain

$$\kappa_p(j) = \pi\left(1 - \sum_{u=1}^{j-1}\binom{j-1}{u-1}\kappa_p(u)\right). \qquad (j = 1, 2, \dots) \tag{10.57}$$

By induction, we easily see that $\kappa_p(j)$ can be expressed as a polynomial of order j in π.

The following theorem gives an exact expression for the cumulants of p.

Theorem 10.9 *If p is the Bernoulli distribution* Bern(π), *then, for any positive integer r, we have*

$$\kappa_p(j) = \sum_{n=1}^{r} n^{j-1}(-1)^{n+1} \sum_{l=n}^{r} \pi^l \binom{l-1}{n-1}. \qquad (j = 1, 2, \ldots, r) \tag{10.58}$$

Proof For $j = 1, 2, \ldots, r$, insertion of (6.20) in (6.6) gives

$$\kappa_p(j) = \mu_{\varphi_p}(j-1) = \sum_{n=1}^{\infty} n^{j-1}\varphi_p(n) = -\sum_{n=1}^{\infty} n^{j-1}\left(\frac{\pi}{\pi-1}\right)^n$$
$$= \sum_{n=1}^{\infty} n^{j-1}(-1)^{n+1}\pi^n(1-\pi)^{-n} = \sum_{n=1}^{\infty} n^{j-1}(-1)^{n+1}\pi^n \sum_{l=0}^{\infty}\binom{l+n-1}{n-1}\pi^l$$
$$= \sum_{n=1}^{\infty} n^{j-1}(-1)^{n+1} \sum_{l=n}^{\infty}\binom{l-1}{n-1}\pi^l = \sum_{l=1}^{\infty}\pi^l \sum_{n=1}^{l} n^{j-1}(-1)^{n+1}\binom{l-1}{n-1}.$$

As $\kappa_p(j)$ is a polynomial of order j in π, the inner summation must be equal to zero for all $l > r$, so that

$$\kappa_p(j) = \sum_{l=1}^{r}\pi^l \sum_{n=1}^{l} n^{j-1}(-1)^{n+1}\binom{l-1}{n-1}, \tag{10.59}$$

and by once more changing the order of summation, we obtain (10.58). □

By comparing (10.53) and (10.58), we obtain

$$\varphi_{p^{(r)}}(n) = (-1)^{n+1}\sum_{l=n}^{r}\pi^l\binom{l-1}{n-1}, \qquad (n = 1, 2, \ldots, r)$$

that is,

$$\varphi_{p^{(r)}}(n) = (-1)^{n+1} n \sum_{l=n}^{r}\frac{\pi^l}{l}\binom{l}{n}. \qquad (n = 1, 2, \ldots, r) \tag{10.60}$$

Insertion in (10.54) gives

$$p^{(r)}(0) = \exp\left(-\sum_{n=1}^{r}(-1)^{n+1}\sum_{l=n}^{r}\frac{\pi^l}{l}\binom{l}{n}\right) = \exp\left(\sum_{l=1}^{r}\frac{\pi^l}{l}\sum_{n=1}^{l}(-1)^n\binom{l}{n}\right),$$

from which we obtain

$$p^{(r)}(0) = \exp\left(-\sum_{l=1}^{r}\frac{\pi^l}{l}\right). \tag{10.61}$$

Letting r go to infinity in (10.60) and (10.61) gives

$$\varphi_p(n) = (-1)^{n+1} n \sum_{l=n}^{\infty} \frac{\pi^l}{l} \binom{l}{n} \qquad (n = 1, 2, \dots) \tag{10.62}$$

$$p(0) = \exp\left(-\sum_{l=1}^{\infty} \frac{\pi^l}{l}\right). \tag{10.63}$$

Let us now turn to $\delta(p, p^{(r)})$. By insertion of (10.60)–(10.63) in (10.27), we obtain

$$\begin{aligned}
\delta(p, p^{(r)}) &= \left| \ln \frac{\exp(-\sum_{l=1}^{r} \pi^l/l)}{\exp(-\sum_{l=1}^{\infty} \pi^l/l)} \right| + \sum_{n=1}^{r} \left| (-1)^{n+1} \sum_{l=r+1}^{\infty} \frac{\pi^l}{l} \binom{l}{n} \right| \\
&\quad + \sum_{n=r+1}^{\infty} \left| (-1)^{n+1} \sum_{l=n}^{\infty} \frac{\pi^l}{l} \binom{l}{n} \right| \\
&= \sum_{l=r+1}^{\infty} \frac{\pi^l}{l} + \sum_{n=1}^{\infty} \sum_{l=\max(r+1,n)}^{\infty} \frac{\pi^l}{l} \binom{l}{n} \\
&\le \sum_{l=r+1}^{\infty} \frac{\pi^l}{l} + \sum_{n=1}^{\infty} \sum_{l=r+1}^{\infty} \frac{\pi^l}{l} \binom{l}{n} = \sum_{l=r+1}^{\infty} \frac{\pi^l}{l} \sum_{n=0}^{l} \binom{l}{n} = \sum_{l=r+1}^{\infty} \frac{(2\pi)^l}{l} \\
&\le \frac{1}{r+1} \sum_{l=r+1}^{\infty} (2\pi)^l.
\end{aligned}$$

If $\pi < 1/2$, then this gives

$$\delta(p, p^{(r)}) \le \tilde{\delta}(p, p^{(r)}) \le \frac{1}{r+1} \frac{(2\pi)^{r+1}}{1 - 2\pi} \tag{10.64}$$

with

$$\tilde{\delta}(p, p^{(r)}) = -\ln(1 - 2\pi) - \sum_{l=1}^{r} \frac{(2\pi)^l}{l},$$

which we can evaluate recursively by

$$\tilde{\delta}(p, p^{(r)}) = \begin{cases} \tilde{\delta}(p, p^{(r-1)}) - \frac{(2\pi)^r}{r} & (r = 2, 3, \dots) \\ -\ln(1 - 2\pi) - 2\pi. & (r = 1) \end{cases}$$

10.7.3 De Pril's Individual Model

Let us now return to De Pril's individual model. By insertion of (10.60) in (10.22) and interchanging summations, we obtain

$$\varphi_{f^{(r)}}(x) = x\sum_{l=1}^{r}\frac{1}{l}\sum_{n=1}^{l}(-1)^{n+1}\binom{l}{n}\sum_{j=1}^{J}\pi_j^l\sum_{i=1}^{I}M_{ij}h_i^{n*}(x).$$
$$(x = 1, 2, \dots)$$

Insertion of (10.61) in (10.23) gives

$$f^{(r)}(0) = \exp\left(-\sum_{l=1}^{r}\sum_{j=1}^{J}M_{\bullet j}\frac{\pi_j^l}{l}\right).$$

If $\pi_j < 1/2$ for $j = 1, 2, \dots, J$, then insertion of (10.64) in (10.24) gives

$$\delta(f, f^{(r)}) \le \tilde{\delta}(f, f^{(r)}) \le \frac{1}{r+1}\sum_{j=1}^{J}M_{\bullet j}\frac{(2\pi_j)^{r+1}}{1-2\pi_j} \tag{10.65}$$

with

$$\tilde{\delta}(f, f^{(r)}) = -\sum_{j=1}^{J}M_{\bullet j}\left(\ln(1-2\pi_j) + \sum_{l=1}^{r}\frac{(2\pi_j)^l}{l}\right),$$

which can be evaluated recursively by

$$\tilde{\delta}(f, f^{(r)}) = \begin{cases} \tilde{\delta}(f, f^{(r-1)}) - \sum_{j=1}^{J}M_{\bullet j}\frac{(2\pi_j)^r}{r} & (r = 2, 3, \dots) \\ -\sum_{j=1}^{J}M_{\bullet j}(\ln(1-2\pi_j) + 2\pi_j). & (r = 1) \end{cases}$$

As for each j the moments of p_j are matched by those of its approximation, that is also the case with f.

Let us consider the special case $r = 1$. Then (10.53) gives that for $j = 1, 2, \dots, J$,

$$\varphi_{p_j^{(1)}}(1) = \kappa_{p_j}(1) = \mu_{p_j}(1) = \pi_j,$$

and from (4.17) we see that $p_j^{(1)}$ is the Poisson distribution $\mathrm{Po}(\pi_j)$, that is, $f^{(1)}$ is the collective compound Poisson approximation of Sect. 7.4. For this approximation, (10.65) gives

$$\delta(f, f^{(1)}) \le -\sum_{j=1}^{J}M_{\bullet j}(\ln(1-2\pi_j) + 2\pi_j) \le 2\sum_{j=1}^{J}M_{\bullet j}\frac{\pi_j^2}{1-2\pi_j}. \tag{10.66}$$

By comparison with (10.52), we see that the weaker bound is twice the corresponding bound for the Kornya approximation. On the other hand, from the discussion at the end of Sect. 10.6.3 follows that for the stop loss transform, the first order Hipp approximation gives a closer approximation than the first order Kornya approximation.

In the semi-collective compound Poisson approximation of Sect. 7.4, we suggested using the present collective compound Poisson approximation for the normal policies and exact evaluation for some special policies. Under the efficiency criterion introduced there, we found that it was for the policies with the highest mean that we should use exact evaluation. We see that it is the policies with the highest π_j that contribute most to the weaker bound in (10.66), so that in our present setting, it seems reasonable to choose those policies for exact evaluation.

For the individual life model, insertion of (10.60) in (10.26) gives

$$f^{(r)}(x) = \frac{1}{x}\sum_{i=1}^{I} s_i \sum_{n=1}^{[r/s_i]} (-1)^{n+1} n f^{(r)}(x - n s_i) \sum_{j=1}^{J} M_{ij} \sum_{l=n}^{r} \frac{\pi_j^l}{l}\binom{l}{n}.$$
$$(x = 0, 1, 2, \dots)$$

10.8 Numerical Example

Let us now compare the approximations of De Pril, Kornya, and Hipp on the dataset of Table 7.1. With these data, we consider the exact aggregate claims distribution as well as the approximations of De Pril, Kornya, and Hipp of order 1, 2, 3, and 4. We see that $f(x) = 0$ for all $x \geq k = \sum_{i=1}^{I} i \sum_{j=1}^{J} M_{ij} = 97$.

In Table 10.3, we display the upper bounds for $\varepsilon(f, f^{(r)})$ obtained by insertion of the upper bounds for $\delta(f, f^{(r)})$ from Table 10.2 in (10.7). The bounds in Table 10.2 are rather simple and are often used in the literature. However, in (10.39), (10.51), and (10.65), we also gave sharper upper bounds for $\delta(f, f^{(r)})$. In Table 10.4, we have inserted these sharper bounds in (10.7). We see that the improvement by using these bounds is rather small.

To check how good the bounds in Table 10.4 are, it would have been interesting to compare them with the exact values of $\varepsilon(f, f^{(r)})$. Unfortunately, this is difficult in practice. However, as we know that the value of $f(x)$ is equal to zero for all $x \geq k$,

Table 10.3 Upper bounds for $\varepsilon(f, f^{(r)})$ based on upper bounds for $\delta(f, f^{(r)})$ from Table 10.2

r	1	2	3	4
Method		Bound		
De Pril	0.040014867	0.001394498	0.000057886	0.000002641
Kornya	0.077354499	0.002644503	0.000109536	0.000004991
Hipp	0.160692717	0.010061612	0.000784806	0.000067409

Table 10.4 Sharper upper bounds for $\varepsilon(f, f^{(r)})$

r	1	2	3	4
Method		Bound		
De Pril	0.039269708	0.001374488	0.000057199	0.000002614
Kornya	0.077236372	0.002641161	0.000109415	0.000004986
Hipp	0.154574226	0.009779147	0.000766601	0.000066069

Table 10.5 $\tilde{\varepsilon}(f, f^{(r)})$

r	1	2	3	4
Method		$\tilde{\varepsilon}(f, f^{(r)})$		
De Pril	0.036532060	0.001263337	0.000052213	0.000002372
Kornya	0.043662436	0.001903977	0.000087262	0.000004299
Hipp	0.026290081	0.001718855	0.000135298	0.000010774

we are mainly interested in the performance of the approximations on the range $\{0, 1, 2, \ldots, k\}$. Hence, in Table 10.5, we have displayed $\tilde{\varepsilon}(f, f^{(r)}) = \sum_{x=0}^{k} |f(x) - f^{(r)}(x)|$. We have

$$\tilde{\varepsilon}(f, f^{(r)}) = \varepsilon(f, f^{(r)}) - \sum_{x=k+1}^{\infty} |f(x) - f^{(r)}(x)| < \varepsilon(f, f^{(r)}).$$

Table 10.5 gives quite a different impression of the Hipp approximation than Tables 10.3 and 10.4. It is better than the Kornya approximation for $r = 1, 2$, and for $r = 1$, it is even better than the De Pril approximation. Considering the discussion after the proof of Theorem 10.4, it is not that surprising that the upper bound in (10.7) seems to give a too flattering impression of the De Pril approximation compared to other approximations.

In Table 10.6, we display exact values of f as well as the approximations.

In Table 10.7, we display the bounds for $\Gamma f/\Gamma f^{(r)}$ from formula (10.18) with $\delta(f, f^{(r)})$ replaced with the sharp bounds of (10.39), (10.51), and (10.65).

In Table 10.8, we display values of Λf together with the approximations of De Pril and Kornya found by subtracting the approximation of the cumulative distribution from one. These figures illustrates a feature that has often appeared in numerical studies of these approximations. Let $f^{(r)\mathrm{D}}$ and $f^{(r)\mathrm{K}}$ denote the rth order approximation of De Pril and Kornya respectively. In numerical examples, in particular in our present example, it often seems that

$$\Gamma f^{(2)\mathrm{D}} \leq \Gamma f^{(4)\mathrm{D}} \leq \Gamma f^{(6)\mathrm{D}} \leq \cdots \leq \Gamma f \leq \cdots \leq \Gamma f^{(5)\mathrm{D}} \leq \Gamma f^{(3)\mathrm{D}} \leq \Gamma f^{(1)\mathrm{D}} \tag{10.67}$$

Table 10.6 Exact evaluation and approximations of f

r		1			2			3			4		
x	Exact	De Pril	Kornya	Hipp	De Pril	Kornya	Hipp	De Pril	Kornya	Hipp	De Pril	Kornya	Hipp
0	2.3819E-01	2.3819E-01	2.2980E-01	2.4660E-01	2.3819E-01	2.3850E-01	2.3847E-01	2.3819E-01	2.3818E-01	2.3821E-01	2.3819E-01	2.3820E-01	2.3820E-01
1	1.4734E-02	1.4734E-02	1.4214E-02	1.4796E-02	1.4734E-02	1.4752E-02	1.4738E-02	1.4734E-02	1.4733E-02	1.4734E-02	1.4734E-02	1.4734E-02	1.4734E-02
2	8.7734E-02	8.7962E-02	8.4862E-02	8.6753E-02	8.7734E-02	8.7845E-02	8.7641E-02	8.7734E-02	8.7730E-02	8.7728E-02	8.7734E-02	8.7734E-02	8.7734E-02
3	1.1318E-01	1.1319E-01	1.0920E-01	1.1122E-01	1.1318E-01	1.1332E-01	1.1302E-01	1.1318E-01	1.1318E-01	1.1317E-01	1.1318E-01	1.1318E-01	1.1318E-01
4	1.1071E-01	1.1297E-01	1.0899E-01	1.1040E-01	1.1071E-01	1.1085E-01	1.1074E-01	1.1071E-01	1.1070E-01	1.1071E-01	1.1071E-01	1.1071E-01	1.1071E-01
5	9.6327E-02	9.6569E-02	9.3165E-02	9.2859E-02	9.6326E-02	9.6447E-02	9.6105E-02	9.6327E-02	9.6322E-02	9.6315E-02	9.6327E-02	9.6328E-02	9.6327E-02
6	6.1549E-02	6.5197E-02	6.2900E-02	6.1008E-02	6.1469E-02	6.1547E-02	6.1587E-02	6.1549E-02	6.1546E-02	6.1551E-02	6.1549E-02	6.1549E-02	6.1549E-02
7	6.9022E-02	7.0317E-02	6.7839E-02	6.5427E-02	6.9015E-02	6.9103E-02	6.8853E-02	6.9022E-02	6.9019E-02	6.9015E-02	6.9022E-02	6.9022E-02	6.9022E-02
8	5.4817E-02	5.9184E-02	5.7099E-02	5.4577E-02	5.4790E-02	5.4859E-02	5.4951E-02	5.4820E-02	5.4818E-02	5.4834E-02	5.4817E-02	5.4817E-02	5.4819E-02
9	4.3147E-02	4.5436E-02	4.3835E-02	4.1321E-02	4.3008E-02	4.3062E-02	4.3012E-02	4.3147E-02	4.3145E-02	4.3132E-02	4.3147E-02	4.3147E-02	4.3146E-02
10	3.0107E-02	3.4211E-02	3.3005E-02	3.0579E-02	3.0065E-02	3.0103E-02	3.0269E-02	3.0108E-02	3.0107E-02	3.0123E-02	3.0107E-02	3.0107E-02	3.0108E-02
11	2.3529E-02	2.6444E-02	2.5512E-02	2.3308E-02	2.3461E-02	2.3491E-02	2.3581E-02	2.3531E-02	2.3529E-02	2.3534E-02	2.3529E-02	2.3529E-02	2.3529E-02
12	1.8282E-02	2.1005E-02	2.0265E-02	1.8344E-02	1.8139E-02	1.8161E-02	1.8272E-02	1.8288E-02	1.8287E-02	1.8280E-02	1.8282E-02	1.8282E-02	1.8282E-02
13	1.2509E-02	1.5250E-02	1.4713E-02	1.3149E-02	1.2436E-02	1.2451E-02	1.2601E-02	1.2511E-02	1.2510E-02	1.2515E-02	1.2509E-02	1.2509E-02	1.2510E-02
14	8.7108E-03	1.0850E-02	1.0467E-02	9.2180E-03	8.6241E-03	8.6350E-03	8.7533E-03	8.7131E-03	8.7126E-03	8.7119E-03	8.7107E-03	8.7107E-03	8.7107E-03
15	5.9116E-03	7.7623E-03	7.4887E-03	6.5043E-03	5.7860E-03	5.7933E-03	5.9146E-03	5.9143E-03	5.9140E-03	5.9067E-03	5.9114E-03	5.9114E-03	5.9108E-03
16	4.1519E-03	5.5492E-03	5.3536E-03	4.5955E-03	4.0780E-03	4.0832E-03	4.1700E-03	4.1579E-03	4.1577E-03	4.1543E-03	4.1519E-03	4.1519E-03	4.1524E-03
17	2.7150E-03	3.8808E-03	3.7440E-03	3.1764E-03	2.6355E-03	2.6388E-03	2.7209E-03	2.7174E-03	2.7173E-03	2.7127E-03	2.7149E-03	2.7149E-03	2.7147E-03
18	1.7409E-03	2.6296E-03	2.5370E-03	2.1234E-03	1.6748E-03	1.6770E-03	1.7408E-03	1.7436E-03	1.7435E-03	1.7390E-03	1.7408E-03	1.7408E-03	1.7407E-03
19	1.1174E-03	1.7747E-03	1.7122E-03	1.4139E-03	1.0571E-03	1.0584E-03	1.1079E-03	1.1204E-03	1.1203E-03	1.1153E-03	1.1172E-03	1.1173E-03	1.1172E-03
20	7.1102E-04	1.1941E-03	1.1520E-03	9.3953E-04	6.6418E-04	6.6502E-04	7.0177E-04	7.1563E-04	7.1560E-04	7.1113E-04	7.1077E-04	7.1077E-04	7.1116E-04
30	3.0943E-06	1.2435E-05	1.1996E-05	8.6329E-06	1.4524E-06	1.4542E-06	2.1291E-06	3.4588E-06	3.4586E-06	3.2110E-06	3.0383E-06	3.0384E-06	3.0938E-06
40	3.5351E-09	5.9356E-08	5.7264E-08	3.6416E-08	−5.1519E-09	−5.1584E-09	−3.7538E-09	9.0839E-09	9.0835E-09	6.8691E-09	1.2847E-09	1.2847E-09	2.5890E-09

Table 10.7 Bounds based on (10.18)

r	Bound	Method		
		De Pril	Kornya	Hipp
1	Lower	0.962214	0.928301	0.866120
2		0.998627	0.997366	0.990316
3		0.999943	0.999891	0.999234
4		0.999997	0.999995	0.999934
4	Upper	1.000003	1.000005	1.000066
3		1.000057	1.000109	1.000767
2		1.001376	1.002648	1.009876
1		1.040875	1.083701	1.182836

$$\Gamma f^{(1)\mathrm{K}} \le \Gamma f^{(3)\mathrm{K}} \le \Gamma f^{(5)\mathrm{K}} \le \cdots \le \Gamma f \le \cdots \le \Gamma f^{(6)\mathrm{K}} \le \Gamma f^{(4)\mathrm{K}} \le \Gamma f^{(2)\mathrm{K}}. \tag{10.68}$$

These inequalities would imply

$$\Gamma f^{(r)\mathrm{D}} \le \Gamma f \le \Gamma f^{(r)\mathrm{K}} = \frac{f^{(r)\mathrm{K}}(0)}{f^{(r)\mathrm{D}}(0)}\Gamma f^{(r)\mathrm{D}} \qquad (r = 2, 4, 6, \dots)$$

$$\Gamma f^{(r)\mathrm{D}} \ge \Gamma f \ge \Gamma f^{(r)\mathrm{K}} = \frac{f^{(r)\mathrm{K}}(0)}{f^{(r)\mathrm{D}}(0)}\Gamma f^{(r)\mathrm{D}}, \qquad (r = 1, 3, 5, \dots)$$

which can be expressed as

$$\min\left(1, \frac{f^{(r)\mathrm{K}}(0)}{f^{(r)\mathrm{D}}(0)}\right) \le \frac{\Gamma f(x)}{\Gamma f^{(r)\mathrm{D}}(x)} \le \max\left(1, \frac{f^{(r)\mathrm{K}}(0)}{f^{(r)\mathrm{D}}(0)}\right) \tag{10.69}$$

when $\Gamma f^{(r)\mathrm{D}}(x) > 0$. Insertion of (10.50) and (10.38) gives

$$\frac{f^{(r)\mathrm{K}}(0)}{f^{(r)\mathrm{D}}(0)} = \frac{\exp(\sum_{n=1}^{r} \frac{1}{n} \sum_{j=1}^{J} M_{\bullet j} (\frac{\pi_j}{\pi_j - 1})^n)}{\prod_{j=1}^{J}(1-\pi_j)^{M_{\bullet j}}}$$

$$= \prod_{j=1}^{J}\left(\frac{\exp(\sum_{n=1}^{r} \frac{1}{n} (\frac{\pi_j}{\pi_j - 1})^n)}{1-\pi_j}\right)^{M_{\bullet j}}.$$

In Table 10.9, we have displayed the span of the bounds in (10.69), that is, the upper bound minus the lower bound,

$$\max\left(1, \frac{\Gamma f^{(r)\mathrm{K}}(0)}{\Gamma f^{(r)\mathrm{D}}(0)}\right) - \min\left(1, \frac{\Gamma f^{(r)\mathrm{K}}(0)}{\Gamma f^{(r)\mathrm{D}}(0)}\right) = \left|\frac{\Gamma f^{(r)\mathrm{K}}(0)}{\Gamma f^{(r)\mathrm{D}}(0)} - 1\right|,$$

together with the span of the bounds for the De Pril approximation from Table 10.7. We see that the span of the bounds from Table 10.7 is more than twice the span of

Table 10.8 Exact evaluation and approximations of Λf

r		1			2			3			4		
x	Exact	De Pril	Kornya	Hipp	De Pril	Kornya	Hipp	De Pril	Kornya	Hipp	De Pril	Kornya	Hipp
0	7.6181E-01	7.6181E-01	7.7020E-01	7.5340E-01	7.6181E-01	7.6150E-01	7.6153E-01	7.6181E-01	7.6182E-01	7.6179E-01	7.6181E-01	7.6180E-01	7.6180E-01
1	7.4707E-01	7.4707E-01	7.5599E-01	7.3861E-01	7.4707E-01	7.4675E-01	7.4679E-01	7.4707E-01	7.4708E-01	7.4706E-01	7.4707E-01	7.4707E-01	7.4707E-01
2	6.5934E-01	6.5911E-01	6.7112E-01	6.5185E-01	6.5934E-01	6.5891E-01	6.5915E-01	6.5934E-01	6.5936E-01	6.5933E-01	6.5934E-01	6.5934E-01	6.5934E-01
3	5.4615E-01	5.4592E-01	5.6192E-01	5.4063E-01	5.4616E-01	5.4558E-01	5.4613E-01	5.4615E-01	5.4618E-01	5.4616E-01	5.4615E-01	5.4615E-01	5.4615E-01
4	4.3544E-01	4.3295E-01	4.5293E-01	4.3023E-01	4.3545E-01	4.3474E-01	4.3539E-01	4.3544E-01	4.3547E-01	4.3544E-01	4.3544E-01	4.3544E-01	4.3544E-01
5	3.3912E-01	3.3638E-01	3.5977E-01	3.3737E-01	3.3912E-01	3.3829E-01	3.3928E-01	3.3912E-01	3.3915E-01	3.3913E-01	3.3912E-01	3.3912E-01	3.3912E-01
6	2.7757E-01	2.7118E-01	2.9687E-01	2.7637E-01	2.7765E-01	2.7674E-01	2.7770E-01	2.7757E-01	2.7761E-01	2.7758E-01	2.7757E-01	2.7757E-01	2.7757E-01
7	2.0855E-01	2.0086E-01	2.2903E-01	2.1094E-01	2.0864E-01	2.0764E-01	2.0884E-01	2.0855E-01	2.0859E-01	2.0856E-01	2.0855E-01	2.0854E-01	2.0855E-01
8	1.5373E-01	1.4168E-01	1.7193E-01	1.5636E-01	1.5385E-01	1.5278E-01	1.5389E-01	1.5373E-01	1.5377E-01	1.5373E-01	1.5373E-01	1.5373E-01	1.5373E-01
9	1.1058E-01	9.6241E-02	1.2809E-01	1.1504E-01	1.1084E-01	1.0972E-01	1.1088E-01	1.1058E-01	1.1063E-01	1.1060E-01	1.1058E-01	1.1058E-01	1.1058E-01
10	8.0475E-02	6.2030E-02	9.5088E-02	8.4463E-02	8.0777E-02	7.9614E-02	8.0611E-02	8.0470E-02	8.0518E-02	8.0475E-02	8.0475E-02	8.0473E-02	8.0475E-02
11	5.6946E-02	3.5586E-02	6.9576E-02	6.1155E-02	5.7316E-02	5.6123E-02	5.7030E-02	5.6940E-02	5.6989E-02	5.6941E-02	5.6946E-02	5.6944E-02	5.6946E-02
12	3.8664E-02	1.4581E-02	4.9311E-02	4.2811E-02	3.9177E-02	3.7962E-02	3.8758E-02	3.8652E-02	3.8702E-02	3.8662E-02	3.8664E-02	3.8662E-02	3.8663E-02
13	2.6154E-02	−6.6952E-04	3.4599E-02	2.9662E-02	2.6742E-02	2.5511E-02	2.6157E-02	2.6141E-02	2.6192E-02	2.6146E-02	2.6155E-02	2.6152E-02	2.6154E-02
14	1.7444E-02	−1.1519E-02	2.4131E-02	2.0444E-02	1.8118E-02	1.6875E-02	1.7404E-02	1.7428E-02	1.7479E-02	1.7434E-02	1.7444E-02	1.7442E-02	1.7443E-02
15	1.1532E-02	−1.9281E-02	1.6643E-02	1.3939E-02	1.2332E-02	1.1082E-02	1.1489E-02	1.1513E-02	1.1565E-02	1.1528E-02	1.1533E-02	1.1530E-02	1.1532E-02
16	7.3801E-03	−2.4831E-02	1.1289E-02	9.3437E-03	8.2535E-03	6.9990E-03	7.3195E-03	7.3555E-03	7.4073E-03	7.3734E-03	7.3808E-03	7.3784E-03	7.3797E-03
17	4.6651E-03	−2.8711E-02	7.5450E-03	6.1673E-03	5.6181E-03	4.3603E-03	4.5985E-03	4.6380E-03	4.6900E-03	4.6607E-03	4.6658E-03	4.6635E-03	4.6650E-03
18	2.9241E-03	−3.1341E-02	5.0081E-03	4.0439E-03	3.9432E-03	2.6833E-03	2.8578E-03	2.8945E-03	2.9465E-03	2.9217E-03	2.9250E-03	2.9226E-03	2.9243E-03
19	1.8068E-03	−3.3116E-02	3.2959E-03	2.6301E-03	2.8862E-03	1.6249E-03	1.7499E-03	1.7741E-03	1.8262E-03	1.8065E-03	1.8078E-03	1.8054E-03	1.8071E-03
20	1.0958E-03	−3.4310E-02	2.1439E-03	1.6905E-03	2.2220E-03	9.5989E-04	1.0481E-03	1.0585E-03	1.1106E-03	1.0953E-03	1.0970E-03	1.0946E-03	1.0959E-03
30	3.4984E-06	−3.6514E-02	1.7845E-05	1.2462E-05	1.2641E-03	7.2998E-07	1.6883E-06	−4.7881E-05	4.3310E-06	3.8605E-06	5.6963E-06	3.3241E-06	3.4715E-06
40	3.1083E-09	−3.6532E-02	7.3588E-08	4.5530E-08	1.2633E-03	−5.9969E-09	−5.2239E-09	−5.2202E-05	1.0932E-08	8.0202E-09	2.3713E-06	−8.0764E-10	1.1901E-09

Table 10.9 Span of the bounds of Table 10.7 and (10.69)

r	Table 10.7	(10.69)
1	0.07866072	0.03524451
2	0.00274898	0.00126493
3	0.00011440	0.00005221
4	0.00000523	0.00000237

the bounds of (10.69), so it is tempting to go for the latter bounds. However, we should keep in mind that we know that the bounds in Table 10.7 hold, whereas we do not know what regularity conditions are needed for the bounds in (10.69).

The reason that we have represented the inequalities in terms of Γ and not Λ, is that, as the De Pril approximation normally does not sum to one, the figures for the De Pril approximation in Table 10.8 are not values of $\Lambda f^{(r)\mathrm{D}}$.

Further Remarks and References

The definition of the De Pril transform was extended to functions in $\mathcal{F}_{1\underline{0}}$ by Dhaene and Sundt (1998), who studied the properties of functions in $\mathcal{F}_{1\underline{0}}$, in particular in relation to their De Pril transforms and in the context of approximating distributions in $\mathcal{P}_{1\underline{0}}$ with functions in $\mathcal{F}_{1\underline{0}}$. Sections 10.1–10.5 are to a large extent based on that paper. Within the same context, Sundt et al. (1998) studied moments and cumulants of functions in $\mathcal{F}_{1\underline{0}}$.

The Kornya approximation was introduced by Kornya (1983) for the individual life model and by Hipp (1986) for the general model. Both these papers gave error bounds. In particular, Hipp (1986) proved that the cumulative first order Kornya approximation gives a lower bound to the exact distribution in De Pril's individual model.

The De Pril approximation was introduced by Vandebroek and De Pril (1988) and De Pril (1988) for the individual life model and De Pril (1989) for the general model. In the published discussion to Kornya (1983), within the context of the individual life model, Elias Shiu suggested the De Pril approximation as a possible improvement to the Kornya approximation, arguing that it matches the first probabilities of the exact distribution.

The Hipp approximation was introduced by Hipp (1986); see also Hipp and Michel (1990). Dhaene et al. (1996) showed that the moments of that approximation match the moments of the exact distribution up to the order of the approximation. Sundt et al. (1998) deduced the Hipp approximation by using the moment matching as constraints. The algorithm of Theorem 10.8 was deduced by Christian Irgens in a note to Bjørn Sundt in 1997.

De Pril (1989) compared the approximations of De Pril, Kornya, and Hipp and deduced separately for each of the approximations the upper bounds for $\varepsilon(f, f^{(r)})$ obtained by insertion in (10.7) of the upper bounds for $\delta(f, f^{(r)})$ given in Table 10.2; for the De Pril approximation in the individual life model, the bound for

$\varepsilon(f, f^{(r)})$ had been given by De Pril (1988). A unified approach to such approximations and error bounds was presented by Dhaene and De Pril (1994). In particular, they proved Theorem 10.4 and gave the bounds for $\delta(f, f^{(r)})$ displayed in Table 10.2. Furthermore, they proved Corollaries 10.1 and 10.3. Their results were expressed in terms of the Dhaene–De Pril transform.

For the error measure ε, Dhaene and Sundt (1997) deduced inequalities that are not expressed in terms of De Pril transforms. In particular, they proved Theorem 10.7.

Dhaene and De Pril (1994) and Dhaene and Sundt (1997, 1998) also deduced error bounds for approximations of stop loss transforms based on the approximations for the associated aggregate claims distribution.

Kornya (1983) gave a proof of (10.68) within the context of the individual life model, but in the published discussion to that paper, David C. McIntosh and Donald P. Minassian pointed out an error in the proof. The latter discussant gave counterexamples. Hence, one needs some regularity conditions. Vernic et al. (2009) give sufficient conditions for the ordering (10.67) to hold for the distribution of the number of policies with claims and discuss both that ordering and the ordering (10.68) more generally.

Within the individual life model, Vandebroek and De Pril (1988) made a numerical comparison between the approximations of De Pril and Kornya and the compound Poisson approximation obtained as the first order Hipp approximation, as well as the natural approximation presented in Sect. 7.4.

Hipp (1986) studied the approximations of Hipp and Kornya for a more general class of severity distributions in the setting of signed Poisson measures. This approach has been further studied in the non-actuarial literature, see e.g. Roos (2005) and references therein.

Within the actuarial framework, Kornya (2007) gave upper bounds for $|\Gamma f - \Gamma f^{(r)}|$ under the approximations of Kornya and Hipp, illustrated with numerical examples. In particular for large portfolios, these bounds seem to perform well compared with the bounds presented in the present chapter.

Chapter 11
Extension to Distributions in $\mathcal{P}_{1\_}$

Summary

In Chap. 10, we discussed approximations in $\mathcal{F}_{1\underline{0}}$ of distributions in $\mathcal{P}_{1\underline{0}}$ and error measures for such approximations. In the present chapter, we shall extend this setting to approximations in $\mathcal{F}_{1\_}$ of distributions in $\mathcal{P}_{1\_}$. Our trick is shifting, like when we deduced Theorem 2.10 from Theorem 2.8.

Section 11.1 is devoted to De Pril transforms. In Sect. 11.1.1, the definition of the De Pril transform is extended to functions in $\mathcal{F}_{1\_}$. For a function in $\mathcal{F}_{1\underline{0}}$, we defined the De Pril transform as a function in $\mathcal{F}_{11}$; for a function in $\mathcal{F}_{1\_}$, we let the shifting parameter needed for shifting that function to $\mathcal{F}_{1\underline{0}}$ be the value of its De Pril transform at zero, and, at positive integers, we let it have the same value as the De Pril transform of that shifted function. In Sect. 11.1.2, we extend some results on the De Pril transform to our present setting.

Finally, in Sect. 11.2, we extend the error measure ε and some results on bounds for this measure to approximations in $\mathcal{F}_{1\_}$ of distributions in $\mathcal{P}_{1\_}$.

11.1 De Pril Transforms

11.1.1 Definition

In the present and next chapter, we shall extend some results to distributions of random variables that can take negative values. In the present chapter, we concentrate on distributions in $\mathcal{P}_{1\_}$ and approximations in $\mathcal{F}_{1\_}$.

To sort out our basic idea, let us consider M independent random variables $X_1, X_2, \dots, X_M$ with distributions in $\mathcal{P}_{1\_}$. For $j = 1, 2, \dots, M$, let f_j denote the distribution of X_j. As $f_j \in \mathcal{P}_{1\_}$, there exists an integer l_j such that $f_j \in \mathcal{P}_{1\underline{l_j}}$. We want to evaluate the distribution f of $X = X_{\bullet M}$.

In Theorem 2.10, we have already encountered such a situation in the special case with identically distributed X_js. In that case, let l denote the common value of the l_js. What we did in the proof of Theorem 2.10, was to replace each X_j with $\tilde{X}_j = X_j - l$. The distribution of $\tilde{X}_j$ was then easily found by shifting the distribution of X_j, and this shifted distribution was in $\mathcal{P}_{1\underline{0}}$. Then we could evaluate the distribution $\tilde{X} = \tilde{X}_{\bullet M}$ as the M-fold convolution of the distribution of $\tilde{X}_1$ by the recursion of Theorem 2.8. As $X = \tilde{X} + Ml$, its distribution was easily found by shifting the distribution of $\tilde{X}$.

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_11, © Springer-Verlag Berlin Heidelberg 2009

It seems obvious how to extend this procedure to the general case: For each j, let $\tilde{X}_j = X_j - l_j$, and let

$$\tilde{X} = \sum_{j=1}^{M} (X_j - l_j) = X - l$$

with $l = l_{\bullet M}$. The distribution of $\tilde{X}$ is now the convolution of M distributions in $\mathcal{P}_{1\underline{0}}$, and we know how to evaluate such a convolution by De Pril's first method. Finally, the distribution of X is found by shifting the distribution of $\tilde{X}$.

Let us try to express this procedure in terms of De Pril transforms. For each j, we let $\tilde{f}_j$ denote the distribution of $\tilde{X}_j$, and we let $\tilde{f}$ denote the distribution of $\tilde{X}$, that is, $\tilde{f} = *_{j=1}^{M} \tilde{f}_j$. Then

$$\varphi_{\tilde{f}} = \sum_{j=1}^{M} \varphi_{\tilde{f}_j}.$$

For each j, $\varphi_{\tilde{f}_j}$ is uniquely determined by $\tilde{f}_j$, and $\tilde{f}$ is uniquely determined by $\varphi_{\tilde{f}}$. Thus, what we still need, is the shifting parameter l_j of each f_j and l of f. We have already seen that the shifting parameters are additive for convolutions, that is, that the shifting parameter of the convolution is equal to the sum of the shifting parameters of the distributions that we convolute. From Corollary 6.1, we know that the De Pril transform also has this property. Furthermore, we know that a distribution in $\mathcal{P}_{1\_}$ is uniquely determined by the De Pril transform of the shifted distribution in $\mathcal{P}_{1\underline{0}}$ and the shifting parameter. It is therefore natural to define the De Pril transform of a distribution in $\mathcal{P}_{1\_}$ by these two quantities. Till now, we have defined the De Pril transform of a distribution only on the positive integers. To also include the shifting parameter, it seems natural to extend the range of the De Pril transform to zero and let its value at zero be equal to the shifting parameter of the shifted distribution in $\mathcal{P}_{1\underline{0}}$. For positive integers, we let the De Pril transform of a distribution in $\mathcal{P}_{1\_}$ be equal to the De Pril transform of the shifted distribution in $\mathcal{P}_{1\underline{0}}$.

If $f \in \mathcal{P}_{1\_}$, then

$$\varphi_f(0) = \min\{x : f(x) > 0\}. \tag{11.1}$$

Let $\tilde{f}$ denote f shifted to $\mathcal{P}_{1\underline{0}}$. Then

$$\varphi_f(x) = \varphi_{\tilde{f}}(x). \qquad (x = 1, 2, \dots) \tag{11.2}$$

For $x = \varphi_f(0) + 1, \varphi_f(0) + 2, \dots$, application of (6.2) gives

$$f(x) = \tilde{f}(x - \varphi_f(0)) = \frac{1}{x - \varphi_f(0)} \sum_{y=1}^{x-\varphi_f(0)} \varphi_{\tilde{f}}(y) \tilde{f}(x - \varphi_f(0) - y),$$

from which we obtain that (6.2) extends to

$$f(x) = \frac{1}{x - \varphi_f(0)} \sum_{y=1}^{x-\varphi_f(0)} \varphi_f(y) f(x-y) \qquad (x = \varphi_f(0)+1, \varphi_f(0)+2, \dots) \tag{11.3}$$

for $f \in \mathcal{P}_{1\_}$. Solving for $\varphi_f(x - \varphi_f(0))$ gives

$$\varphi_f(x - \varphi_f(0)) = \frac{1}{f(\varphi_f(0))} \left((x - \varphi_f(0)) f(x) - \sum_{y=1}^{x-\varphi_f(0)-1} \varphi_f(y) f(x-y) \right),$$

and by change of variable we obtain

$$\varphi_f(x) = \frac{1}{f(\varphi_f(0))} \left(x f(x + \varphi_f(0)) - \sum_{y=1}^{x-1} \varphi_f(y) f(x + \varphi_f(0) - y) \right), \quad (x = 1, 2, \dots) \tag{11.4}$$

which extends (6.1) to distributions in $\mathcal{P}_{1\_}$.

Also for functions $f \in \mathcal{F}_{1\_}$, we define the De Pril transform φ_f by (11.1) and (11.4).

11.1.2 Extension of Results

From the way we constructed the De Pril transform of a function in $\mathcal{F}_{1\_}$, it is obvious that Theorem 10.2 holds also for functions in $\mathcal{F}_{1\_}$.

The following theorem extends Theorem 10.3 to compound functions with counting function in $\mathcal{F}_{10}$.

Theorem 11.1 *If $p \in \mathcal{F}_{10}$ and $h \in \mathcal{P}_{11}$, then*

$$\varphi_{p \vee h}(x) = \varphi_p(0)\varphi_h(x) + x \sum_{n=1}^{x} \frac{\varphi_p(n)}{n} h^{n*}(x). \qquad (x = 0, 1, 2, \dots) \tag{11.5}$$

Proof Let $\tilde{p}$ denote the function p shifted to $\mathcal{F}_{1\underline{0}}$. Then $p \vee h = h^{\varphi_p(0)*} * (\tilde{p} \vee h)$, and application of Theorems 10.2 and 10.3 gives

$$\varphi_{p \vee h} = \varphi_p(0)\varphi_h + \varphi_{\tilde{p} \vee h} = \varphi_p(0)\varphi_h + \Phi(\Psi\varphi_{\tilde{p}} \vee h).$$

By application of (11.2), we obtain (11.5). □

It seems tempting to write (11.5) in the more compact form $\varphi_{p \vee h} = \varphi_p(0)\varphi_h + \Phi(\Psi\varphi_p \vee h)$. However, as $\varphi_p(0)$ is now not necessarily equal to zero, φ_p is not necessarily in $\mathcal{F}_{11}$, and we have defined the operator Ψ only for functions in $\mathcal{F}_{11}$.

The following theorem extends Theorem 6.1 to functions in $\mathcal{F}_{1\_}$.

Theorem 11.2 *If* $p \in \mathcal{F}_{1\_}$ *and* $\sum_{n=1}^{\infty} |\varphi_p(n)|/n < \infty$, *then*

$$\kappa_p(j) = I(j=0)\ln p(0) + I(j=1)\varphi_p(0) + \sum_{x=1}^{\infty} x^{j-1}\varphi_p(x). \qquad (j=0,1,2,\dots) \tag{11.6}$$

Proof Let $\tilde{p}$ denote the function p shifted to $\mathcal{F}_{1\underline{0}}$. From (1.21), we obtain that

$$\kappa_p(j) = I(j=1)\varphi_p(0) + \kappa_{\tilde{p}}(j), \qquad (j=0,1,2,\dots)$$

and application of (6.6), Theorem 6.1, and (11.2), gives (11.6). □

11.2 Error Bounds

Let us now try to extend some of the results of Sect. 10.3. The obvious extension of the error measure ε to $\mathcal{F}_{1\_}$ is

$$\varepsilon(f,\hat{f}) = \mu_{|f-\hat{f}|}(0) = \sum_{x=-\infty}^{\infty} |f(x) - \hat{f}(x)|. \qquad (f,\hat{f} \in \mathcal{F}_{1\_})$$

However, it seems non-trivial to extend $\delta(f,\hat{f})$, and Theorem 10.7, to all choices of f and $\hat{f}$ in $\mathcal{F}_{1\_}$, so we restrict to the case with $\hat{f} \in \mathcal{F}_{1\underline{\varphi_f}}$, and for that case we let

$$\delta(f,\hat{f}) = \left|\ln \frac{\hat{f}(\varphi_f(0))}{f(\varphi_f(0))}\right| + \sum_{x=1}^{\infty} \frac{|\varphi_f(x) - \varphi_{\hat{f}}(x)|}{x}.$$

With this definition, we obtain by simple shifting that Theorem 10.4 more generally holds for $f \in \mathcal{P}_{1\_}$ and $\hat{f} \in \mathcal{F}_{1\underline{\varphi_f}}$. Analogously, Theorem 10.5 holds more generally when $f_1, f_2, \dots, f_M \in \mathcal{P}_{1\_}$ and $\hat{f}_j \in \mathcal{F}_{1\underline{\varphi_{f_j}}}$ for $j = 1, 2, \dots, M$.

Let us now extend Theorem 10.6.

Theorem 11.3 *If* $p \in \mathcal{P}_{10}$, $\hat{p} \in \mathcal{F}_{1\underline{\varphi_p}}$, *and* $h \in \mathcal{P}_{11}$, *then*

$$\delta(p \vee h, \hat{p} \vee h) \le \delta(p,\hat{p}).$$

Proof Let $\tilde{p}$ and $\tilde{\hat{p}}$ denote p and $\hat{p}$ shifted to $\mathcal{P}_{1\underline{0}}$ and $\mathcal{F}_{1\underline{0}}$ respectively. Then

$$\begin{aligned}\delta(p \vee h, \hat{p} \vee h) &= \delta(h^{\varphi_p(0)*} * (\tilde{p} \vee h), h^{\varphi_p(0)*} * (\tilde{\hat{p}} \vee h))\\ &\le \delta(h^{\varphi_p(0)*}, h^{\varphi_p(0)*}) + \delta(\tilde{p} \vee h, \tilde{\hat{p}} \vee h) = \delta(\tilde{p} \vee h, \tilde{\hat{p}} \vee h)\\ &\le \delta(\tilde{p}, \tilde{\hat{p}}) = \delta(p,\hat{p}).\end{aligned}$$

□

Further Remarks and References

Sundt (1995) pointed out that by shifting one could extend De Pril's first method for evaluating the convolution of a finite number of distributions in $\mathcal{P}_{1\underline{0}}$ to distributions in $\mathcal{P}_{1\_}$.

Sundt (1998) extended the definition of the De Pril transform to functions in $\mathcal{F}_{1\_}$ and proved the results that we have presented in the present chapter.

Chapter 12
Allowing for Negative Severities

Summary

In Chap. 2, we started with developing recursions for compound distributions with severity distribution in $\mathcal{P}_{11}$ and counting distribution in $\mathcal{P}_{10}$, the latter satisfying a linear recursion of order one. In Sect. 2.7, we extended the setting to severity distributions in $\mathcal{P}_{10}$, and in Chap. 5, we allowed for recursions of higher order for the counting distribution. Section 9.2 was devoted to recursions for moments of compound distributions, and there we allowed for negative severities; we did not even require them to be integer-valued. But what about extending the recursions for compound distributions to severity distributions in $\mathcal{P}_1$, that is, allowing for negative integer-valued severities? That is the topic of the present chapter.

Section 12.1 gives a general introduction, pointing out the problems that arise when extending the recursions by allowing for negative severities. As the extension is non-trivial even when the counting distribution is in the Panjer class, we restrict to that case. In Sects. 12.2–12.4, we treat the binomial, Poisson, and negative binomial case respectively.

12.1 Introduction

Let us consider the compound distribution $f = p \vee h$ with $p \in \mathcal{P}_{10}$ and $h \in \mathcal{P}_1$. If p satisfies the recursion (5.6), then we would have to extend the relation (5.8) to severity distribution $h \in \mathcal{P}_1$. We actually did that in (9.16).

For simplicity, we restrict to the case where p is $R_1[a, b]$, that is, p is in the Panjer class $\mathcal{R}_1$. Then (9.16) reduces to

$$xf(x) = \sum_{y=-\infty}^{\infty} (ax + by)h(y)f(x - y). \tag{12.1}$$

When $h \in \mathcal{P}_{10}$, this gives

$$xf(x) = \sum_{y=0}^{x} (ax + by)h(y)f(x - y).$$

By solving this equation for $f(x)$, we obtain (2.79).

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_12, © Springer-Verlag Berlin Heidelberg 2009

That it is not always trivial to extend this procedure to allow for negative severities, is clearly seen even in the apparently simple case where h is non-degenerate with support $\{-1, 1\}$. Then (12.1) reduces to

$$xf(x) = (ax - b)h(-1)f(x + 1) + (ax + b)h(1)f(x - 1),$$

which gives

$$f(x + 1) = \frac{xf(x) - (ax + b)h(1)f(x - 1)}{(ax - b)h(-1)} \tag{12.2}$$

$$f(x - 1) = \frac{xf(x) - (ax - b)h(-1)f(x + 1)}{(ax + b)h(1)} \tag{12.3}$$

if the denominators are not equal to zero. In principle, these are recursions for f, but where should we start the recursion? In the negative binomial and Poisson cases, the support of f would be unbounded both upwards and downwards. However, if p is the binomial distribution $\text{bin}(M, \pi)$, then $f(x) = 0$ for all integers x such that $|x| > M$. Furthermore, we have $f(-M) = (\pi h(-1))^M$ and $f(M) = (\pi h(1))^M$, so that we can evaluate f recursively by (12.2) or (12.3). More generally, in Sect. 12.2, we shall extend this procedure to the case when the support of h is bounded on at least one side.

From this example, we see that even when p is in the Panjer class, a general extension to severity distributions allowing for negative severities would be non-trivial. To avoid making our presentation too complex, we shall therefore restrict to this case in the following.

12.2 Binomial Counting Distribution

Let p be the binomial distribution $\text{bin}(M, \pi)$.

If $h \in \mathcal{P}_{1\underline{l}}$ for some negative integer l, then $f(Ml) = (\pi h(l))^M$ and $f(x) = 0$ for all integers $x < Ml$. For $x = Ml + 1, Ml + 2, \ldots$, (12.1) gives

$$xf(x) = \sum_{y=l}^{x-Ml} (ax + by)h(y)f(x - y).$$

The highest argument of f in this equation is $x - l$, so we solve for $f(x - l)$ and obtain

$$f(x - l) = \frac{1}{(ax + bl)h(l)}\left(xf(x) - \sum_{y=l+1}^{x-Ml} (ax + by)h(y)f(x - y)\right).$$

Replacing x with $x + l$ and y with $y + l$ gives

$$f(x) = \frac{1}{(ax + (a+b)l)h(l)} \times \left((x+l) f(x+l) - \sum_{y=1}^{x-Ml} (ax + by + (a+b)l) h(y+l) f(x-y) \right).$$

Finally, by insertion of the expressions for a and b from Table 2.1, we obtain

$$f(x) = \frac{1}{(x - Ml)h(l)} \left(\sum_{y=1}^{x-Ml} (Ml + (M+1)y - x) h(y+l) f(x-y) - \left(\frac{1}{\pi} - 1\right)(x+l) f(x+l) \right). \qquad (x = Ml+1, Ml+2, \ldots) \qquad (12.4)$$

On the other hand, let us now assume that the range of h is bounded upwards so that there exists an integer k such that $h(k) > 0$ and $h(y) = 0$ for all integers $y > k$. Then $f(Mk) = (\pi h(k))^M$ and $f(x) = 0$ for all integers $x > Mk$. For $x = Mk - 1, Mk - 2, \ldots$, (12.1) gives

$$x f(x) = \sum_{y=x-Mk}^{k} (ax + by) h(y) f(x-y).$$

The lowest argument of f in this equation is $x - k$ so we solve for $f(x-k)$ and obtain

$$f(x-k) = \frac{1}{(ax + bk)h(k)} \left(x f(x) - \sum_{y=x-Mk}^{k-1} (ax + by) h(y) f(x-y) \right).$$

Replacing x with $x + k$ and y with $k - y$ gives

$$f(x) = \frac{1}{(ax + (a+b)k)h(k)} \times \left((x+k) f(x+k) - \sum_{y=1}^{Mk-x} (ax - by + (a+b)k) h(k-y) f(x+y) \right).$$

Finally, by insertion of the expressions from Table 2.1 for a and b, we obtain

$$f(x) = \frac{1}{(Mk - x)h(k)} \left(\left(\frac{1}{\pi} - 1\right)(x+k) f(x+k) - \sum_{y=1}^{Mk-x} (Mk - x - (M+1)y) h(k-y) f(x+y) \right).$$

$$(x = Mk - 1, Mk - 2, \ldots)$$

If the support of the severity distribution h is unbounded both from above and below, then the procedures presented above do not work. It seems then to be advisable to approximate the severity distribution with a function whose support is bounded on one side. As the left tail would normally be much lighter than the right tail, we consider truncation from the left.

Let $Y_1, Y_2, \ldots,$ be independent and identically distributed random variables with distribution h independent of the random variable N, which has distribution p. Then $X = Y_{\bullet N}$ has distribution $f = p \vee h$. For some negative integer l, we introduce $Y_{il} = \max(l, Y_i)$ for $i = 1, 2, \ldots$. The distribution $h_l \in \mathcal{P}_{1\underline{l}}$ of the Y_{il}s is given by

$$h_l(y) = \begin{cases} h(y) & (y = l+1, l+2, \ldots) \\ \Gamma h(l), & (y = l) \end{cases}$$

and $X_l = \sum_{i=1}^{N} Y_{il}$ has distribution $f_l = p \vee h_l$, which can be evaluated recursively by (12.4). As $X_l \geq X$, we have $\Gamma f \geq \Gamma f_l$. Thus, from (1.46), we obtain $\Pi_{f_l} - \varepsilon_l \leq \Pi_f \leq \Pi_{f_l}$ with

$$\begin{aligned} \varepsilon_l &= \mu_{f_l}(1) - \mu_f(1) = \mathrm{E}(X_l - X) = \mu_p(1)\,\mathrm{E}(Y_{1l} - Y_1) \\ &= \mu_p(1)(\Pi_h(l) + l - \mu_h(1)), \end{aligned}$$

and application of Corollary 1.5 gives

$$\Lambda f_l - \varepsilon_l \leq \Lambda f \leq \Lambda f_l; \qquad \Gamma f_l \leq \Gamma f \leq \Gamma f_l + \varepsilon_l; \qquad |f - f_l| \leq \varepsilon_l.$$

12.3 Poisson Counting Distribution

Let p be the Poisson distribution Po(λ). Then we obtain from Table 2.1 that $a = 0$ and $b = \lambda$ so that (12.1) reduces to

$$xf(x) = \lambda \sum_{y=-\infty}^{\infty} yh(y)f(x-y).$$

Somewhat simpler, but still not too encouraging. However, let us consider the setting in terms of random variables. Let the severities $Y_1, Y_2, \ldots$ be mutually independent and identically distributed with distribution h and independent of the counting variable N, which has distribution p. For $i = 1, 2, \ldots,$ we also introduce the variables $Y_i^+ = (Y_i)_+$ and $Y_i^- = (-Y_i)_+$. Then $Y_i = Y_i^+ - Y_i^-$. We denote the distributions of the Y_i^+s and the Y_i^-s by h^+ and h^- respectively. Then $h^+, h^- \in \mathcal{P}_{10}$, and these distributions are given by

$$h^+(y) = \begin{cases} h(y) & (y = 1, 2, \ldots) \\ \Gamma h(0) & (y = 0) \end{cases} \tag{12.5}$$

$$h^-(y) = \begin{cases} h(-y) & (y = 1, 2, \dots) \\ \Lambda h(-1). & (y = 0) \end{cases} \tag{12.6}$$

Let $X = Y_{\bullet N}$, $X^+ = Y^+_{\bullet N}$, $X^- = Y^-_{\bullet N}$. Then X, X^+, and X^- have distributions $f = p \vee h$, $f^+ = p \vee h^+$, and $f^- = p \vee h^-$ respectively, and we have $X = X^+ - X^-$. The compound Poisson distributions f^+ and f^- can be evaluated recursively by (2.7); from Table 2.3, (12.5), and (12.6), we obtain the initial values $f^+(0) = e^{-\lambda \Lambda h(0)}$ and $f^-(0) = e^{-\lambda \Gamma h(-1)}$.

Lemma 12.1 *Under the above assumptions, the variables X^+ and X^- are independent.*

Proof For convenience, in this proof, we denote the characteristic function of the distribution of a random variable by ζ with the symbol of that random variable as subscript.

We have

$$\zeta_Y(s) = \sum_{y=-\infty}^{\infty} e^{isy} h(y) = h(0) + \sum_{y=1}^{\infty} e^{isy} h^+(y) + \sum_{y=1}^{\infty} e^{-isy} h^-(y)$$
$$= h(0) + \zeta_{Y^+}(s) - \Gamma h(0) + \zeta_{Y^-}(-s) - \Lambda h(-1) = \zeta_{Y^+}(s) + \zeta_{Y^-}(-s) - 1.$$

By application of (1.29) and (2.12), we obtain

$$\zeta_X(s) = \tau_p(\zeta_Y(s)) = e^{\lambda(\zeta_Y(s)-1)} = e^{\lambda(\zeta_{Y^+}(s)+\zeta_{Y^-}(-s)-2)}$$
$$= e^{\lambda(\zeta_{Y^+}(s)-1)} e^{\lambda(\zeta_{Y^-}(-s)-1)} = \tau_p(\zeta_{Y^+}(s))\tau_p(\zeta_{Y^-}(-s))$$
$$= \zeta_{X^+}(s)\zeta_{X^-}(-s) = \zeta_{X^+}(s)\zeta_{-X^-}(s).$$

Hence, X^+ and $-X^-$ are independent, and this implies that X^+ and X^- are independent. □

From Lemma 12.1, we obtain

$$f(x) = \sum_{z=\max(0,-x)}^{\infty} f^+(x+z) f^-(z).$$

At least, we have now got an explicit expression for $f(x)$ in the shape of a sum of quantities that we are able to evaluate. However, unfortunately, the sum has an infinite number of terms. Hence, we need some approximation. We introduce the random variables $X^-_m = \min(m, X^-)$ and $X_m = X^+ - X^-_m$. The distribution f^-_m of X^-_m is given by

$$f^-_m(x) = \begin{cases} f^-(x) & (x = 0, 1, 2, \dots, m-1) \\ \Lambda f^-(m-1), & (x = m) \end{cases}$$

and for the distribution f_m of X_m we obtain

$$f_m(x) = \sum_{z=\max(0,-x)}^{m} f^+(x+z) f_m^-(z)$$
$$= \sum_{z=\max(0,-x)}^{m-1} f^+(x+z) f^-(z) + f^+(x+m)\Lambda f^-(m-1).$$

We obviously have $X_m^- \leq X^-$ and $X_m \geq X$. This implies that $\Gamma f \geq \Gamma f_m$. From (1.46), we obtain $\Pi_{f_m} - \varepsilon_m \leq \Pi_f \leq \Pi_{f_m}$ with

$$\varepsilon_m = \mu_{f_m}(1) - \mu_f(1) = \mathrm{E}(X_m - X) = \mathrm{E}(X^- - X_m^-) = \Pi_{f^-}(m),$$

and application of Corollary 1.5 gives

$$\Lambda f_m - \varepsilon_m \leq \Lambda f \leq \Lambda f_m; \qquad \Gamma f_m \leq \Gamma f \leq \Gamma f_m + \varepsilon_m; \qquad |f - f_m| \leq \varepsilon_m.$$

12.4 Negative Binomial Counting Distribution

Let p be the negative binomial distribution $\mathrm{NB}(\alpha, \pi)$. For convenience, we assume that h has finite range $\{l, l+1, l+2, \dots, r\}$ with $l < 0 < r$ and $f(l) \neq 0 \neq f(r)$. If this assumption is not fulfilled from the origin, then we could approximate the original severity distribution with a distribution with finite range, analogous to what we did in Sect. 12.2.

By insertion of the expressions for a and b from Table 2.1 in (12.1), we obtain that

$$xf(x) = \pi \sum_{y=l}^{r} (x + (\alpha-1)y) h(y) f(x-y)$$

for any integer x. Solving for $f(x-l)$ gives

$$f(x-l) = \frac{1}{(x+(\alpha-1)l)h(l)} \left(\frac{xf(x)}{\pi} - \sum_{y=l+1}^{r} (x+(\alpha-1)y)h(y)f(x-y) \right),$$

and by replacing x with $x+l$, we obtain

$$f(x) = \frac{1}{(x+\alpha l)h(l)} \left(\frac{(x+l)f(x+l)}{\pi} - \sum_{y=l+1}^{r} (x+l+(\alpha-1)y)h(y)f(x+l-y) \right). \tag{12.7}$$

We still have the problem where to start the recursion. In this connection, the following asymptotic result can be useful.

Theorem 12.1 *Let p be the negative binomial distribution* $\mathrm{NB}(\alpha, \pi)$ *and* $h \in \mathcal{P}_1$.
i) *If there exists a* $\sigma_+ > 1$ *such that* $\tau_h(\sigma_+) = 1/\pi$, *then*

$$(p \vee h)(x) \sim \frac{1}{\Gamma(\alpha)} \left(\frac{1-\pi}{\pi\sigma_+ \tau_h'(\sigma_+)} \right)^\alpha x^{\alpha-1} \sigma_+^{-x}. \qquad (x \uparrow \infty)$$

ii) *If there exists a* $\sigma_- \in (0, 1)$ *such that* $\tau_h(\sigma_-) = 1/\pi$, *then*

$$(p \vee h)(x) \sim \frac{1}{\Gamma(\alpha)} \left(\frac{1-\pi}{\pi\sigma_- \tau_h'(\sigma_-)} \right)^\alpha |x|^{\alpha-1} \sigma_+^{-x}. \qquad (x \downarrow -\infty) \tag{12.8}$$

We suggest to start the recursion (12.7) at x equal to some small negative integer $L < l$, approximating $f(x)$ by the asymptotic expression (12.8) for $x = L + l - 1$, $L + l - 2, \ldots, L + l - r$. To get an idea whether we have chosen L sufficiently small, we can check whether the first couple of values of f from the recursion are reasonably close to the corresponding asymptotic values obtain from (12.8).

Further Remarks and References

The question of how to evaluate a compound distribution with counting distribution in the Panjer class and severity distribution that allows for negative severities, was briefly addressed by Sundt and Jewell (1981). They presented the recursion (12.4) for a compound binomial distribution, pointed out that in the Poisson case, one could utilise that X^+ and X^- are independent, and suggested that one could apply an iterative approach in the negative binomial case. This research was followed up in Milidiu's (1985) doctoral thesis, supervised by Jewell; some results from that thesis were presented in Jewell and Milidiu (1986). Theorem 12.1 was proved by Milidiu (1985). Of other literature on asymptotic expressions for compound distributions, we mention Sundt (1982), Embrechts et al. (1985), Panjer and Willmot (1986, Chap. 10), and Willmot (1989, 1990).

The error bounds presented in Sect. 12.3 were deduced by Sundt (1986b). The results were generalised by Hürlimann (1991).

Chapter 13
Underflow and Overflow

Summary

One issue that we have ignored till now, is numerical problems. In the present chapter, we shall concentrate on underflow and overflow. In our recursions for distributions, underflow could typically occur at the initiation of the recursion. For the compound Poisson recursion of Theorem 2.2, the initial value at zero is $e^{-\lambda}$. The parameter λ is the mean of the counting variable; in insurance applications, that could typically be the expected number of claims in an insurance portfolio. If λ is very large, then $e^{-\lambda}$ could be so small that the computer would round it to zero, and then all values of the compound distribution evaluated by the recursion (2.7) will be set to zero. With a common spreadsheet program, that seems to happen when the expected number of claims exceeds 708.

For simplicity, we restrict to recursions for distributions in $\mathcal{P}_{1\underline{0}}$.

In a short Sect. 13.1, we discuss avoiding underflow and overflow by simple scaling, that is, for the recursive evaluation of the distribution, we multiply the distribution by a constant, and afterwards we divide it by that constant. Exponential scaling is a more advanced scaling procedure to be discussed in Sect. 13.2. Finally, in Sect. 13.3, we consider handling underflow and overflow by expressing the distribution as a convolution.

13.1 Simple Scaling

We assume that we want to evaluate a function $f \in \mathcal{P}_{1\underline{0}}$ recursively. The most obvious way to proceed if $f(0)$ is so small that we get underflow, is to evaluate $f_c = cf$ for some large number c, that is, we let the initial value be $cf(0)$ instead of $f(0)$. When we have done the recursive evaluation of cf, then we scale back by dividing the evaluated values by c.

If c is large, then we risk overflow when evaluating $cf(x)$ for x sufficiently large. Then we can scale back to f or df for some smaller constant d.

13.2 Exponential Scaling

Let us now consider a more advanced scaling method. We assume that $p \in \mathcal{P}_{1\underline{0}}$ satisfies the recursion (5.6).

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_13,

For some number u and some positive integer v, let

$$\ddot{p}(n) = p(n)e^{u(v-n)} \qquad (n = 0, 1, 2, \dots)$$
$$\ddot{q}(n) = q(n)e^{u(v-n)} \qquad (n = 1, 2, \dots)$$
$$\ddot{a}(i) = a(i)e^{-ui}; \qquad \ddot{b}(i) = b(i)e^{-ui}. \qquad (i = 1, 2, \dots, k)$$

Then (5.6) gives

$$\ddot{p}(n) = \ddot{q}(n) + \sum_{i=1}^{k}\left(\ddot{a}(i) + \frac{\ddot{b}(i)}{n}\right)\ddot{p}(n-i). \qquad (n = 1, 2, \dots) \tag{13.1}$$

It is suggested to apply this transformed recursion for n up to v, and then apply the original recursion (5.6).

By letting $q = a \equiv 0$, we obtain that

$$\varphi_{\ddot{p}}(n) = e^{-un}\varphi_p(n). \qquad (n = 1, 2, \dots)$$

The constants u and v can be adjusted so that the transformed function $\ddot{p}$ gets a shape less susceptible to underflow and overflow than the original distribution p. It might be reasonable to start the recursion such that $\ddot{p}(0) = 1$, that is,

$$u = -\frac{\ln p(0)}{v}.$$

It is difficult to decide on how to choose v. To avoid underflow and overflow, one may have to do some trial and error. It could be an idea to include a test in the program that gives a warning when $\ddot{p}(n)$ exceeds a limit close to the maximum number that the computer can handle. Then one could adjust u and v and rescale the values of $\ddot{p}$ that have already been evaluated.

The transformed recursion (13.1) is in the same form as the original recursion (5.6). We need only to change the initial value $p(0)$ to $\ddot{p}(0) = p(0)e^{uv}$ and replace the functions q, a, and b with $\ddot{q}$, $\ddot{a}$, and $\ddot{b}$. In particular when evaluating $\ddot{p}(0)$, one should take care that one does not run into underflow again. Thus, it might be better to first evaluate $\ddot{p}(0)$ by $\ddot{p}(0) = \exp(\ln p(0) + uv)$ instead of using straight forward multiplication.

Now let $f = p \vee h$ with $h \in \mathcal{P}_{11}$ and p being $R_1[a, b]$. Then f satisfies the recursion (2.39), which is a special case of (5.6). The transformed recursion (13.1) then becomes

$$\ddot{f}(x) = \sum_{y=1}^{x}\left(a + b\frac{y}{x}\right)\ddot{h}(y)\ddot{f}(x-y) \qquad (x = 1, 2, \dots)$$

with

$$\ddot{f}(x) = f(x)e^{u(v-x)} \qquad (x = 0, 1, 2, \dots)$$
$$\ddot{h}(y) = h(y)e^{-uy}. \qquad (y = 1, 2, \dots)$$

13.3 Convolutions

A distribution $f \in \mathcal{P}_{1\underline{0}}$ can often be decomposed as the convolution $f = *_{j=1}^{M} f_j$ with $f_1, f_2, \dots, f_M \in \mathcal{P}_{1\underline{0}}$. We assume that all these distributions can be evaluated recursively. If we run into problems with underflow and/or overflow with the recursion for f, then it might be better to evaluate the f_js separately by recursion and finally evaluate f by brute force convolution of the f_js.

If all the f_js are equal to some distribution g, then it might seem tempting to evaluate f by the recursion of Theorem 2.8. However, then we are likely to run into the same underflow problem so that we have to use brute force convolution.

As we have seen in Sect. 5.1, we can do the brute force convolution of two distributions in $\mathcal{F}_{1\underline{0}}$ more efficiently when these distributions are equal. The following algorithm for evaluating $f = g^{M*}$ seems to be the procedure that takes as much as possible advantage of this feature. We introduce the binary expansion $M = \sum_{i=0}^{k_M} w_{Mi} 2^i$ with $w_{Mk_M} = 1$ and $w_{Mi} \in \{0, 1\}$ for $i = 0, 1, 2, \dots, k_M - 1$. Let $g_i = g^{2^i *}$ for $i = 0, 1, 2, \dots, k_M$. We evaluate the g_is recursively by evaluating $g_i = g_{i-1} * g_{i-1}$ by (5.15) for $i = 1, 2, \dots, k_M$, and finally we evaluate $f = *_{\{i: w_{Mi}=1\}} g_i$ by brute force convolution.

In particular, if f is infinitely divisible, then we can choose M freely. It then seems most efficient to let it be a power of two, that is, $M = 2^k$ for some integer k. Then we evaluate $g_1, g_2, \dots, g_k$ recursively as described above, and finally we let $f = g_M$.

Further Remarks and References

The present chapter is based on Panjer and Willmot (1986). However, they restricted their discussion to the compound Panjer case. The binary convolution procedure of Sect. 13.3 is discussed in Sundt and Dickson (2000).

Underflow and overflow in connection with recursive evaluation of distributions is discussed by Waldmann (1994, 1995) for the aggregate claims distribution in the individual life model and by Waldmann (1996) for a compound distribution with counting distribution in $\bigcup_{l=0}^{\infty} \mathcal{S}_l$ and severity distribution in $\mathcal{P}_{11}$. He argues that for numerical reasons it is advantageous to apply recursions for non-negative increasing convex functions, that is, if $f \in \mathcal{P}_{10}$, it could be advantageous to use recursions for Γf and $\Gamma^2 f$ instead of f.

Underflow and overflow are also discussed by Wang and Panjer (1994).

Stability is another numerical issue. When doing calculations on real numbers, a computer will never be completely accurate; there will always be rounding errors. That could be particularly important in connection with recursive evaluation of a function, where the evaluation of any new value of the function will depend on the whole history of evaluation of values of the function done earlier. Thus, the rounding errors could accumulate and affect the numerical stability of the recursion. To give a more or less self-contained account of stability properties of the sort of recursions that we study in the present book, would involve theory rather different from the rest of the book, and we feel that it would break the scope of the book. We therefore rather refer to more specialised literature on that area. Stability of recursions for distributions of the type presented in this book, has been studied by Panjer and Wang (1993, 1995), Wang and Panjer (1993), and Wang (1994, 1995). Gerhold et al. (2008) discuss modifying recursions to get them more stable, but then they also get more time-consuming.

Part II
Multivariate Distributions

Chapter 14
Introduction

Summary

In Chap. 1, we introduced some prerequisites for this book, but restricted to concepts needed in the univariate setting of Part I. In the present chapter, we shall supplement this with concepts needed in the multivariate setting of Part II.

In Sect. 14.1, we discuss how it would be interesting to consider aggregate claims distributions in a multivariate setting.

When working in a multivariate setting, we shall need notation for vectors and matrices. This is the topic of Sect. 14.2.

When extending to a multivariate setting results that we have proved in the univariate case, the proof is often so similar to the proof we presented in the univariate case, that we omit the proof. However, as it might be difficult to see immediately how to extend an induction proof to the multivariate case, we discuss that in Sect. 14.3.

In Sect. 14.4, we extend to the multivariate case the notation of Sect. 1.3 for classes of distributions and other functions.

Sections 14.5 and 14.7 are devoted to convolutions and moments respectively.

Extension of the concept of compound distributions to the multivariate case is the topic of Sect. 14.6. There are two main types of extension, Type 1 with univariate counting distribution and multivariate severity distribution and Type 2 with multivariate counting distribution and univariate severity distribution. We also have a combined Type 3 with multivariate counting distribution and multivariate severity distributions.

14.1 Aggregate Claims Distributions

In Sect. 1.1, we discussed univariate aggregate claims distributions. In this section, we shall extend that discussion to multivariate distributions.

In an individual model, a natural setting could be that each policyholder has some policies, between which we believe that there might be dependence; a careless person could have high fire risk in his home, be a risky car driver, etc. For such cases, we can extend the concept of individual model from Sect. 1.1 by replacing the one-dimensional aggregate claims of each policyholder with a vector of elements representing his aggregate claims in home insurance, his aggregate claims in automobile insurance, etc.

In insurance, it is not always so that the amount of a claim is known and paid immediately when a claim event occurs. Firstly, it can take some time before the claim

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_14, © Springer-Verlag Berlin Heidelberg 2009

is reported to the insurance company. This could e.g. be the case for product liability of a pharmaceutical company. Let us say that a child has got some medical defect, and that it turns out that this is caused by some drug that the mother used during the pregnancy. Depending on the conditions of the policy, this could be covered by the policy that was in force when the damage was caused, not when it was discovered. Also, after a claim is made, it can take long until the final amount is known, and in the meantime there could be made partial payments. In personal injury, it could take long until one knows the total costs of the treatment and the final outcome of the injury. Furthermore, there could be disagreement between the insurance company and the injured person, so that the final amount would have to be settled by a court.

In such cases, one could split the aggregate claim amount of a policy into a vector whose first element is the payments made in the year of the claim event, the second element the payments made next year, and so on.

Analogously, one could let the first element of the vector be the amount paid by the cedant after reinsurance, the second element the amount paid by the first reinsurer, and so on. This would not work with stop loss reinsurance where the part paid by the reinsurer is not determined by the aggregate claims of the individual policy, only by the aggregate claims of the whole portfolio.

The concept of splitting between years of payment can of course also be applied in a collective model. Like in the univariate case, we let the counting variable be the number of claims. In the univariate case, each of these claims generated a one-dimensional amount whereas now we split this amount into a vector where the first element is the payment made in the year of the claim event, the second amount is the payment made in the next year, and so on. We assume that these vectors are mutually independent and identically distributed and independent of the number of claims, that is, the aggregate claims distribution is now a compound distribution with univariate counting distribution and multivariate severity distribution.

This way of modelling could also be applied with the concept of splitting between what is paid by the cedant and the various reinsurers, provided that the part paid of each claim is determined only by that claim, which is not the case with stop loss reinsurance.

As an example, we assume that there is only one reinsurer, and that the reinsurance is an unlimited excess of loss reinsurance that for each claim pays the amount that exceeds a fixed retention. This is like stop loss reinsurance with the difference that in excess of loss reinsurance, the retention is applied on each claim whereas in stop loss reinsurance it is applied only on the aggregate claims of the whole portfolio.

A natural question is now, is it of interest to know the joint distribution of what is covered by the cedant and what is paid by the reinsurer? Perhaps not too interesting in the present case, but let us modify the reinsurance treaty by adding an *aggregate limit*, that is, a fixed upper limit on the payments to be made by the reinsurer. The excess of the aggregate limit is paid by the cedant unless he has got reinsurance protection elsewhere. Let N denote the number of claims, Y_i the amount of the ith of these claims, r the retention, and l the aggregate limit. Before the aggregate limit, the amount of the nth claim paid by the cedant is $Y_{i1} = \min(Y_i, r)$, and the reinsurer pays $Y_{i2} = (Y_i - r)_+$. The aggregate claims vector without the aggregate limit is $\mathbf{X} = (X_1, X_2)' = \sum_{i=1}^{N} \mathbf{Y}_i$ with $\mathbf{Y}_i = (Y_{i1}, Y_{i2})'$, and the part paid by the cedant after the aggregate limit, is $Z = X_1 + (X_2 - l)_+$. It seems that to be able to evaluate the distribution of Z, we have to go by the distribution of $\mathbf{X}$. We shall return to this situation in Example 15.4.

Let us now consider a situation where a claim event can cause claims on more than one policy. A typical example is hurricane insurance. In such situations, we could use a collective model where claim event n produces a claim amount vector whose jth element is the claim amount caused to the jth policy by this claim event. We assume that these vectors are mutually independent and identically distributed and independent of the number of claims.

In the first example of individual models discussed in this section, the aggregate claims vector was comprised of aggregate claims for different types of insurance, e.g. home insurance and automobile insurance. In the collective models we have mentioned till now in this section, we have assumed that the aggregate claims distribution is a compound distribution with univariate counting distribution and multivariate severity distribution. In principle, we can apply such a model also in the present case, letting the counting variable be the number of claim events and the ith severity vector the vector whose jth element is the amount caused to insurance type j by this event. However, in this setting, we would typically have that most of these vectors would have only one non-zero element; it is not that common that a claim event hits both home insurance and automobile insurance. In this setting, it seems more natural to consider a collective model where the aggregate claims distribution is a compound distribution with a multivariate, say m-variate, claim number distribution, and m univariate severity distributions. Here we have an m-variate claim number vector whose jth element is the number of claims to insurance type j. Furthermore, it is assumed that all the claim amounts are mutually independent and independent of the claim number vector, and that the claim amounts for the same type of insurance are identically distributed.

We have now introduced two different types of compound distributions for aggregate claims in collective models, the first one with univariate counting distribution and multivariate severity distribution and the second one with multivariate counting distribution and univariate severity distributions. We can also combine these two settings; in the setting of the previous paragraph, we could assume that each claim amount is replaced with a vector with that amount split on payment year, and the dimension of these severity vectors could be different for different insurance types.

14.2 Vectors and Matrices

We shall normally denote the dimension of aggregate claims vectors by m.

We shall denote an $m \times 1$ vector by a bold letter and for $j = 1, 2, \dots, m$ its jth element by the corresponding italic with subscript j; when the subscript is $\bullet$, we mean the sum of these elements, e.g. the column vector $\mathbf{x}$ has elements $x_1, x_2, \dots, x_m$ and we have $x_{\bullet} = x_{\bullet m}$. For $j = 1, 2, \ldots, m$, we let $\mathbf{e}_j$ denote the vector whose jth element is equal to one and all other elements equal to zero, and introduce the notation $\mathbf{e}_{j_1 j_2 \cdots j_s} = \sum_{r=1}^{s} \mathbf{e}_{j_r}$ for different integers $j_1, j_2, \ldots, j_s \in \{1, 2, \ldots, m\}$. In particular, we let $\mathbf{e} = \mathbf{e}_{12\cdots m}$. Furthermore, we introduce the zero vector $\mathbf{0}$ with all elements equal to zero. We hope that it will not cause problems for the readers that the dimension m does not appear explicitly in some of the vector notation.

For a sequence of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots,$ we let $\mathbf{x}_{\bullet n} = \sum_{j=1}^{n} \mathbf{x}_j$ and, for $i = 1, 2, \ldots, m$, we denote the ith element of this vector by $x_{\bullet ni}$; the sum of these elements is denoted by $x_{\bullet n \bullet}$.

We let $\mathbb{Z}_m$ and $\mathbb{R}_m$ denote the set of all $m \times 1$ vectors with integer-valued and real-valued elements respectively. If $\mathbf{x}, \mathbf{y} \in \mathbb{R}_m$, then by $\mathbf{y} < \mathbf{x}$, we shall mean that $y_j < x_j$ for $j = 1, 2, \ldots, m$, and by $\mathbf{y} \leq \mathbf{x}$ that $y_j \leq x_j$ for $j = 1, 2, \ldots, m$. When indicating a range like e.g. $\mathbf{0} \leq \mathbf{y} \leq \mathbf{x}$ for fixed $\mathbf{x} \in \mathbb{Z}_m$, it will always be tacitly assumed that the vector $\mathbf{y}$ has integer-valued elements. We let $\mathbb{N}_m = \{\mathbf{x} \in \mathbb{Z}_m : \mathbf{x} \geq \mathbf{0}\}$ and $\mathbb{N}_{m+} = \{\mathbf{x} \in \mathbb{Z}_m : \mathbf{x} > \mathbf{0}\}$.

Matrices will be denoted by bold upper-case letters. The columns of a matrix will be denoted by the corresponding bold lower-case letter with a subscript indicating the number of the column. We denote the elements of a matrix by the corresponding italic with subscripts indicating the row and the column. Thus, if $\mathbf{A}$ is an $m \times s$ matrix, then, for $j = 1, 2, \ldots, s$, $\mathbf{a}_j$ denotes its jth column, and, for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, s$, a_{ij} denotes the element in the ith row and jth column; using $\bullet$ to indicate summation like we did for vectors, we let $a_{i\bullet}$ denote the sum of the elements in the ith row, $a_{\bullet j}$ the sum of the elements of the jth column, and $a_{\bullet\bullet}$ the sum of the elements of the whole matrix.

14.3 Induction Proofs in a Multivariate Setting

In the following chapters, we shall present multivariate versions of many of the results that we have proved in the univariate case in Part I. We shall often omit the proof of such an extension when it is a straight-forward modification of the proof of the univariate case. Some of these omitted proofs involve induction, so let us indicate how such induction proofs can be extended from the univariate case to the multivariate case.

Let us assume that in the univariate case, we want to prove by induction the hypothesis that some result holds for all non-negative integers x. Then we first show that it holds for $x = 0$. Next, we show that if it holds for all non-negative integers x less than some positive integer z, then it must also hold for $x = z$. When this is the case, then the result must hold for all non-negative integers x. Hence, our hypothesis is proved.

In the multivariate extension, we want to prove by induction the hypothesis that the result holds for all $\mathbf{x} \in \mathbb{N}_m$. Then we first show that it holds for $\mathbf{x} = \mathbf{0}$. Next, we show that if it holds for all $\mathbf{x} \in \mathbb{N}_m$ such that $\mathbf{0} \leq \mathbf{x} < \mathbf{z}$ for some $\mathbf{z} \in \mathbb{N}_{m+}$, then it must also hold for $\mathbf{x} = \mathbf{z}$. When this is the case, then the result must hold for all $\mathbf{x} \in \mathbb{N}_m$. Hence, our hypothesis is proved.

If, in the univariate case, we want to prove the hypothesis that the result holds for all positive integers, then we proceed as indicated above, but start by showing that the result holds for $x = 1$.

In the multivariate case, this corresponds to proving the hypothesis that the result holds for all $\mathbf{x} \in \mathbb{N}_{m+}$. We now start by showing that for $j = 1, 2, \ldots, m$, it holds for $\mathbf{x} = \mathbf{e}_j$. Next, we show that if it holds for all $\mathbf{x} \in \mathbb{N}_{m+}$ such that $\mathbf{e}_j \leq \mathbf{x} < \mathbf{z}$ for some $\mathbf{z} \in \mathbb{N}_{m+}$, then it must also hold for $\mathbf{x} = \mathbf{z}$.

In the present book, an induction proof in a multivariate setting following the ideas outlined in the present section, has been written out for Theorem 16.4.

14.4 Classes of Distributions and Functions

In Part II, we shall mainly concentrate on distributions on vectors with integer-valued elements. Our main attention will be on the probability function of such distributions, and then it will be convenient to associate the distribution with its probability function. Hence, when referring to a distribution, we shall normally mean its probability function. Such functions will be denoted by lower case italics.

We let $\mathcal{P}_m$ denote the class of all distributions on $\mathbb{Z}_m$. For all $\mathbf{l} \in \mathbb{Z}_m$, we let $\mathcal{P}_{m\mathbf{l}}$ denote the class of all distributions $f \in \mathcal{P}_m$ for which $f(\mathbf{x}) = 0$ when $\mathbf{x} < \mathbf{l}$. By $\mathcal{P}_{m\underline{\mathbf{l}}}$, we denote the class of all distributions in $\mathcal{P}_{m\mathbf{l}}$ with a positive mass at $\mathbf{l}$. Let $\mathcal{P}_{m\_} = \bigcup_{\mathbf{l} \in \mathbb{Z}_m} \mathcal{P}_{m\underline{\mathbf{l}}}$, that is, the set of all distributions in $\mathcal{P}_m$ whose support is bounded from below. We let $\mathcal{P}_{m+} = \bigcup_{j=1}^{m} \mathcal{P}_{m\mathbf{e}_j}$, that is, the set of distributions on $\mathbb{N}_{m+}$; in particular, we have $\mathcal{P}_{1+} = \mathcal{P}_{11}$.

We let $\mathcal{F}_m$ denote the class of all functions on $\mathbb{Z}_m$. For all $\mathbf{l} \in \mathbb{Z}_m$, we let $\mathcal{F}_{m\mathbf{l}}$ denote the set of all functions $f \in \mathcal{F}_m$ for which $f(\mathbf{x}) = 0$ when $\mathbf{x} < \mathbf{l}$, and we let $\mathcal{F}_{m\underline{\mathbf{l}}}$ denote the set of functions $f \in \mathcal{F}_{m\mathbf{l}}$ with a positive mass at $\mathbf{l}$. Let $\mathcal{F}_{m\_} = \bigcup_{\mathbf{l} \in \mathbb{Z}_m} \mathcal{F}_{m\underline{\mathbf{l}}}$. We let $\mathcal{F}_{m+} = \bigcup_{j=1}^{m} \mathcal{F}_{m\mathbf{e}_j}$.

We introduce the cumulation operator Γ given by $\Gamma f(\mathbf{x}) = \sum_{\mathbf{y} \leq \mathbf{x}} f(\mathbf{y})$ for all $\mathbf{x} \in \mathbb{Z}_m$ and all functions $f \in \mathcal{F}_m$ for which the summation exists. When $f \in \mathcal{F}_{m\_}$, Γf always exists as then there will always exist a vector $\mathbf{l} \in \mathbb{Z}_m$ such that $\Gamma f(\mathbf{x}) = \sum_{\mathbf{l} \leq \mathbf{y} \leq \mathbf{x}} f(\mathbf{y})$ for all $\mathbf{x} \in \mathbb{Z}_m$ such that $\mathbf{x} \geq \mathbf{l}$.

The operator Φ is applied to functions in $\mathcal{F}_{m\mathbf{0}}$ and gives the function multiplied by the sum of its arguments, that is, if $f \in \mathcal{F}_{m\mathbf{0}}$, then $\Phi f \in \mathcal{F}_{m\mathbf{0}}$ is given by $\Phi f(\mathbf{x}) = x_\bullet f(\mathbf{x})$ for $\mathbf{x} \in \mathbb{N}_m$.

14.5 Convolutions

The definition of convolution trivially extends to the m-variate case.

We define the convolution $f * g$ of two functions $f, g \in \mathcal{F}_m$ by

$$(f * g)(\mathbf{z}) = \sum_{\mathbf{x} \in \mathbb{Z}_m} f(\mathbf{x}) g(\mathbf{z} - \mathbf{x}). \qquad (\mathbf{z} \in \mathbb{Z}_m)$$

Like in the univariate case, we let $f^{M*} = f^{(M-1)*} * f$ for $M = 1, 2, \dots$ with f^{0*} being the function concentrated in $\mathbf{0}$ with mass one.

If $f, g \in \mathcal{F}_{m\underline{\mathbf{0}}}$, then

$$(f * g)(\mathbf{z}) = \sum_{\mathbf{0} \le \mathbf{x} \le \mathbf{z}} f(\mathbf{x}) g(\mathbf{z} - \mathbf{x}); \qquad (\mathbf{z} \in \mathbb{N}_m)$$

in particular, we have $(f * g)(\mathbf{0}) = f(\mathbf{0}) g(\mathbf{0})$.

Now let $f, g \in \mathcal{P}_m$. If $\mathbf{X}$ and $\mathbf{Y}$ are random vectors with distribution f and g respectively, then $f * g$ is the distribution of $\mathbf{X} + \mathbf{Y}$.

In terms of cumulative distributions, we define the convolution $F * G$ of two m-variate distributions F and G by

$$(F * G)(\mathbf{z}) = \int_{\mathbb{R}_m} G(\mathbf{z} - \mathbf{x}) \, \mathrm{d}F(\mathbf{x}), \qquad (\mathbf{x} \in \mathbb{R}_m)$$

and we have $F^{M*} = F^{(M-1)*} * F$ for $M = 1, 2, \dots$ with F^{0*} being the distribution concentrated in zero.

14.6 Compound Distributions

As discussed in Sect. 14.1, we can extend the definition of univariate compound distributions of Sect. 1.6 to multivariate distributions in the following two ways:

1. *Univariate counting distribution and multivariate severity distribution*. The extension of the definition of univariate compound distributions of Sect. 1.6 to this case is trivial; the compound distribution $p \vee H$ where $p \in \mathcal{P}_{10}$ and H is a multivariate distribution, is still given by (1.5), and if $h \in \mathcal{P}_m$, then (1.6) still holds. If $h \in \mathcal{P}_{m+}$, then

$$(p \vee h)(\mathbf{x}) = \sum_{n=0}^{x_\bullet} p(n) h^{n*}(\mathbf{x}). \qquad (\mathbf{x} \in \mathbb{N}_m) \tag{14.1}$$

2. *Multivariate counting distribution and univariate severity distributions*. Let $p \in \mathcal{P}_{m\mathbf{0}}$ and $\mathbf{H} = (H_1, H_2, \dots, H_m)$ where $H_1, H_2, \dots, H_m$ are univariate distributions. Then the compound distribution $p \vee \mathbf{H}$ with counting distribution p and severity distributions $\mathbf{H}$ is given by

$$(p \vee \mathbf{H})(\mathbf{x}) = \sum_{\mathbf{n} \in \mathbb{N}_m} p(\mathbf{n}) \prod_{j=1}^{m} H_j^{n_j *}(x_j). \qquad (\mathbf{x} \in \mathbb{R}_m)$$

   If $\mathbf{h} = (h_1, h_2, \dots, h_m)$ with $h_1, h_2, \dots, h_m \in \mathcal{P}_1$, then $p \vee \mathbf{h}$ is given by

$$(p \vee \mathbf{h})(\mathbf{x}) = \sum_{\mathbf{n} \in \mathbb{N}_m} p(\mathbf{n}) \prod_{j=1}^{m} h_j^{n_j *}(x_j), \qquad (\mathbf{x} \in \mathbb{Z}_m) \tag{14.2}$$

and in the special case $h_1, h_2, \ldots, h_m \in \mathcal{P}_{11}$, we obtain

$$(p \vee \mathbf{h})(\mathbf{x}) = \sum_{\mathbf{0} \le \mathbf{n} \le \mathbf{x}} p(\mathbf{n}) \prod_{j=1}^{m} h_j^{n_j *}(x_j); \qquad (\mathbf{x} \in \mathbb{N}_m)$$

in particular, this gives $(p \vee \mathbf{h})(\mathbf{x}) = p(\mathbf{0})$.

As Type 2 is a compounding of a multivariate counting distribution, we could call that a compound multivariate distribution. On the other hand, Type 1 is a multivariate compounding of a univariate counting distribution, so we could call that a multivariate compound distribution. However, it can also be convenient to have one term covering both these multivariate types of compound distributions, and then it seems natural to say multivariate compound distribution as it is a multivariate version of a compound distribution. Hence, we shall refer to the two types simply as Types 1 and 2 and let the term multivariate compound distribution cover both cases.

In Sect. 14.1, we discussed various insurance applications of these two types of compound distributions. Furthermore, we pointed out that the two types can be combined by letting the h_js in Type 2 be multivariate. Let us call that combination a multivariate compound distribution of Type 3. Letting $m = 1$ in that setting, brings us to Type 1.

Let us try to compare Types 1 and 2. We consider an insurance portfolio with m policies. Let $f \in \mathcal{P}_m$ be the joint distribution of the aggregate claims of each of the m policies during a given period; there could be dependences between these variables.

We see that in both approaches, the claim amounts from one policy are independent of the claim number(s) and mutually independent and identically distributed.

Type 1 is more general than Type 2 in the sense that in Type 2 the claim amounts of different policies are always independent. On the other hand, Type 1 is less general in the sense that there we have more specific assumptions on the occurrence of claim events.

From the above discussion, it seems that it would be possible to find cases that could be modelled in both ways if we have independent claim amounts from different policies and assumptions on claim events. Let us look at such a case. We let $p_1 \in \mathcal{P}_1$ and $\mathbf{h}$ be defined as above and introduce an *m-variate Bernoulli distribution* (that is, each variable can take only the values zero and one) $q \in \mathcal{P}_m$. Then we let $p_m = p_1 \vee q$ and $h = q \vee \mathbf{h}$. Both these distributions are in $\mathcal{P}_m$, and we have $f = p_1 \vee h = p_m \vee \mathbf{h}$. When letting all the h_js be concentrated in one, h reduces to q and f to p_m.

Example 14.1 Let us consider the special case when q is given by

$$q(\mathbf{e}_j) = \pi_j. \qquad (0 \le \pi_j \le 1;\ j = 1, 2, \ldots, m;\ \pi_\bullet = 1) \tag{14.3}$$

Then, for $M = 0, 1, 2, \ldots$, we have that q^{M*} is the *multinomial distribution* $\mathrm{mnom}(M, \boldsymbol{\pi})$ given by

$$q^{M*}(\mathbf{n}) = M! \prod_{j=1}^{m} \frac{\pi_j^{n_j}}{n_j!}. \tag{14.4}$$

$$(\mathbf{0} < \mathbf{n} \leq M\mathbf{e};\ n_\bullet = M;\ 0 \leq \pi_j \leq 1;\ j = 1, 2, \ldots, m;\ \pi_\bullet = 1)$$

Hence, for all $\mathbf{n} \in \mathbb{N}_m$, $q^{s*}(\mathbf{n}) > 0$ only when $s = n_\bullet$, so that (14.1) gives

$$p_m(\mathbf{n}) = p_1(n_\bullet) q^{n_\bullet *}(\mathbf{n}) = p_1(n_\bullet) n_\bullet! \prod_{j=1}^{m} \frac{\pi_j^{n_j}}{n_j!}. \qquad (\mathbf{n} \in \mathbb{N}_m) \tag{14.5}$$

Furthermore,

$$h(\mathbf{y}) = \pi_j h_j(y_j). \qquad (\mathbf{y} = y_j \mathbf{e}_j;\ y_j = 1, 2, \ldots;\ j = 1, 2, \ldots, m) \tag{14.6}$$

□

We extend the definition (1.6) of $p \vee h$ to $p \in \mathcal{F}_{10}$ and $h \in \mathcal{F}_m$ and the definition (14.2) of $p \vee \mathbf{h}$ to $p \in \mathcal{F}_{m\mathbf{0}}$ and $h_1, h_2, \ldots, h_m \in \mathcal{F}_1$.

14.7 Moments

Let $\mathbf{X}$ be an $m \times 1$ random vector with distribution F. For $\mathbf{j} \in \mathbb{Z}_m$ and $\mathbf{c} \in \mathbb{R}_m$, we denote the $\mathbf{j}$th order moment of F around $\mathbf{c}$ by $\mu_F(\mathbf{j}; \mathbf{c})$, that is,

$$\mu_F(\mathbf{j}; \mathbf{c}) = \mathrm{E} \prod_{i=1}^{m} (X_i - c_i)^{j_i} = \int_{\mathbb{R}_m} \left(\prod_{i=1}^{m} (x_i - c_i)^{j_i} \right) \mathrm{d}F(\mathbf{x}).$$

In particular, we have $\mu_F(\mathbf{0}; \mathbf{c}) = 1$ for any value of $\mathbf{c}$. For convenience, we let $\mu_F(\mathbf{j}, \mathbf{0}) = \mu_F(\mathbf{j})$.

If $f \in \mathcal{F}_m$, then we let

$$\mu_f(\mathbf{j}; \mathbf{c}) = \sum_{\mathbf{x} \in \mathbb{Z}_m} f(\mathbf{x}) \prod_{i=1}^{m} (x_i - c_i)^{j_i}$$

and $\mu_f(\mathbf{j}) = \mu_f(\mathbf{j}; \mathbf{0})$.

Further Remarks and References

The discussion in Sect. 14.6 on the two types of compound distributions is based on Sect. 3 in Sundt and Vernic (2004).

Johnson et al. (1997) give a broad overview of discrete multivariate distributions and their properties. Kocherlakota and Kocherlakota (1992) consider discrete bivariate distributions.

Chapter 15
Multivariate Compound Distributions of Type 1

Summary

The main purpose of the present chapter is to develop recursions for multivariate compound distributions of Type 1, that is, compound distributions with univariate counting distribution and multivariate severity distribution.

We first give some results on covariances in Sect. 15.1.

Section 15.2 is devoted to the case where the counting distribution belongs to the Panjer class. We give some examples, in particular related to reinsurance. Like in the univariate case, we can deduce recursions for convolutions of multivariate distributions from recursions for compound binomial distributions. This is done in Sect. 15.3.

A short Sect. 15.4 is devoted to infinitely divisible distributions. Its main result is a multivariate extension of Theorem 4.2.

In Sect. 15.5, we extend some recursions from Chap. 5 to multivariate severity distributions.

Finally, in Sect. 15.6, we deduce some recursions for compound distributions with univariate counting distribution and multivariate Bernoulli severity distribution.

15.1 Covariances

Let $\mathbf{Y}_1, \mathbf{Y}_2, \dots$ be mutually independent and identically distributed non-degenerate random $m \times 1$ vectors with non-negative elements, independent of the non-degenerate random variable N. In this section, we shall study the covariance between two elements X_j and X_k of $\mathbf{X} = \mathbf{Y}_{\bullet N}$. We have

$$\mathrm{Cov}(X_j, X_k) = \mathrm{E}\,\mathrm{Cov}[Y_{\bullet Nj}, Y_{\bullet Nk}|N] + \mathrm{Cov}(\mathrm{E}[Y_{\bullet Nj}|N], \mathrm{E}[Y_{\bullet Nk}|N]),$$

which gives

$$\begin{aligned}\mathrm{Cov}(X_j, X_k) &= \mathrm{E}\,N\,\mathrm{Cov}(Y_{1j}, Y_{1k}) + \mathrm{E}\,Y_{1j}\,\mathrm{E}\,Y_{1k}\,\mathrm{Var}\,N \\ &= \mathrm{E}\,N\,\mathrm{E}\,Y_{1j}Y_{1k} + \mathrm{E}\,Y_{1j}\,\mathrm{E}\,Y_{1k}(\mathrm{Var}\,N - \mathrm{E}\,N).\end{aligned} \qquad (15.1)$$

If Y_{1j} and Y_{1k} are uncorrelated, then

$$\mathrm{Cov}(X_j, X_k) = \mathrm{E}\,Y_{1j}\,\mathrm{E}\,Y_{1k}\,\mathrm{Var}\,N > 0,$$

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_15, © Springer-Verlag Berlin Heidelberg 2009

that is, X_j and X_k are positively correlated. This seems intuitively reasonable; if X_j is large, then that could indicate that N is large, and a large value of N could indicate a large value of X_k.

If N is Poisson distributed, then $\mathrm{Var}\, N = \mathrm{E}\, N$ and the last expression in (15.1) reduces to

$$\mathrm{Cov}(X_j, X_k) = \mathrm{E}\, N\, \mathrm{E}\, Y_{1j} Y_{1k} \geq 0,$$

with equality iff $\mathrm{E}\, Y_{1j} Y_{1k} = 0$, that is, when the probability that both Y_{1j} and Y_{1k} are positive, is zero. In the situation when the elements of the severity vectors are claim amounts of different policies from the same claim event, this means that the same claim event cannot affect both policy j and k.

More generally, if we leave the Poisson assumption, but keep the assumption that $\mathrm{E}\, Y_{1j} Y_{1k} = 0$, the second expression in (15.1) gives

$$\mathrm{Cov}(X_j, X_k) = \mathrm{E}\, Y_{1j}\, \mathrm{E}\, Y_{1k} (\mathrm{Var}\, N - \mathrm{E}\, N),$$

that is, the sign of $\mathrm{Cov}(X_j, X_k)$ is the same as the sign of $\mathrm{Var}\, N - \mathrm{E}\, N$. From the discussion at the end of Sect. 2.3.2 follows that this covariance is positive when the counting distribution is negative binomial, equal to zero when the counting distribution is Poisson (as we have already shown), and negative when the counting distribution is binomial.

15.2 Counting Distribution in the Panjer Class

In this section, we shall consider evaluation of $f = p \vee h$ where p is $R_1[a, b]$ and $h \in \mathcal{P}_{m\mathbf{0}}$. In the univariate case, our faithful workhorse was (5.2), so a crucial question seems to be how to extend this relation to $m > 1$.

Let the random $m \times 1$ vectors $\mathbf{Y}_1, \mathbf{Y}_2, \ldots$ be mutually independent and identically distributed with distribution h and independent of the random variable N, which has distribution p. Then $\mathbf{X} = \mathbf{Y}_{\bullet N}$ has distribution f. With analogous reasoning to what we used to set up (5.2), we obtain

$$\mathrm{E}[a x_j + b Y_{1j} | \mathbf{Y}_{\bullet n} = \mathbf{x}] = \left(a + \frac{b}{n}\right) x_j. \tag{15.2}$$

$$(\mathbf{x} \in \mathbb{N}_m;\ n = 1, 2, \ldots;\ j = 1, 2, \ldots, m)$$

This gives that for $\mathbf{x} \in \mathbb{N}_{m+}$ and $j = 1, 2, \ldots, m$,

$$x_j f(\mathbf{x}) = x_j \sum_{n=1}^{\infty} p(n) h^{n*}(\mathbf{x}) = x_j \sum_{n=1}^{\infty} p(n-1)\left(a + \frac{b}{n}\right) h^{n*}(\mathbf{x})$$

$$= \sum_{n=1}^{\infty} p(n-1)\, \mathrm{E}[a x_j + b Y_{1j} | \mathbf{Y}_{\bullet n} = \mathbf{x}] h^{n*}(\mathbf{x})$$

$$
\begin{aligned}
&= \sum_{n=1}^{\infty} p(n-1) \sum_{\mathbf{0} \le \mathbf{y} \le \mathbf{x}} (ax_j + by_j) h(\mathbf{y}) h^{(n-1)*}(\mathbf{x}-\mathbf{y}) \\
&= \sum_{\mathbf{0} \le \mathbf{y} \le \mathbf{x}} (ax_j + by_j) h(\mathbf{y}) \sum_{n=1}^{\infty} p(n-1) h^{(n-1)*}(\mathbf{x}-\mathbf{y}) \\
&= \sum_{\mathbf{0} \le \mathbf{y} \le \mathbf{x}} (ax_j + by_j) h(\mathbf{y}) f(\mathbf{x}-\mathbf{y}) \\
&= ax_j h(\mathbf{0}) f(\mathbf{x}) + \sum_{\mathbf{0} < \mathbf{y} \le \mathbf{x}} (ax_j + by_j) h(\mathbf{y}) f(\mathbf{x}-\mathbf{y}),
\end{aligned}
$$

from which we obtain

$$
x_j f(\mathbf{x}) = \frac{1}{1 - ah(\mathbf{0})} \sum_{\mathbf{0} < \mathbf{y} \le \mathbf{x}} (ax_j + by_j) h(\mathbf{y}) f(\mathbf{x}-\mathbf{y}). \qquad (\mathbf{x} \in \mathbb{N}_{m+};\ j = 1, 2, \dots, m) \tag{15.3}
$$

When $x_j > 0$, we can divide by x_j, and this gives

$$
f(\mathbf{x}) = \frac{1}{1 - ah(\mathbf{0})} \sum_{\mathbf{0} < \mathbf{y} \le \mathbf{x}} \left(a + b \frac{y_j}{x_j} \right) h(\mathbf{y}) f(\mathbf{x}-\mathbf{y}). \qquad (\mathbf{x} \ge \mathbf{e}_j;\ j = 1, 2, \dots, m) \tag{15.4}
$$

For any $\mathbf{x} \in \mathbb{N}_{m+}$, there must exist at least one $j \in \{1, 2, \dots, m\}$ for which $x_j > 0$. Hence, together with the initial condition $f(\mathbf{0}) = \tau_p(h(\mathbf{0}))$, (15.4) gives a procedure for recursive evaluation of $f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{N}_m$.

If $Y_{1j} > 0$ almost surely, then $h(\mathbf{x}) = f(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathbb{N}_{m+}$ with $x_j = 0$. In this case, we do not need to apply (15.4) for any other j. However, if there exists no j such that $Y_{1j} > 0$ almost surely, then we cannot apply the same j in (15.4) for all $\mathbf{x} \in \mathbb{N}_{m+}$. This complicates programming the procedure. However, we shall now see how we can avoid that problem.

Let $\mathbf{c}$ be an $m \times 1$ vector. Multiplication of (15.3) by c_j and summation over j gives

$$
\mathbf{c}'\mathbf{x} f(\mathbf{x}) = \frac{1}{1 - ah(\mathbf{0})} \sum_{\mathbf{0} < \mathbf{y} \le \mathbf{x}} (a\mathbf{c}'\mathbf{x} + b\mathbf{c}'\mathbf{y}) h(\mathbf{y}) f(\mathbf{x}-\mathbf{y}). \qquad (\mathbf{x} \in \mathbb{N}_{m+})
$$

When $\mathbf{c}'\mathbf{x} \neq 0$, we can divide by $\mathbf{c}'\mathbf{x}$ and obtain

$$
f(\mathbf{x}) = \frac{1}{1 - ah(\mathbf{0})} \sum_{\mathbf{0} < \mathbf{y} \le \mathbf{x}} \left(a + b \frac{\mathbf{c}'\mathbf{y}}{\mathbf{c}'\mathbf{x}} \right) h(\mathbf{y}) f(\mathbf{x}-\mathbf{y}). \qquad (\mathbf{x} \in \mathbb{N}_{m+};\ \mathbf{c}'\mathbf{x} \neq 0)
$$

For $\mathbf{c} = \mathbf{e}$, we have $\mathbf{c}'\mathbf{x} = x_\bullet$, which is greater than zero for all $\mathbf{x} \in \mathbb{N}_{m+}$. Hence,

$$f(\mathbf{x}) = \frac{1}{1 - ah(\mathbf{0})} \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} \left(a + b\frac{y_\bullet}{x_\bullet}\right) h(\mathbf{y}) f(\mathbf{x} - \mathbf{y}). \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{15.5}$$

Example 15.1 Let us return to the setting of Example 14.1 with the additional assumption that p_1 is $R_1[a, b]$. Insertion of (14.6) in (15.4) and (15.5) gives

$$f(\mathbf{x}) = a \sum_{k=1}^{m} \pi_k \sum_{y=1}^{x_k} h_k(y) f(\mathbf{x} - y\mathbf{e}_k) + b\frac{\pi_j}{x_j} \sum_{y=1}^{x_j} y h_j(y) f(\mathbf{x} - y\mathbf{e}_j)$$
$$(\mathbf{x} \geq \mathbf{e}_j;\ j = 1, 2, \ldots, m) \tag{15.6}$$

$$f(\mathbf{x}) = \sum_{j=1}^{m} \pi_j \sum_{y=1}^{x_j} \left(a + b\frac{y}{x_\bullet}\right) h_j(y) f(\mathbf{x} - y\mathbf{e}_j). \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{15.7}$$

When letting all the h_js be concentrated in one, h reduces to q and f to p_m. Hence, we obtain

$$p_m(\mathbf{n}) = a \sum_{k=1}^{m} \pi_k p_m(\mathbf{n} - \mathbf{e}_k) + b\frac{\pi_j}{n_j} p_m(\mathbf{n} - \mathbf{e}_j)$$
$$(\mathbf{n} \geq \mathbf{e}_j;\ j = 1, 2, \ldots, m) \tag{15.8}$$

$$p_m(\mathbf{n}) = \left(a + \frac{b}{n_\bullet}\right) \sum_{j=1}^{m} \pi_j p_m(\mathbf{n} - \mathbf{e}_j). \qquad (\mathbf{n} \in \mathbb{N}_{m+}) \tag{15.9}$$

When p_1 is the Poisson distribution Po(λ), that is, $a = 0$ and $b = \lambda$ (see Table 2.1), (15.6) reduces to

$$f(\mathbf{x}) = \lambda\frac{\pi_j}{x_j} \sum_{y=1}^{x_j} y h_j(y) f(\mathbf{x} - y\mathbf{e}_j). \qquad (\mathbf{x} \geq \mathbf{e}_j;\ j = 1, 2, \ldots, m)$$

By comparison with Theorem 2.2, we see that there exists a function $\tilde{f}_j$ such that

$$f(\mathbf{x}) = f_j(x_j)\tilde{f}_j(x_1, x_2, \ldots, x_{j-1}, x_{j+1}, \ldots, x_m),$$
$$(\mathbf{x} \geq \mathbf{e}_j;\ j = 1, 2, \ldots, m)$$

where f_j is the compound Poisson distribution with Poisson parameter $\lambda\pi_j$ and severity distribution h_j. Then f_j must be the marginal distribution of X_j and $\tilde{f}_j$ the joint distribution of the other elements of $\mathbf{X}$. It follows that in this case, the X_js are independent and each X_j compound Poisson distributed with Poisson parameter $\lambda\pi_j$ and severity distribution h_j. □

Example 15.2 Let us now consider the bivariate case with $m = 2$. With some abuse of notation, to avoid carrying subscripts 1 and 2 all the time, we assume that the severity vectors $(U_1, V_1)', (U_2, V_2)', \dots$ are mutually independent, identically distributed with distribution $h \in \mathcal{P}_{2+}$, and independent of the counting variable N whose distribution p is $R_1[a, b]$. Then $(X, Y)' = (U_{\bullet N}, V_{\bullet N})'$ has distribution $f = p \vee h$.

In this case, (15.4) gives

$$f(x, y) = \sum_{u=0}^{x}\left(a + b\frac{u}{x}\right)\sum_{v=0}^{y} h(u, v) f(x-u, y-v) \qquad (15.10)$$

$$(x = 1, 2, \dots; y = 0, 1, 2, \dots)$$

$$f(x, y) = \sum_{v=0}^{y}\left(a + b\frac{v}{y}\right)\sum_{u=0}^{x} h(u, v) f(x-u, y-v), \qquad (15.11)$$

$$(x = 0, 1, 2, \dots; y = 1, 2, \dots)$$

and from (15.5) we obtain

$$f(x, y) = \sum_{u=0}^{x}\sum_{v=0}^{y}\left(a + b\frac{u+v}{x+y}\right) h(u, v) f(x-u, y-v). \qquad (15.12)$$

$$((x, y) > (0, 0))$$

Note that to simplify the formulae, we have, unlike in (15.4) and (15.5), included $(u, v) = (0, 0)$. This does not change the value as $h(0, 0) = 0$.

To assume that U_1 and V_1 are independent does not seem to give any significant simplification; with h_U and h_V denoting their marginal distributions, (15.10) gives

$$f(x, y) = \sum_{u=0}^{x}\left(a + b\frac{u}{x}\right) h_U(u) \sum_{v=0}^{y} h_V(v) f(x-u, y-v).$$

$$(x = 1, 2, \dots; y = 0, 1, 2, \dots)$$

□

Example 15.3 Let us now consider a univariate situation where N is the number of claims in an insurance portfolio and W_i the amount of the ith of these claims. We assume that the distribution p of N is $R_1[a, b]$ and that the W_is are independent of N and mutually independent and identically distributed with distribution $g \in \mathcal{P}_{11}$. We want to evaluate the joint distribution f of the aggregate claim amount $X = \sum_{i=1}^{N} W_i$ and the number of claims N. We are now in the setting of Example 15.2 with $U_i = W_i$ and $V_i = 1$ for $i = 1, 2, \dots$. This gives of course $Y = N$. We obtain

$$h(u, v) = g(u) \qquad (u = 1, 2, \dots; v = 1) \qquad (15.13)$$

$$f(x, y) = p(y) g^{y*}(x). \qquad (x, y = 0, 1, 2, \dots)$$

Insertion of (15.13) in (15.10) and (15.11) gives that for $x, y = 1, 2, \dots$

$$f(x, y) = \sum_{u=1}^{x} \left(a + b\frac{u}{x}\right) g(u) f(x-u, y-1)$$
$$= \left(a + \frac{b}{y}\right) \sum_{u=1}^{x} g(u) f(x-u, y-1); \qquad (x, y = 1, 2, \dots)$$

the latter recursion is the most convenient of the two.

In the present case, (15.1) gives that $\mathrm{Cov}(X, N) = \mathrm{E}\, W \,\mathrm{Var}\, N > 0$. □

Example 15.4 Let us now assume that the portfolio of Example 15.3 is protected by an unlimited excess of loss reinsurance where the retention is a positive integer r. We want to evaluate the joint distribution of the aggregate payments X of the cedant and Y of the reinsurer. We are now in the setting of Example 15.2 with $U_i = \min(W_i, r)$ and $V_i = (W_i - r)_+$ for $i = 1, 2, \dots$. This gives

$$h(u, v) = \begin{cases} g(u) & (u = 1, 2, \dots, r;\ v = 0) \\ g(r+v). & (u = r;\ v = 1, 2, \dots) \end{cases} \tag{15.14}$$

Insertion in (15.10) and (15.11) gives

$$f(x, y) = \sum_{u=1}^{r} \left(a + b\frac{u}{x}\right) g(u) f(x-u, y) + \left(a + b\frac{r}{x}\right) \sum_{v=1}^{y} g(r+v) f(x-r, y-v)$$
$$(x = 1, 2, \dots;\ y = 0, 1, 2, \dots)$$

$$f(x, y) = a \sum_{u=1}^{r} g(u) f(x-u, y) + \sum_{v=1}^{y} \left(a + b\frac{v}{y}\right) g(r+v) f(x-r, y-v).$$
$$(x = r, r+1, r+2, \dots;\ y = 1, 2, \dots)$$

As the U_is are strictly positive, the former recursion specifies f completely together with the initial value $f(0, 0)$. The latter recursion may be more convenient for small values of y, but then we still need the former recursion to evaluate $f(x, 0)$ for $x = 1, 2, \dots$. Insertion of (15.14) in (15.12) gives

$$f(x, y) = \sum_{u=1}^{r} \left(a + b\frac{u}{x+y}\right) g(u) f(x-u, y) + \sum_{v=1}^{y} \left(a + b\frac{r+v}{x+y}\right) g(r+v) f(x-r, y-v).$$
$$(x = 1, 2, \dots;\ y = 0, 1, 2, \dots)$$

If the reinsurance treaty also has an aggregate limit l where l is a positive integer, then the cedant pays $X + (Y - l)_+$ with distribution k given by

$$k(z) = \sum_{y=0}^{l} f(z, y) + \sum_{y=l+1}^{z+l-r} f(z - y + l, y). \qquad (z = 0, 1, 2, \ldots)$$

□

The following multivariate extension of Theorem 2.3 is proved like in the univariate case.

Theorem 15.1 *If p is the Poisson distribution* $\text{Po}(\lambda)$ *and $h \in \mathcal{P}_{m\mathbf{0}}$ satisfies the relation*

$$y_j h(\mathbf{y}) = \eta(\mathbf{y}) + \sum_{\mathbf{0}<\mathbf{z}\le\mathbf{r}} (y_j - z_j)\chi(\mathbf{z})h(\mathbf{y} - \mathbf{z}) \qquad (\mathbf{y} \in \mathbb{N}_{m+}) \tag{15.15}$$

for functions η and χ on $\{\mathbf{y} \in \mathbb{N}_{m+} : \mathbf{y} \le \mathbf{r}\}$ with $\mathbf{r} = (r_1, r_2, \ldots, r_m)'$ with r_j being a positive integer or infinity for $j = 1, 2, \ldots, m$, then $f = p \vee h$ satisfies the relation

$$x_j f(\mathbf{x}) = \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{r}} (\lambda\eta(\mathbf{y}) + (x_j - y_j)\chi(\mathbf{y}))f(\mathbf{x} - \mathbf{y}). \tag{15.16}$$

$$(\mathbf{x} \in \mathbb{N}_{m+};\ j = 1, 2, \ldots, m)$$

By proceeding from (15.16) in the way we deduced (15.4) and (15.5), we obtain the recursions

$$f(\mathbf{x}) = \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{r}} \left(\frac{\lambda}{x_j}\eta(\mathbf{y}) + \left(1 - \frac{y_j}{x_j}\right)\chi(\mathbf{y})\right) f(\mathbf{x} - \mathbf{y}) \qquad (\mathbf{x} \ge \mathbf{e}_j;\ j = 1, 2, \ldots, m)$$

$$f(\mathbf{x}) = \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{r}} \left(\frac{m}{x_\bullet}\lambda\eta(\mathbf{y}) + \left(1 - \frac{y_\bullet}{x_\bullet}\right)\chi(\mathbf{y})\right) f(\mathbf{x} - \mathbf{y}); \quad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{15.17}$$

for the latter recursion, we can replace the condition (15.15) with the weaker

$$h(\mathbf{y}) = \frac{m}{y_\bullet}\eta(\mathbf{y}) + \sum_{\mathbf{0}<\mathbf{z}\le\mathbf{r}} \left(1 - \frac{z_\bullet}{y_\bullet}\right)\chi(\mathbf{z})h(\mathbf{y} - \mathbf{z}). \qquad (\mathbf{y} \in \mathbb{N}_{m+}) \tag{15.18}$$

15.3 Convolutions of a Distribution

In this section, we shall extend the theory of Sect. 2.4 to the multivariate case. In that section, we gave a direct proof of Theorem 2.8 and afterwards indicated how it could also be deduced from the binomial case of Theorem 2.4. For the multivariate case, we go straight to the binomial case of (15.4) and (15.5). The advantage is that then we can get straight to both of these recursions at the same time. With $h \in \mathcal{P}_{m+}$

and p being bin(M, π), insertion of the values of a and b from Table 2.1 in (15.4) and (15.5) gives

$$f(\mathbf{x}) = \frac{\pi}{1-\pi} \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} \left((M+1)\frac{y_j}{x_j} - 1\right) h(\mathbf{y}) f(\mathbf{x}-\mathbf{y}) \tag{15.19}$$

$$(\mathbf{x} \geq \mathbf{e}_j;\ j = 1, 2, \ldots, m)$$

$$f(\mathbf{x}) = \frac{\pi}{1-\pi} \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} \left((M+1)\frac{y_\bullet}{x_\bullet} - 1\right) h(\mathbf{y}) f(\mathbf{x}-\mathbf{y}). \quad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{15.20}$$

We are now ready to prove the following multivariate version of Theorem 2.8.

Theorem 15.2 *The M-fold convolution $f = g^{M*}$ of $g \in \mathcal{P}_{m\mathbf{0}}$ satisfies the recursions*

$$f(\mathbf{x}) = \frac{1}{g(\mathbf{0})} \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} \left((M+1)\frac{y_j}{x_j} - 1\right) g(\mathbf{y}) f(\mathbf{x}-\mathbf{y}) \tag{15.21}$$

$$(\mathbf{x} \geq \mathbf{e}_j;\ j = 1, 2, \ldots, m)$$

$$f(\mathbf{x}) = \frac{1}{g(\mathbf{0})} \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} \left((M+1)\frac{y_\bullet}{x_\bullet} - 1\right) g(\mathbf{y}) f(\mathbf{x}-\mathbf{y}) \quad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{15.22}$$

with initial value $f(\mathbf{0}) = g(\mathbf{0})^M$.

Proof We represent g as a compound Bernoulli distribution with Bernoulli counting distribution Bern(π) with $\pi = 1 - g(\mathbf{0})$ and severity distribution $h \in \mathcal{P}_{m+}$ given by $h(\mathbf{y}) = g(\mathbf{y})/\pi$ for $\mathbf{y} \in \mathbb{N}_{m+}$. Then $f = p \vee h$ with p being the binomial distribution bin(M, π), and the theorem follows from (15.19) and (15.20). □

The following multivariate extensions of Theorems 2.9 and 2.10 are proved like in the univariate case.

Theorem 15.3 *If the distribution $g \in \mathcal{P}_{m\mathbf{0}}$ has a finite range $\{\mathbf{x} \in \mathbb{N}_m : \mathbf{0} \leq \mathbf{x} \leq \mathbf{k}\}$ with $g(\mathbf{k}) > 0$ for some $\mathbf{k} \in \mathbb{N}_m$, then $f = g^{M*}$ satisfies the recursion*

$$f(\mathbf{x}) = \frac{1}{g(\mathbf{k})} \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{k}} \left(\frac{(M+1)y_j}{Mk_j - x_j} - 1\right) g(\mathbf{k}-\mathbf{y}) f(\mathbf{x}+\mathbf{y})$$

$$(\mathbf{0} \leq \mathbf{x} \leq M\mathbf{k} - \mathbf{e}_j;\ j = 1, 2, \ldots, m)$$

$$f(\mathbf{x}) = \frac{1}{g(\mathbf{k})} \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{k}} \left(\frac{(M+1)y_\bullet}{Mk_\bullet - x_\bullet} - 1\right) g(\mathbf{k}-\mathbf{y}) f(\mathbf{x}+\mathbf{y}) \qquad (\mathbf{0} \leq \mathbf{x} \leq M\mathbf{k})$$

with initial value $f(M\mathbf{k}) = g(\mathbf{k})^M$.

Theorem 15.4 *If $g \in \mathcal{P}_{m\underline{\mathbf{l}}}$ for some $\mathbf{l} \in \mathbb{Z}_m$, then $f = g^{M*}$ satisfies the recursions*

$$f(\mathbf{x}) = \frac{1}{g(\mathbf{l})} \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}-M\mathbf{l}} \left(\frac{(M+1)y_j}{x_j - Ml_j} - 1\right) g(\mathbf{l}+\mathbf{y}) f(\mathbf{x}-\mathbf{y})$$
$$(\mathbf{x} \ge \mathbf{e}_j;\ j = 1, 2, \ldots, m)$$

$$f(\mathbf{x}) = \frac{1}{g(\mathbf{l})} \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}-M\mathbf{l}} \left(\frac{(M+1)y_\bullet}{x_\bullet - Ml_\bullet} - 1\right) g(\mathbf{l}+\mathbf{y}) f(\mathbf{x}-\mathbf{y}) \qquad (\mathbf{x} \in \mathbb{N}_{m+})$$

with initial value $f(M\mathbf{l}) = g(\mathbf{l})^M$.

In the following example, we shall extend the univariate recursion (2.54) to the multivariate case.

Example 15.5 Let $f \in \mathcal{P}_{m+}$ be the multinomial distribution mnom$(M, \boldsymbol{\pi})$. Then $f = g^{M*}$ where $g \in \mathcal{P}_{m+}$ is the m-variate Bernoulli distribution given by (14.3). At first glance, it seems that we cannot apply Theorem 15.4 in this case as there exists no $\mathbf{l} \in \mathbb{Z}_m$ such that $f \in \mathcal{P}_{m\underline{\mathbf{l}}}$ although the support of f is bounded from below by $\mathbf{0}$. However, as $x_m = M - \sum_{j=1}^{m-1} x_j$ when $f(\mathbf{x}) > 0$, we can just as well consider $\tilde{f} \in \mathcal{P}_{m-1,\underline{\mathbf{0}}}$ given by

$$\tilde{f}(x_1, x_2, \ldots, x_{m-1}) = f\left(x_1, x_2, \ldots, x_{m-1}, M - \sum_{j=1}^{m-1} x_j\right),$$

which satisfies $\tilde{f} = \tilde{g}^{M*}$ with $\tilde{g} \in \mathcal{P}_{m-1,\underline{\mathbf{0}}}$ given by

$$\tilde{g}(y_1, y_2, \ldots, y_{m-1}) = g\left(y_1, y_2, \ldots, y_{m-1}, 1 - \sum_{j=1}^{m-1} y_j\right).$$

When $x_m < M$, insertion in (15.22) gives

$$\begin{aligned} f(\mathbf{x}) &= \frac{1}{g(\mathbf{e}_m)} \sum_{i=1}^{m-1} \left(\frac{M+1}{\sum_{j=1}^{m-1} x_j} - 1\right) g(\mathbf{e}_i) f(\mathbf{x}-\mathbf{e}_i) \\ &= \frac{1}{\pi_m} \sum_{i=1}^{m-1} \left(\frac{M+1}{M - x_m} - 1\right) \pi_i f(\mathbf{x}-\mathbf{e}_i) = \frac{1}{\pi_m} \frac{x_m+1}{M-x_m} \sum_{i=1}^{m-1} \pi_i f(\mathbf{x}-\mathbf{e}_i). \end{aligned}$$

More generally, we obtain

$$f(\mathbf{x}) = \frac{1}{\pi_j} \frac{x_j+1}{M-x_j} \sum_{i\ne j} \pi_i f(\mathbf{x}-\mathbf{e}_i). \quad (\mathbf{0} < \mathbf{x} \le M\mathbf{e} - \mathbf{e}_j;\ x_\bullet = M;\ j = 1, 2, \ldots, m)$$

In addition, we have $f(M\mathbf{e}_j) = \pi_j^M$ for $j = 1, 2, \ldots, m$. □

15.4 Infinite Divisibility

The main purpose of the present section is to extend Theorem 4.2 and its proof to the multivariate case.

The definitions of infinite divisibility at the beginning of Chap. 4 are immediately applicable for multivariate distributions. Many of the results presented in that chapter for univariate distributions, are easily extended to the multivariate case, in particular Theorem 4.1. The following multivariate extension of its item vii) is proved like in the univariate case.

Theorem 15.5 *An infinitely divisible distribution in $\mathcal{P}_{m\mathbf{0}}$ has a positive probability in* $\mathbf{0}$.

From Table 2.1, we know that the Poisson distribution Po(λ) is $R_1[0, \lambda]$. Insertion in (15.5) gives that the compound Poisson distribution f with Poisson parameter $\lambda > 0$ and severity distribution $h \in \mathcal{P}_{m+}$ satisfies the recursion

$$f(\mathbf{x}) = \frac{\lambda}{x_\bullet} \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} y_\bullet h(\mathbf{y}) f(\mathbf{x}-\mathbf{y}) \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{15.23}$$

with initial condition

$$f(\mathbf{0}) = \mathrm{e}^{-\lambda}. \tag{15.24}$$

The following multivariate version of Theorem 4.2 can be proved by a trivial modification of the proof of the univariate case.

Theorem 15.6 *A non-degenerate distribution in $\mathcal{P}_{m\underline{\mathbf{0}}}$ is infinitely divisible iff it can be expressed as a compound Poisson distribution with severity distribution in $\mathcal{P}_{m+}$.*

By interpreting a distribution concentrated in zero as a Poisson distribution with parameter zero, we can drop the assumption in Theorem 15.6 that the distribution should be non-degenerate.

15.5 Counting Distribution with Recursion of Higher Order

Let us return to the setting of Sect. 15.2 with the weaker assumption on the counting distribution p that it belongs to $\mathcal{P}_{10}$. Proceeding like in the deduction of (5.4), (15.2) easily extends to

$$\mathrm{E}\left[a + \frac{b}{i}\frac{\sum_{l=1}^{i} Y_{lj}}{x_j}\middle|\mathbf{Y}_{\bullet n} = \mathbf{x}\right] = a + \frac{b}{n}.$$
$$(\mathbf{x} \geq \mathbf{e}_j;\ j = 1, 2, \dots, m;\ n = i, i+1, i+2, \dots;\ i = 1, 2, \dots)$$

Analogous to the deduction of (15.4) and (15.5), representing p in the form (5.6) with a and b now being functions, we extend (5.8) to

$$f(\mathbf{x}) = \frac{1}{1-\tau_a(h(\mathbf{0}))} \times \left((q \vee h)(\mathbf{x}) + \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} \left((a \vee h)(\mathbf{y}) + (\Psi b \vee h)(\mathbf{y})\frac{y_j}{x_j}\right) f(\mathbf{x}-\mathbf{y})\right)$$
$$(\mathbf{x} \geq \mathbf{e}_j;\ j = 1, 2, \ldots, m) \tag{15.25}$$

$$f(\mathbf{x}) = \frac{1}{1-\tau_a(h(\mathbf{0}))} \times \left((q \vee h)(\mathbf{x}) + \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} \left((a \vee h)(\mathbf{y}) + (\Psi b \vee h)(\mathbf{y})\frac{y_\bullet}{x_\bullet}\right) f(\mathbf{x}-\mathbf{y})\right).$$
$$(\mathbf{x} \in \mathbb{N}_{m+}) \tag{15.26}$$

When p is $R_k[a,b]$, we obtain

$$f(\mathbf{x}) = \frac{1}{1-\tau_a(h(\mathbf{0}))} \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} \left((a \vee h)(\mathbf{y}) + (\Psi b \vee h)(\mathbf{y})\frac{y_j}{x_j}\right) f(\mathbf{x}-\mathbf{y})$$
$$(\mathbf{x} \geq \mathbf{e}_j;\ j = 1, 2, \ldots, m) \tag{15.27}$$

$$f(\mathbf{x}) = \frac{1}{1-\tau_a(h(\mathbf{0}))} \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} \left((a \vee h)(\mathbf{y}) + (\Psi b \vee h)(\mathbf{y})\frac{y_\bullet}{x_\bullet}\right) f(\mathbf{x}-\mathbf{y}).$$
$$(\mathbf{x} \in \mathbb{N}_{m+}) \tag{15.28}$$

In particular, if $p \in \mathcal{P}_{m\underline{\mathbf{0}}}$ and $h \in \mathcal{P}_{m+}$, we have

$$f(\mathbf{x}) = \frac{1}{x_j} \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} y_j (\Psi\varphi_p \vee h)(\mathbf{y}) f(\mathbf{x}-\mathbf{y}) \quad (\mathbf{x} \geq \mathbf{e}_j;\ j = 1, 2, \ldots, m) \tag{15.29}$$

$$f(\mathbf{x}) = \frac{1}{x_\bullet} \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} y_\bullet (\Psi\varphi_p \vee h)(\mathbf{y}) f(\mathbf{x}-\mathbf{y}). \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{15.30}$$

With

$$c(\mathbf{y}) = \frac{(a \vee h)(\mathbf{y})}{1-\tau_a(h(\mathbf{0}))}; \qquad d(\mathbf{y}) = \frac{\Phi(\Psi b \vee h)(\mathbf{y})}{1-\tau_a(h(\mathbf{0}))}, \qquad (\mathbf{0} < \mathbf{y} \leq \mathbf{x}) \tag{15.31}$$

we can rewrite (15.28) as

$$f(\mathbf{x}) = \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} \left(c(\mathbf{y}) + \frac{d(\mathbf{y})}{x_\bullet}\right) f(\mathbf{x}-\mathbf{y}). \qquad (\mathbf{x} \in \mathbb{N}_{m+})$$

This is a multivariate version of (5.12) with $k=\infty$. We therefore find it natural to introduce the notation $R_{\mathbf{k}}[\alpha,\beta]$ for a distribution $g\in\mathcal{P}_{m\underline{\mathbf{0}}}$ that satisfies the recursion

$$g(\mathbf{x})=\sum_{\mathbf{0}<\mathbf{y}\le\mathbf{k}}\left(\alpha(\mathbf{y})+\frac{\beta(\mathbf{y})}{x_\bullet}\right)g(\mathbf{x}-\mathbf{y})\qquad(\mathbf{x}\in\mathbb{N}_{m+})\tag{15.32}$$

with $\mathbf{k}=(k_1,k_2,\ldots,k_m)'$ with k_j being a positive integer or infinity for $j=1,2,\ldots,m$ (we hope that it will not cause problems for the readers that the dimension m does not appear explicitly in the notation $R_{\mathbf{k}}[\alpha,\beta]$). The recursion (15.27) can now be expressed as

$$f(\mathbf{x})=\sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}}\left(c(\mathbf{y})+\frac{y_j}{y_\bullet}\frac{d(\mathbf{y})}{x_j}\right)f(\mathbf{x}-\mathbf{y}).\qquad(\mathbf{x}\ge\mathbf{e}_j;\ j=1,2,\ldots,m)$$

15.6 Multivariate Bernoulli Severity Distribution

We shall now give a rather general result for compound distributions where the severity distribution is the multivariate Bernoulli distribution given by (14.3).

Theorem 15.7 *If $p\in\mathcal{P}_{m\mathbf{0}}$ satisfies the recursion*

$$p(n)=b(n)+\sum_{i=1}^{k}b_i(n)p(n-i),\qquad(n=1,2,\ldots)\tag{15.33}$$

and h is the multivariate Bernoulli distribution given by (14.3), *then $f=p\vee h$ satisfies the recursion*

$$\begin{aligned}f(\mathbf{x})&=b(x_\bullet)h^{x_\bullet *}(\mathbf{x})+\sum_{i=1}^{k}b_i(x_\bullet)\sum_{\mathbf{0}<\mathbf{y}<\mathbf{x}}f(\mathbf{x}-\mathbf{y})h^{i*}(\mathbf{y})\\&=x_\bullet!b(x_\bullet)\left(\prod_{j=1}^{m}\frac{\pi_j^{x_j}}{x_j!}\right)+\sum_{i=1}^{k}i!b_i(x_\bullet)\sum_{\substack{\mathbf{0}<\mathbf{y}<\mathbf{x}\\ y_\bullet=i}}\left(\prod_{j=1}^{m}\frac{\pi_j^{y_j}}{y_j!}\right)f(\mathbf{x}-\mathbf{y}).\end{aligned}$$

$$(\mathbf{x}\in\mathbb{N}_{m+})\tag{15.34}$$

Proof Insertion of (15.33) in (14.5) gives that for all $\mathbf{x}\in\mathbb{N}_{m+}$, we have

$$f(\mathbf{x})=\left(b(x_\bullet)+\sum_{i=1}^{k}b_i(x_\bullet)p(x_\bullet-i)\right)h^{x_\bullet *}(\mathbf{x}).\tag{15.35}$$

By application of (14.5), we obtain that for $i = 1, 2, \ldots, k$,

$$p(x_\bullet - i)h^{x_\bullet *}(\mathbf{x}) = p(x_\bullet - i) \sum_{\mathbf{0}<\mathbf{y}<\mathbf{x}} h^{(x_\bullet - i)*}(\mathbf{x}-\mathbf{y})h^{i*}(\mathbf{y})$$
$$= \sum_{\mathbf{0}<\mathbf{y}<\mathbf{x}} f(\mathbf{x}-\mathbf{y})h^{i*}(\mathbf{y}),$$

and insertion in (15.35) gives the first equality in (15.34). The second equality follows by insertion of (14.4).

This completes the proof of Theorem 15.7. □

The following recursion can sometimes be more convenient.

Theorem 15.8 *Let $f = p \vee h$ with $p \in \mathcal{P}_{m\mathbf{0}}$ satisfying the recursion (15.33) and h being the multivariate Bernoulli distribution given by (14.3). For $i = 1, 2, \ldots, k$, let ${}_i\mathbf{y} \in \mathbb{N}_{m+}$ with ${}_iy_\bullet = i$ such that ${}_1\mathbf{y} < {}_2\mathbf{y} < \cdots < {}_k\mathbf{y}$. Then*

$$f(\mathbf{x}) = b(x_\bullet)h^{x_\bullet *}(\mathbf{x}) + \sum_{i=1}^{k} b_i(x_\bullet)x_\bullet^{(i)}\left(\prod_{j=1}^{m} \frac{\pi_j^{{}_iy_j}}{x_j^{({}_iy_j)}}\right) f(\mathbf{x}-{}_i\mathbf{y}). \qquad (\mathbf{x} \geq {}_k\mathbf{y}) \tag{15.36}$$

Proof Application of (14.4) and (14.5) gives that for all $\mathbf{x} \in \mathbb{N}_{m+}$ such that $\mathbf{x} \geq {}_k\mathbf{y}$ and $i = 1, 2, \ldots, k$, we have

$$p(x_\bullet - i)h^{x_\bullet *}(\mathbf{x}) = p(x_\bullet - i)h^{(x_\bullet - i)*}(\mathbf{x}-{}_i\mathbf{y})\frac{h^{x_\bullet *}(\mathbf{x})}{h^{(x_\bullet - i)*}(\mathbf{x}-{}_i\mathbf{y})}$$
$$= f(\mathbf{x}-{}_i\mathbf{y})x_\bullet^{(i)}\left(\prod_{j=1}^{m} \frac{\pi_j^{{}_iy_j}}{x_j^{({}_iy_j)}}\right),$$

and insertion in (15.35) gives (15.36). □

When letting ${}_i\mathbf{y} = i\mathbf{e}_j$ for $i = 1, 2, \ldots, k$ and $j \in \{1, 2, \ldots, m\}$, (15.36) reduces to

$$f(\mathbf{x}) = b(x_\bullet)h^{x_\bullet *}(\mathbf{x}) + \sum_{i=1}^{k} b_i(x_\bullet)\frac{x_\bullet^{(i)}}{x_j^{(i)}}\pi_j^i f(\mathbf{x}-i\mathbf{e}_j).$$
$$(\mathbf{x} \geq k\mathbf{e}_j;\ j = 1, 2, \ldots, m)$$

Further Remarks and References

Sections 15.1–15.3 and 15.5 are based on Sundt (1999c). The recursion (15.4) was deduced independently by Ambagaspitiya (1999), Sundt (1999c), and Walhin and

Paris (2000a). Reinsurance applications have been presented by Mata (2000, 2003), Walhin and Paris (2000a, 2001a), Walhin et al. (2001), Walhin (2002b, 2002c), and Witdouck and Walhin (2004).

Section 15.4 is based on Sundt (2000a). Theorem 15.6 follows from Theorem 2 in Horn and Steutel (1978).

Chapter 16
De Pril Transforms

Summary

The main purpose of the present chapter is to extend the definition of the De Pril transform to functions in $\mathcal{F}_{m\underline{\mathbf{0}}}$ and study some properties of the De Pril transform within that context.

In Sect. 16.1, we extend the definition of the De Pril transform and the Dhaene–De Pril transform to distributions in $\mathcal{P}_{m\underline{\mathbf{0}}}$ and give expressions for these transforms of a multivariate compound distribution of Type 1. Such compound distributions are further discussed in Sect. 16.2, but there we restrict the counting distributions to the $\mathcal{R}_k$ classes.

In Sect. 16.3, we show that a distribution in $\mathcal{P}_{m\underline{\mathbf{0}}}$ is infinitely divisible iff its De Pril transform is non-negative.

In Sect. 16.4, we extend the definition of the De Pril transform to functions $\mathcal{F}_{m\underline{\mathbf{0}}}$ and point out that the most important properties are easily extended to that setting. In particular, we point out that the De Pril transform of a convolution of such functions is the sum of the De Pril transforms of those functions.

Individual models is the topic of Sect. 16.7, and we extend parts of the theory of Chap. 7 to the multivariate case.

Whereas in Sects. 16.1–16.4 and 16.7, we extend definitions and results from the univariate case, Sects. 16.5 and 16.6 are devoted to issues that were not present in the univariate case, namely, respectively, independence and dependence within a random vector. In particular, we deduce a recursion for the multivariate Poisson distribution.

16.1 Definitions

Like in the univariate case, for any distribution in $f \in \mathcal{P}_{m\underline{\mathbf{0}}}$, there exists a unique function φ_f on $\mathcal{F}_{m+}$ such that f can be expressed as $R_\infty[0, \varphi_f]$. We define this function as the *De Pril transform* of f. Insertion in (15.32) gives

$$f(\mathbf{x}) = \frac{1}{x_\bullet} \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} \varphi_f(\mathbf{y}) f(\mathbf{x}-\mathbf{y}), \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{16.1}$$

and by solving for $\varphi_f(\mathbf{x})$, we obtain the recursion

$$\varphi_f(\mathbf{x}) = \frac{1}{f(\mathbf{0})} \left(x_\bullet f(\mathbf{x}) - \sum_{\mathbf{0}<\mathbf{y}<\mathbf{x}} \varphi_f(\mathbf{y}) f(\mathbf{x}-\mathbf{y}) \right). \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{16.2}$$

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_16,

Comparison between (16.1) and (15.30), gives that if $f = p \vee h$ with $p \in \mathcal{P}_{1\underline{0}}$ and $h \in \mathcal{P}_{11}$, then

$$\varphi_f = \Phi(\Psi\varphi_p \vee h). \tag{16.3}$$

By insertion of this in (15.29), we obtain the recursion

$$f(\mathbf{x}) = \frac{1}{x_j} \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}} \frac{y_j}{y_\bullet} \varphi_f(\mathbf{y}) f(\mathbf{x}-\mathbf{y}). \qquad (\mathbf{x} \ge \mathbf{e}_j;\ j = 1, 2, \dots, m) \tag{16.4}$$

As any distribution in $\mathcal{P}_{m\underline{\mathbf{0}}}$ can be represented as a compound distribution with Bernoulli counting distribution in $\mathcal{P}_{1\underline{0}}$ and severity distribution in $\mathcal{P}_{m+}$, this recursion holds for any distribution $f \in \mathcal{P}_{m\underline{\mathbf{0}}}$.

The recursion (16.4) gets a more pleasant shape if we use the Dhaene–De Pril transform instead of the De Pril transform. Extending the univariate case discussed in Sect. 10.3.2, we define the *Dhaene–De Pril transform* of a distribution $f \in \mathcal{P}_{m\underline{\mathbf{0}}}$ by

$$\psi_f(\mathbf{x}) = \begin{cases} \ln f(\mathbf{0}) & (\mathbf{x} = \mathbf{0}) \\ \varphi_f(\mathbf{x})/x_\bullet. & (\mathbf{x} \in \mathbb{N}_{m+}) \end{cases}$$

Insertion in (16.1) and (16.4) gives

$$f(\mathbf{x}) = \frac{1}{x_\bullet} \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}} y_\bullet \psi_f(\mathbf{y}) f(\mathbf{x}-\mathbf{y}) \qquad (\mathbf{x} \in \mathbb{N}_{m+})$$

$$f(\mathbf{x}) = \frac{1}{x_j} \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}} y_j \psi_f(\mathbf{y}) f(\mathbf{x}-\mathbf{y}), \qquad (\mathbf{x} \ge \mathbf{e}_j;\ j = 1, 2, \dots, m)$$

and by solving for $\psi_f(\mathbf{x})$, we obtain the recursions

$$\psi_f(\mathbf{x}) = \frac{1}{f(\mathbf{0})} \left(f(\mathbf{x}) - \frac{1}{x_\bullet} \sum_{\mathbf{0}<\mathbf{y}<\mathbf{x}} y_\bullet \psi_f(\mathbf{y}) f(\mathbf{x}-\mathbf{y}) \right) \qquad (\mathbf{x} \in \mathbb{N}_{m+})$$

$$\psi_f(\mathbf{x}) = \frac{1}{f(\mathbf{0})} \left(f(\mathbf{x}) - \frac{1}{x_j} \sum_{\mathbf{0}<\mathbf{y}<\mathbf{x}} y_j \psi_f(\mathbf{y}) f(\mathbf{x}-\mathbf{y}) \right).$$

$$(\mathbf{x} \ge \mathbf{e}_j;\ j = 1, 2, \dots, m)$$

Like in the univariate case, in terms of Dhaene–De Pril transforms, (16.3) reduces to $\psi_f = \psi_p \vee h$. Although the Dhaene–De Pril transform has some advantages compared to the De Pril transform, we shall not pursue it further in this book.

16.2 The $\mathcal{R}_k$ Classes

The following multivariate extension of Theorem 6.2 is proved from (15.18) and (15.17) by a trivial modification of the proof from the univariate case.

Theorem 16.1 *A distribution $f \in \mathcal{P}_{m\underline{0}}$ satisfies the recursion*

$$f(\mathbf{x}) = \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{r}} \left(\frac{\xi(\mathbf{y})}{x_\bullet} + \left(1 - \frac{y_\bullet}{x_\bullet}\right)\chi(\mathbf{y})\right) f(\mathbf{x}-\mathbf{y}) \qquad (\mathbf{x} \in \mathbb{N}_{m+})$$

for functions ξ and χ on $\{\mathbf{y} \in \mathbb{N}_{m+} : \mathbf{y} \leq \mathbf{r}\}$ with $\mathbf{r} = (r_1, r_2, \ldots, r_m)'$ with r_j being a positive integer or infinity for $j = 1, 2, \ldots, m$ iff its De Pril transform satisfies the recursion

$$\varphi_f(\mathbf{x}) = \xi(\mathbf{x}) + \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{r}} \chi(\mathbf{y})\varphi_f(\mathbf{x}-\mathbf{y}). \qquad (\mathbf{x} \in \mathbb{N}_{m+})$$

By letting $\mathbf{r} = \mathbf{k}$, $\chi = c$, and $\xi = \Phi c + d$, we obtain the following multivariate extension of Corollary 6.8.

Corollary 16.1 *If f is $R_{\mathbf{k}}[c, d]$ with $\mathbf{k} = (k_1, k_2, \ldots, k_m)'$ with k_j being a positive integer or infinity for $j = 1, 2, \ldots, m$, then*

$$\varphi_f(\mathbf{x}) = (\Phi c + d)(\mathbf{x}) + \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{k}} c(\mathbf{y})\varphi_f(\mathbf{x}-\mathbf{y}). \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{16.5}$$

By letting $k_j = \infty$ for $j = 1, 2, \ldots, m$, $c \equiv 0$, and $d = \varphi_f$ in (16.5), we get the obvious result $\varphi_f(\mathbf{x}) = \varphi_f(\mathbf{x})$.

Insertion of (15.31) in (16.5) with $k_j = \infty$ for $j = 1, 2, \ldots, m$ gives the following multivariate extension of Corollary 6.9.

Corollary 16.2 *If $f = p \vee h$ where p is $R_k[a, b]$ and $h \in \mathcal{P}_{m\mathbf{0}}$, then*

$$\varphi_f(\mathbf{x}) = \frac{1}{1 - \tau_a(h(\mathbf{0}))}\left(x_\bullet((a + \Psi b) \vee h)(\mathbf{x}) + \sum_{\mathbf{0}<\mathbf{y}<\mathbf{x}} (a \vee h)(\mathbf{y})\varphi_f(\mathbf{x}-\mathbf{y})\right).$$
$$(\mathbf{x} \in \mathbb{N}_{m+}) \tag{16.6}$$

Letting $k = \infty$, $a \equiv 0$, and $b = \varphi_p$, brings us back to (16.3).

Now let $k = 1$. In this case, (16.6) reduces to

$$\varphi_f(\mathbf{x}) = \frac{1}{1 - ah(\mathbf{0})}\left(x_\bullet(a + b)h(\mathbf{x}) + a \sum_{\mathbf{0}<\mathbf{y}<\mathbf{x}} h(\mathbf{y})\varphi_f(\mathbf{x}-\mathbf{y})\right).$$
$$(\mathbf{x} \in \mathbb{N}_{m+})$$

Insertion of (6.17) in (16.3) extends (6.18) to

$$\varphi_f(\mathbf{x}) = x_\bullet(a + b)\sum_{n=1}^{x_\bullet} \frac{a^{n-1}}{n} h^{n*}(\mathbf{x}). \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{16.7}$$

16.3 Infinite Divisibility

The following theorem extends Theorem 4.5 to multivariate distributions.

Theorem 16.2 *A distribution in $\mathcal{P}_{m\underline{0}}$ is infinitely divisible iff its De Pril transform is non-negative.*

Analogous to the univariate case, we shall need the following lemma to prove this theorem.

Lemma 16.1 *If a distribution $f \in \mathcal{P}_{m\underline{0}}$ has non-negative De Pril transform, then $\sum_{\mathbf{x}\in\mathbb{N}_{m+}} \varphi_f(\mathbf{x})/x_\bullet < \infty$.*

Proof Application of (16.2) gives

$$\sum_{\mathbf{x}\in\mathbb{N}_{m+}} \frac{\varphi_f(\mathbf{x})}{x_\bullet} = \sum_{\mathbf{x}\in\mathbb{N}_{m+}} \frac{1}{f(\mathbf{0})}\left(f(\mathbf{x}) - \frac{1}{x_\bullet}\sum_{\mathbf{0}<\mathbf{y}<\mathbf{x}} \varphi_f(\mathbf{y})f(\mathbf{x}-\mathbf{y})\right)$$
$$\leq \sum_{\mathbf{x}\in\mathbb{N}_{m+}} \frac{f(\mathbf{x})}{f(\mathbf{0})} = \frac{1-f(\mathbf{0})}{f(\mathbf{0})} < \infty.$$

□

Proof of Theorem 16.2 From Theorem 15.6, we know that a distribution in $\mathcal{P}_{m\underline{0}}$ is infinitely divisible iff it can be expressed as a compound Poisson distribution with severity distribution in $\mathcal{P}_{m+}$. Hence, it is sufficient to show that a distribution in $\mathcal{P}_{m\underline{0}}$ has non-negative De Pril transform iff it can be expressed as a compound Poisson distribution with severity distribution in $\mathcal{P}_{m+}$.

Let f be a compound Poisson distribution with Poisson parameter λ and severity distribution $h \in \mathcal{P}_{m+}$. From Table 2.1, we know that the counting distribution is $R_1[0, \lambda]$, and insertion in (16.7) gives

$$\varphi_f(\mathbf{x}) = \lambda x_\bullet h(\mathbf{x}) \geq 0, \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{16.8}$$

that is, a compound Poisson distribution with severity distribution in $\mathcal{P}_{m+}$ has non-negative De Pril transform.

Now, let f be a distribution with non-negative De Pril transform. From Lemma 16.1, we have that

$$\lambda = \sum_{\mathbf{x}\in\mathbb{N}_{m+}} \frac{\varphi_f(\mathbf{x})}{x_\bullet} < \infty.$$

Let $h(\mathbf{x}) = \varphi_f(\mathbf{x})/(\lambda x_\bullet)$ for $\mathbf{x} \in \mathbb{N}_{m+}$. Then h is non-negative and sums to one and is therefore a distribution in $\mathcal{P}_{m+}$. Insertion of $\varphi_f(\mathbf{x}) = \lambda x_\bullet h(\mathbf{x})$ in (16.1) gives the recursion (15.23) for $p \vee h$ where p denotes the Poisson distribution with parameter λ. Hence, f and $p \vee h$ must be proportional, and, as both of them are distributions and therefore sum to one, they must be equal. Thus, any distributions in $\mathcal{P}_{m\underline{\mathbf{0}}}$ with non-negative De Pril transform can be expressed as a compound Poisson distribution with severity distribution in $\mathcal{P}_{m+}$.

This completes the proof of Theorem 16.2. □

16.4 Extension to Functions in $\mathcal{F}_{m\underline{\mathbf{0}}}$

In Chap. 6, we studied some properties of the De Pril transform of distributions in $\mathcal{P}_{1\underline{0}}$. Later we found that it would be convenient to apply De Pril transforms more generally for functions in $\mathcal{F}_{1\underline{0}}$ in connection with cumulative functions and tails (Chap. 8) as well as approximations to distributions (Chap. 10). In Chap. 10, we discussed extension of properties of De Pril transforms to functions in $\mathcal{F}_{1\underline{0}}$. Analogously, in the multivariate setting, it can be convenient to extend the definition of the De Pril transform to functions in $\mathcal{F}_{m\underline{\mathbf{0}}}$, and we therefore use (16.2) as definition of the De Pril transform of any function $f \in \mathcal{F}_{m\underline{\mathbf{0}}}$.

Like in the univariate case, a distribution in $\mathcal{P}_{m\underline{\mathbf{0}}}$ is uniquely determined by its De Pril transform as a distribution sums to one, whereas the De Pril transform of a function in $\mathcal{F}_{m\underline{\mathbf{0}}}$ determines that function only up to a multiplicative constant.

The proofs of Lemma 10.1 and Theorems 10.1–10.3 easily extend to the multivariate case. Hence, the De Pril transform of the convolution of a finite number of functions in $\mathcal{F}_{m\underline{\mathbf{0}}}$ is still the sum of the De Pril transforms of these functions, and (16.3) still holds when $p \in \mathcal{F}_{1\underline{0}}$ and $h \in \mathcal{F}_{m+}$.

16.5 Vectors of Independent Random Subvectors

Let $\mathbf{X}$ be a random $m \times 1$ vector with distribution $f \in \mathcal{P}_{m\underline{\mathbf{0}}}$. We express $\mathbf{X}$ as $\mathbf{X} = (\mathbf{X}^{(1)\prime}, \mathbf{X}^{(2)\prime}, \dots, \mathbf{X}^{(s)\prime})'$, where $\mathbf{X}^{(j)}$ is an $m_j \times 1$ subvector for $j = 1, 2, \dots, s$ with $\sum_{j=1}^{s} m_j = m$. With analogous notation, we split other $m \times 1$ vectors in the same way. The main purpose of this section is to show how to express the De Pril transform of f in terms of the De Pril transforms of the marginal distributions of the subvectors $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(s)}$ when these subvectors are independent. Furthermore, we shall characterise such independence in terms of De Pril transforms.

We introduce the sets

$$C_j = \{\mathbf{x} \in \mathbb{N}_m : \mathbf{x}^{(i)} = \mathbf{0}; i \neq j\}; \qquad C_{j+} = C_j \sim \{\mathbf{0}\}; \qquad (j = 1, 2, \dots, s)$$

the C_{j+}s are disjoint. For functions $f \in \mathcal{F}_{m\underline{\mathbf{0}}}$ and $j = 1, 2, \dots, s$, we define the function $\hat{f}_j \in \mathcal{F}_{m_j\underline{\mathbf{0}}}$ by

$$\hat{f}_j(\mathbf{x}^{(j)}) = f(\mathbf{x}). \qquad (\mathbf{x} \in C_j) \tag{16.9}$$

For the following, it will be convenient to rewrite (16.2) as

$$\varphi_f(\mathbf{x}) = \frac{1}{f(\mathbf{0})}\left(x_\bullet f(\mathbf{x}) - \sum_{\mathbf{0}<\mathbf{y}<\mathbf{x}} \varphi_f(\mathbf{x}-\mathbf{y}) f(\mathbf{y})\right). \quad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{16.10}$$

Lemma 16.2 *For $f \in \mathcal{F}_{m\underline{\mathbf{0}}}$, we have*

$$\varphi_f(\mathbf{x}) = \varphi_{\hat{f}_j}(\mathbf{x}^{(j)}). \qquad (\mathbf{x} \in C_{j+};\ j = 1, 2, \ldots, s)$$

Proof From (16.10) and (16.9), we obtain that

$$\varphi_f(\mathbf{x}) = \frac{1}{\hat{f}_j(\mathbf{0})}\left(x_\bullet^{(j)} \hat{f}_j(\mathbf{x}^{(j)}) - \sum_{\mathbf{0}<\mathbf{y}<\mathbf{x}} \varphi_f(\mathbf{x}-\mathbf{y}) \hat{f}_j(\mathbf{y}^{(j)})\right).$$
$$(\mathbf{x} \in C_{j+};\ j = 1, 2, \ldots, s)$$

This is the same recursion as the recursion (16.10) for $\varphi_{\hat{f}_j}$, and, thus, Lemma 16.2 is proved. □

The following theorem follows immediately from Lemma 16.2.

Theorem 16.3 *If $f, g \in \mathcal{F}_{m\underline{\mathbf{0}}}$ with $f(\mathbf{x}) = cg(\mathbf{x})$ for all $\mathbf{x} \in \mathbf{C}_{j+}$ for some j and some positive constant c, then $\varphi_f(\mathbf{x}) = \varphi_g(\mathbf{x})$ for all $\mathbf{x} \in \mathbf{C}_{j+}$.*

Theorem 16.4 *Let $j \in \{1, 2, \ldots, s\}$. Then the function $f \in \mathcal{F}_{m\underline{\mathbf{0}}}$ satisfies*

$$f(\mathbf{x}) = 0 \qquad (\mathbf{x} \in \mathbb{N}_m \sim C_j) \tag{16.11}$$

iff

$$\varphi_f(\mathbf{x}) = 0. \qquad (\mathbf{x} \in \mathbb{N}_{m+} \sim C_{j+}) \tag{16.12}$$

Proof We first assume that (16.11) holds. Insertion of (16.11) in (16.10) gives

$$\varphi_f(\mathbf{x}) = -\frac{1}{f(\mathbf{0})} \sum_{\{\mathbf{y}\in C_{j+}:\mathbf{0}<\mathbf{y}<\mathbf{x}\}} \varphi_f(\mathbf{x}-\mathbf{y}) f(\mathbf{y}). \qquad (\mathbf{x} \in \mathbb{N}_{m+} \sim C_{j+})$$

This gives a recursion for $\varphi_f(\mathbf{x})$ for $\mathbf{x} \in \mathbb{N}_{m+} \sim C_{j+}$ as $\mathbf{x} - \mathbf{y} \in \mathbb{N}_{m+} \sim C_{j+}$ when $\mathbf{x} \in \mathbb{N}_{m+} \sim C_{j+}$ and $\mathbf{y} \in C_{j+}$. In particular, if $\mathbf{x}^{(j)}$ is a unit vector, we obtain that $\varphi_f(\mathbf{x}) = 0$. By induction follows that $\varphi_f(\mathbf{x}) = 0$ also for all other $\mathbf{x} \in \mathbb{N}_{m+} \sim C_{j+}$.

We now assume that (16.12) holds. Let

$$f_j(\mathbf{x}) = f(\mathbf{x}) I(\mathbf{x} \in C_j). \qquad (\mathbf{x} \in \mathbb{N}_m) \tag{16.13}$$

Then Theorem 16.3 gives that $\varphi_{f_j}(\mathbf{x}) = \varphi_f(\mathbf{x})$ for all $\mathbf{x} \in C_{j+}$. As f_j satisfies (16.11), it must also satisfy (16.12) so that for all $\mathbf{x} \in \mathbb{N}_{m+} \sim C_{j+}$, $\varphi_{f_j}(\mathbf{x}) = 0 = \varphi_f(\mathbf{x})$. Hence, $\varphi_f = \varphi_{f_j}$ so that f and f_j must be proportional; as also $f(\mathbf{0}) = f_j(\mathbf{0})$, they are even equal. Hence, f satisfies (16.11).

This completes the proof of Theorem 16.4. □

Theorem 16.5 *If $f \in \mathcal{F}_{m\underline{\mathbf{0}}}$ can be written in the form*

$$f(\mathbf{x}) = \prod_{j=1}^{s} f_j(\mathbf{x}^{(j)}) \qquad (\mathbf{x} \in \mathbb{N}_m) \tag{16.14}$$

with $f_j \in \mathcal{F}_{m_j\underline{\mathbf{0}}}$ for $j = 1, 2, \ldots, s$, then

$$\varphi_f(\mathbf{x}) = \begin{cases} \varphi_{f_j}(\mathbf{x}^{(j)}) & (\mathbf{x} \in C_{j+};\ j = 1, 2, \ldots, s) \\ 0. & (\textit{otherwise}) \end{cases} \tag{16.15}$$

Proof We have $f = *_{j=1}^{s} \dot{f}_j$ with

$$\dot{f}_j(\mathbf{x}) = f_j(\mathbf{x}^{(j)}) I(\mathbf{x} \in C_j). \qquad (\mathbf{x} \in \mathbb{N}_m;\ j = 1, 2, \ldots, s)$$

Application of Theorems 16.3 and 16.4 gives that

$$\varphi_{\dot{f}_j}(\mathbf{x}) = \varphi_{f_j}(\mathbf{x}^{(j)}) I(\mathbf{x} \in C_{j+}), \qquad (\mathbf{x} \in \mathbb{N}_{m+};\ j = 1, 2, \ldots, s)$$

and (16.15) follows as $\varphi_f = \sum_{j=1}^{s} \varphi_{\dot{f}_j}$. □

In the setting at the beginning of this section, the condition (16.14) is satisfied iff the $\mathbf{X}^{(j)}$s are independent and $\mathbf{X}^{(j)}$ has marginal distribution f_j for $j = 1, 2, \ldots, s$.

Corollary 16.3 *A function $f \in \mathcal{F}_{m\underline{\mathbf{0}}}$ can be expressed in the form (16.14) iff*

$$\varphi_f(\mathbf{x}) = 0. \qquad \left(\mathbf{x} \in \mathbb{N}_{m+} \sim \bigcup_{j=1}^{s} C_{j+}\right) \tag{16.16}$$

Proof Theorem 16.5 gives that if $f \in \mathcal{F}_{m\underline{\mathbf{0}}}$ can be expressed in the form (16.14), then (16.16) holds.

We now assume that $f \in \mathcal{F}_{m\underline{\mathbf{0}}}$ satisfies (16.16) and define $\tilde{f} \in \mathcal{F}_{m\underline{\mathbf{0}}}$ by

$$\tilde{f}(\mathbf{x}) = \prod_{j=1}^{s} f_j(\mathbf{x}^{(j)}) \qquad (\mathbf{x} \in \mathbb{N}_m)$$

with $f_j \in \mathcal{F}_{m_j\underline{\mathbf{0}}}$ given by $f_j = f(\mathbf{0})^{1/s-1} \hat{f}_j$ for $j = 1, 2, \ldots, s$. Then $\tilde{f}(\mathbf{x}) = f(\mathbf{x})$ when $\mathbf{x} \in \bigcup_{j=1}^{s} C_{j+}$, and, thus, Theorem 16.3 gives that $\varphi_f(\mathbf{x}) = \varphi_{\tilde{f}}(\mathbf{x})$ when $\mathbf{x} \in \bigcup_{j=1}^{s} C_{j+}$. Furthermore, from Theorem 16.5 and (16.16) follows that $\varphi_{\tilde{f}}(\mathbf{x}) = 0 = \varphi_f(\mathbf{x})$ when $\mathbf{x} \in \mathbb{N}_{m+} \sim \bigcup_{j=1}^{s} C_{j+}$ so that $\varphi_{\tilde{f}} = \varphi_f$. Hence, f and $\tilde{f}$ are proportional, and, as $f(\mathbf{0}) = \tilde{f}(\mathbf{0})$, we have $f = \tilde{f}$, that is, f can be expressed in the form (16.14).

This completes the proof of Corollary 16.3. □

In particular, this corollary gives a necessary and sufficient condition for independence of subvectors of a random vector in terms of the De Pril transform of the distribution of this vector.

Let us now reformulate Theorem 16.5 in the special case when $s = m$ and $m_j = 1$ for $j = 1, 2, \dots, s$. In the setting of the beginning of this section, this corresponds to the situation that $\mathbf{X}$ is a vector of m independent random variables with distributions $f_1, f_2, \dots, f_m$.

Corollary 16.4 *If $f \in \mathcal{F}_{m\underline{0}}$ can be written in the form*

$$f(\mathbf{x}) = \prod_{j=1}^{m} f_j(x_j) \qquad (\mathbf{x} \in \mathbb{N}_m) \tag{16.17}$$

with $f_j \in \mathcal{F}_{1\underline{0}}$ for $j = 1, 2, \dots, m$, then

$$\varphi_f(\mathbf{x}) = \begin{cases} \varphi_{f_j}(x) & (\mathbf{x} = \mathbf{e}_j x;\ x = 1, 2, \dots;\ j = 1, 2, \dots, m) \\ 0. & (otherwise) \end{cases} \tag{16.18}$$

By application of (16.17), (16.4), and (16.18), we obtain that for all $\mathbf{x} \geq \mathbf{e}_j$ and $j = 1, 2, \dots, m$

$$\begin{aligned} \prod_{i=1}^{m} f_i(x_i) = f(\mathbf{x}) &= \frac{1}{x_j} \sum_{y=1}^{x_j} \varphi_{f_j}(y) f(\mathbf{x} - \mathbf{e}_j y) \\ &= \frac{1}{x_j} \sum_{y=1}^{x_j} \varphi_{f_j}(y) f_j(x_j - y) \prod_{i \neq j}^{m} f_i(x_i). \end{aligned}$$

When $\prod_{i \neq j}^{m} f_i(x_i) \neq 0$, which is the case at least when $x_i = 0$ for all $i \neq j$, division by $\prod_{i \neq j}^{m} f_i(x_i)$ gives

$$f_j(x_j) = \frac{1}{x_j} \sum_{y=1}^{x_j} \varphi_{f_j}(y) f_j(x_j - y),$$

that is, we are back to the univariate recursion (6.2) for f_j.

Let us finally apply Corollary 16.4 to extend (8.13) to the multivariate case.

Theorem 16.6 *For $f \in \mathcal{F}_{m\underline{0}}$ and $t = 0, 1, 2, \dots$, we have*

$$\varphi_{\Gamma^t f}(\mathbf{x}) = \begin{cases} \varphi_f(\mathbf{x}) + t & (\mathbf{x} = \mathbf{e}_j x;\ x = 1, 2, \dots;\ j = 1, 2, \dots, m) \\ \varphi_f(\mathbf{x}). & (otherwise) \end{cases} \tag{16.19}$$

Proof We easily see that $\Gamma f = f * \gamma_m$ with $\gamma_m \in \mathcal{F}_{m\underline{\mathbf{0}}}$ defined by

$$\gamma_m(\mathbf{x}) = \prod_{j=1}^{m} \gamma(x_j) \qquad (\mathbf{x} \in \mathbb{N}_m)$$

with γ given by (1.37). Hence, for $t = 0, 1, 2, \dots, \Gamma^t f = f * \gamma_m^{t*}$ and

$$\varphi_{\Gamma^t f} = \varphi_f + t\varphi_{\gamma_m}. \tag{16.20}$$

By insertion of (8.14) in (16.18), we obtain

$$\varphi_{\gamma_m}(\mathbf{x}) = \begin{cases} 1 & (\mathbf{x} = x\mathbf{e}_j; x = 1, 2, \dots; j = 1, 2, \dots, m) \\ 0, & (\text{otherwise}) \end{cases}$$

and insertion in (16.20) gives (16.19). □

16.6 Vectors of Linear Combinations of Independent Random Variables

Like in the previous section, let $\mathbf{X}$ be a random $m \times 1$ vector with distribution $f \in \mathcal{P}_{m\underline{\mathbf{0}}}$. Whereas in the previous section we studied the effect on the De Pril transform of f of stochastic independence within the vector $\mathbf{X}$, we shall now look at the effect of modelling linear dependence within $\mathbf{X}$. Our framework will be that $\mathbf{X} = \mathbf{AY}$ where $\mathbf{Y}$ is an $s \times 1$ vector of independent random variables $Y_1, Y_2, \dots, Y_s$ with distributions $g_1, g_2, \dots, g_s \in \mathcal{P}_{1\underline{0}}$ and $\mathbf{A}$ is a non-random $m \times s$ matrix of non-negative integers.

We start with the special case $s = 1$.

Theorem 16.7 *Let $f \in \mathcal{F}_{m\underline{\mathbf{0}}}$ be defined by*

$$f(\mathbf{x}) = g(y), \qquad (\mathbf{x} = \mathbf{a}y; y = 1, 2, \dots) \tag{16.21}$$

where $g \in \mathcal{F}_{1\underline{0}}$ and $\mathbf{a}$ is a non-random vector of non-negative integers. Then

$$\varphi_f(\mathbf{x}) = \begin{cases} a_\bullet \varphi_g(y) & (\mathbf{x} = \mathbf{a}y; y = 1, 2, \dots) \\ 0. & (\textit{otherwise}) \end{cases} \tag{16.22}$$

Proof We express f as a compound function $f = g \vee h$ with $h \in \mathcal{P}_{m+}$ being the degenerate distribution concentrated in $\mathbf{a}$. Then, for $y = 0, 1, 2, \dots, h^{y*}$ is the distribution concentrated in $\mathbf{a}y$, and insertion in (16.3) gives (16.22). □

The special case with $f \in \mathcal{P}_{1\underline{0}}$ was treated in Corollary 6.5.

The following corollary follows easily from Theorem 16.7 and characterises functions that can be expressed in the form (16.21), in terms of their De Pril transforms.

Corollary 16.5 *If* $f \in \mathcal{F}_{m\underline{0}}$ *satisfies* $\varphi_f(\mathbf{x}) = 0$ *except when* $\mathbf{x} = \mathbf{a}y$ *for* $y = 1, 2, \dots$ *and some non-zero* $m \times 1$ *vector* $\mathbf{a}$ *of non-negative elements, then* f *can be written in the form* (16.21) *with* $g \in \mathcal{F}_{1\underline{0}}$ *given by* $g(0) = f(\mathbf{0})$ *and*

$$\varphi_g(y) = \frac{\varphi_f(\mathbf{a}y)}{a_\bullet}. \qquad (y = 1, 2, \dots)$$

Let us now turn to general s. In the setting of the beginning of this section, we have $\mathbf{X} = \mathbf{Z}_{\bullet s}$ with $\mathbf{Z}_i = \mathbf{a}_i Y_i$ for $i = 1, 2, \dots, s$. For $i = 1, 2, \dots, s$, let f_i denote the distribution of $\mathbf{Z}_i$. Then Theorem 16.7 gives that

$$\varphi_{f_i}(\mathbf{x}) = \begin{cases} a_{\bullet i}\varphi_{g_i}(y) & (\mathbf{x} = \mathbf{a}_i y;\ y = 1, 2, \dots) \\ 0. & (\text{otherwise}) \end{cases} \qquad (i = 1, 2, \dots, s) \tag{16.23}$$

As the Y_is are mutually independent, this is also the case with the $\mathbf{Z}_i$s so that

$$\varphi_f = \sum_{i=1}^{s} \varphi_{f_i}. \tag{16.24}$$

By insertion of this and (16.23) in (16.1) and (16.4), we obtain

$$f(\mathbf{x}) = \frac{1}{x_\bullet}\sum_{i=1}^{s} a_{\bullet i} \sum_{\{y \in \mathbb{N}_{1+}:\mathbf{a}_i y \le \mathbf{x}\}} \varphi_{g_i}(y) f(\mathbf{x} - \mathbf{a}_i y) \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{16.25}$$

$$f(\mathbf{x}) = \frac{1}{x_j}\sum_{i=1}^{s} a_{ji} \sum_{\{y \in \mathbb{N}_{1+}:\mathbf{a}_i y \le \mathbf{x}\}} \varphi_{g_i}(y) f(\mathbf{x} - \mathbf{a}_i y). \qquad (\mathbf{x} \ge \mathbf{e}_j;\ j = 1, 2, \dots, m) \tag{16.26}$$

If none of the columns of $\mathbf{A}$ are proportional to each other, insertion of (16.23) in (16.24) gives

$$\varphi_f(\mathbf{x}) = \begin{cases} a_{\bullet i}\varphi_{g_i}(y) & (\mathbf{x} = \mathbf{a}_i y;\ y = 1, 2, \dots;\ i = 1, 2, \dots, s) \\ 0. & (\text{otherwise}) \end{cases} \tag{16.27}$$

These procedures also apply when we approximate each g_i with a function in $\mathcal{F}_{1\underline{0}}$. When $s = m$ and $\mathbf{A} = \mathbf{I}$, (16.27) then reduces to (16.18).

Example 16.1 For $i = 1, 2, \dots, s$, let g_i be a compound Poisson distribution with Poisson parameter λ_i and severity distribution $h_i \in \mathcal{P}_{11}$. Then (16.8) gives that $\varphi_{g_i} = \lambda_i \Phi h_i$. Insertion in (16.25) and (16.26) gives

$$f(\mathbf{x}) = \frac{1}{x_\bullet}\sum_{i=1}^{s} a_{\bullet i}\lambda_i \sum_{\{y \in \mathbb{N}_{1+}:\mathbf{a}_i y \le \mathbf{x}\}} y h_i(y) f(\mathbf{x} - \mathbf{a}_i y)\ (\mathbf{x} \in \mathbb{N}_{m+}) \tag{16.28}$$

$$f(\mathbf{x}) = \frac{1}{x_j} \sum_{i=1}^{s} a_{ji} \lambda_i \sum_{\{y \in \mathbb{N}_{1+} : \mathbf{a}_i y \leq \mathbf{x}\}} y h_i(y) f(\mathbf{x} - \mathbf{a}_i y). \tag{16.29}$$

$$(\mathbf{x} \geq \mathbf{e}_j;\ j = 1, 2, \dots, m)$$

By comparing (16.28) with (16.1), we see that $\varphi_f = \lambda_\bullet \Phi h$ with $\lambda_\bullet = \sum_{i=1}^{s} \lambda_i$ and $h \in \mathcal{P}_{m+}$ given by

$$h(\mathbf{x}) = \begin{cases} \frac{\lambda_i}{\lambda_\bullet} h_i(y) & (\mathbf{x} = \mathbf{a}_i y;\ y = 1, 2, \dots;\ i = 1, 2, \dots, s) \\ 0. & (\text{otherwise}) \end{cases} \tag{16.30}$$

Comparison with (16.8) shows that f is a compound Poisson distribution with Poisson parameter $\lambda_\bullet$ and severity distribution h. From (16.27), we obtain that

$$\varphi_f(\mathbf{x}) = \begin{cases} a_{\bullet i} \lambda_i y h_i(y) & (\mathbf{x} = \mathbf{a}_i y;\ y = 1, 2, \dots;\ i = 1, 2, \dots, s) \\ 0 & (\text{otherwise}) \end{cases} \tag{16.31}$$

when none of the columns of $\mathbf{A}$ are proportional to each other.

In the special case when each g_i is the Poisson distribution $\text{Po}(\lambda_i)$, that is, h_i is the degenerate distribution concentrated in one, (16.28)–(16.31) reduce to

$$f(\mathbf{x}) = \frac{1}{x_\bullet} \sum_{i=1}^{s} a_{\bullet i} \lambda_i f(\mathbf{x} - \mathbf{a}_i) \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{16.32}$$

$$f(\mathbf{x}) = \frac{1}{x_j} \sum_{i=1}^{s} a_{ji} \lambda_i f(\mathbf{x} - \mathbf{a}_i) \qquad (\mathbf{x} \geq \mathbf{e}_j;\ j = 1, 2, \dots, m) \tag{16.33}$$

$$h(\mathbf{x}) = \begin{cases} \lambda_i / \lambda_\bullet & (\mathbf{x} = \mathbf{a}_i;\ i = 1, 2, \dots, s) \\ 0 & (\text{otherwise}) \end{cases} \tag{16.34}$$

$$\varphi_f(\mathbf{x}) = \begin{cases} a_{\bullet i} \lambda_i & (\mathbf{x} = \mathbf{a}_i;\ i = 1, 2, \dots, s) \\ 0 & (\text{otherwise}) \end{cases}$$

respectively. When all the elements in $\mathbf{A}$ are equal to zero or one, the distribution f is the *multivariate Poisson distribution* $\text{mPo}(\mathbf{A}, \boldsymbol{\lambda})$. In that case, (16.33) reduces to

$$f(\mathbf{x}) = \frac{1}{x_j} \sum_{\{i : a_{ji} = 1\}} \lambda_i f(\mathbf{x} - \mathbf{a}_i). \qquad (\mathbf{x} \geq \mathbf{e}_j;\ j = 1, 2, \dots, m) \tag{16.35}$$

The reason that we restrict the elements of $\mathbf{A}$ to zero and one, is that in that case, Theorem 4.4 gives that the marginal distributions of $\mathbf{X}$ are Poisson distributed. We shall discuss the multivariate Poisson distribution more thoroughly in Sect. 20.1. □

Example 16.2 A common way to construct dependent random variables from independent ones is to take a set of independent random variables and add to each of them the same random variable which is independent of the other variables. Let

$Y_0, Y_1, Y_2, \dots, Y_m$ be independent random variables and assume that for each i, Y_i has distribution $g_i \in \mathcal{P}_{1\underline{0}}$. For $i = 1, 2, \dots, m$, we let $X_i = Y_0 + Y_i$. Then

$$\mathbf{X} = (X_1, X_2, \dots, X_m)' = \sum_{i=0}^{m} \mathbf{Z}_i$$

with

$$\mathbf{Z}_i = \begin{cases} \mathbf{e}Y_0 & (i = 0) \\ \mathbf{e}_i Y_i. & (i = 1, 2, \dots, m) \end{cases}$$

We denote the distribution of $\mathbf{X}$ by f and the distribution of $\mathbf{Z}_i$ by f_i for $i = 0, 1, 2, \dots, m$. From (16.22), we obtain

$$\varphi_{f_0}(\mathbf{x}) = \begin{cases} m\varphi_{g_0}(y) & (\mathbf{x} = \mathbf{e}y;\ y = 1, 2, \dots) \\ 0 & (\text{otherwise}) \end{cases} \tag{16.36}$$

$$\varphi_{f_i}(\mathbf{x}) = \begin{cases} \varphi_{g_i}(y) & (\mathbf{x} = \mathbf{e}_i y;\ y = 1, 2, \dots) \\ 0. & (\text{otherwise}) \end{cases} \qquad (i = 1, 2, \dots, m) \tag{16.37}$$

As $\varphi_f = \sum_{i=0}^{m} \varphi_{f_i}$, we get

$$\varphi_f(\mathbf{x}) = \begin{cases} m\varphi_{g_0}(y) & (\mathbf{x} = \mathbf{e}y;\ y = 1, 2, \dots) \\ \varphi_{g_i}(y) & (\mathbf{x} = \mathbf{e}_i y;\ y = 1, 2, \dots;\ i = 1, 2, \dots, m) \\ 0, & (\text{otherwise}) \end{cases}$$

and insertion in (16.1) and (16.4) gives

$$f(\mathbf{x}) = \frac{1}{x_\bullet}\left(m \sum_{y=1}^{\min x_i} \varphi_{g_0}(y) f(\mathbf{x} - \mathbf{e}y) + \sum_{i=1}^{m} \sum_{y=1}^{x_i} \varphi_{g_i}(y) f(\mathbf{x} - \mathbf{e}_i y) \right) \qquad (\mathbf{x} \in \mathbb{N}_{m+})$$

$$f(\mathbf{x}) = \frac{1}{x_j}\left(\sum_{y=1}^{\min x_i} \varphi_{g_0}(y) f(\mathbf{x} - \mathbf{e}y) + \sum_{y=1}^{x_j} \varphi_{g_j}(y) f(\mathbf{x} - \mathbf{e}_j y) \right).$$

$$(\mathbf{x} \geq \mathbf{e}_j;\ j = 1, 2, \dots, m)$$

If each g_i is a compound Poisson distribution with Poisson parameter λ_i and severity distribution $h_i \in \mathcal{P}_{11}$, then similar to the previous example, we obtain

$$\varphi_f(\mathbf{x}) = \begin{cases} m\lambda_0 y h_0(y) & (\mathbf{x} = \mathbf{e}y;\ y = 1, 2, \dots) \\ \lambda_i y h_i(y) & (\mathbf{x} = \mathbf{e}_i y;\ y = 1, 2, \dots;\ i = 1, 2, \dots, m) \\ 0 & (\text{otherwise}) \end{cases}$$

$$f(\mathbf{x}) = \frac{1}{x_\bullet}\left(m\lambda_0 \sum_{y=1}^{\min x_j} y h_0(y) f(\mathbf{x} - \mathbf{e}y) + \sum_{i=1}^{m} \lambda_i \sum_{y=1}^{x_i} y h_i(y) f(\mathbf{x} - \mathbf{e}_i y) \right)$$

$$(\mathbf{x} \in \mathbb{N}_{m+})$$

$$f(\mathbf{x}) = \frac{1}{x_j}\left(\lambda_0 \sum_{y=1}^{\min x_i} y h_0(y) f(\mathbf{x} - \mathbf{e}y) + \lambda_j \sum_{y=1}^{x_j} y h_j(y) f(\mathbf{x} - \mathbf{e}_j y)\right).$$
$$(\mathbf{x} \geq \mathbf{e}_j;\ j = 1, 2, \ldots, m)$$

In the special case when each g_i is the Poisson distribution $\mathrm{Po}(\lambda_i)$, these formulae reduce to

$$\varphi_f(\mathbf{x}) = \begin{cases} m\lambda_0 & (\mathbf{x} = \mathbf{e}) \\ \lambda_i & (\mathbf{x} = \mathbf{e}_i;\ i = 1, 2, \ldots, m) \\ 0 & (\text{otherwise}) \end{cases}$$

$$f(\mathbf{x}) = \frac{1}{x_\bullet}\left(m\lambda_0 f(\mathbf{x} - \mathbf{e}) + \sum_{i=1}^{m} \lambda_i f(\mathbf{x} - \mathbf{e}_i)\right) \qquad (\mathbf{x} \in \mathbb{N}_{m+})$$

$$f(\mathbf{x}) = \frac{1}{x_j}(\lambda_0 f(\mathbf{x} - \mathbf{e}) + \lambda_j f(\mathbf{x} - \mathbf{e}_j)). \qquad (\mathbf{x} \geq \mathbf{e}_j;\ j = 1, 2, \ldots, m)$$ □

16.7 Individual Models

In this section, we shall extend parts of the theory of Chap. 7 to multivariate distributions.

Let us assume that we want to evaluate $f = *_{j=1}^{M} f_j$ with $f_1, f_2, \ldots, f_M \in \mathcal{P}_{m\underline{\mathbf{0}}}$.

In De Pril's first method, we first evaluate the De Pril transform of each f_j by (16.2), then we find the De Pril transform of f by summing these De Pril transforms, and finally we evaluate f by (16.1); we can express these last two steps by

$$f(\mathbf{x}) = \frac{1}{x_\bullet}\sum_{j=1}^{M}\sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} \varphi_{f_j}(\mathbf{y}) f(\mathbf{x} - \mathbf{y}). \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{16.38}$$

In De Pril's second method, we express each f_j as a compound distribution $p_j \vee h_j$ where p_j is the Bernoulli distribution $\mathrm{Bern}(\pi_j)$ with $\pi_j = 1 - f_j(\mathbf{0})$ and the severity distribution $h_j \in \mathcal{P}_{m+}$ is given by $h_j(\mathbf{y}) = f_j(\mathbf{y})/\pi_j$ for $\mathbf{y} \in \mathbb{N}_{m+}$. Insertion of (6.20) in (16.3) gives

$$\varphi_{f_j}(\mathbf{x}) = -x_\bullet \sum_{n=1}^{x_\bullet} \frac{1}{n}\left(\frac{\pi_j}{\pi_j - 1}\right)^n h_j^{n*}(\mathbf{x}), \qquad (\mathbf{x} \in \mathbb{N}_{m+})$$

from which we obtain

$$\varphi_f(\mathbf{x}) = -x_\bullet \sum_{n=1}^{x_\bullet} \frac{1}{n} \sum_{j=1}^{M}\left(\frac{\pi_j}{\pi_j - 1}\right)^n h_j^{n*}(\mathbf{x}). \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{16.39}$$

With

$$\sigma_j(\mathbf{x}) = \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}} \varphi_{f_j}(\mathbf{y}) f(\mathbf{x}-\mathbf{y}), \qquad (\mathbf{x} \in \mathbb{N}_{m+};\ j = 1, 2, \dots, M) \tag{16.40}$$

we rewrite (16.38) as

$$f(\mathbf{x}) = \frac{1}{x_\bullet} \sum_{j=1}^{M} \sigma_j(\mathbf{x}). \qquad (\mathbf{x} \in \mathbb{N}_{m+})$$

The following theorem extends Corollary 7.1.

Theorem 16.8 *Let* $f_1, f_2, \dots, f_M \in \mathcal{P}_{m\underline{\mathbf{0}}}$ *and* $f = *_{j=1}^{M} f_j$. *If* f_j *is a compound distribution with counting distribution* $R_k[a, b]$ *and severity distribution* $h \in \mathcal{P}_{m\mathbf{0}}$ *for some* j, *then*

$$\sigma_j(\mathbf{x}) = \frac{1}{1-\tau_a(h(\mathbf{0}))} \times \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}} (y_\bullet((a+\Psi b)\vee h)(\mathbf{y}) f(\mathbf{x}-\mathbf{y}) + (a \vee h)(\mathbf{y})\sigma_j(\mathbf{x}-\mathbf{y})).$$

$$(\mathbf{x} \in \mathbb{N}_{m+}) \tag{16.41}$$

Proof By starting with (16.40) and successive application of Corollary 16.2 and (16.40), we obtain that for all $\mathbf{x} \in \mathbb{N}_{m+}$

$$\begin{aligned}
\sigma_j(\mathbf{x}) &= \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}} \varphi_{f_j}(\mathbf{y}) f(\mathbf{x}-\mathbf{y}) \\
&= \frac{1}{1-\tau_a(h(\mathbf{0}))} \\
&\quad \times \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}} \left(y_\bullet((a+\Psi b)\vee h)(\mathbf{y}) + \sum_{\mathbf{0}<\mathbf{z}<\mathbf{y}} (a\vee h)(\mathbf{z})\varphi_{f_j}(\mathbf{y}-\mathbf{z}) \right) f(\mathbf{x}-\mathbf{y}) \\
&= \frac{1}{1-\tau_a(h(\mathbf{0}))} \sum_{\mathbf{0}<\mathbf{z}\le\mathbf{x}} \left(z_\bullet((a+\Psi b)\vee h)(\mathbf{z}) f(\mathbf{x}-\mathbf{z}) \right. \\
&\quad \left. + (a\vee h)(\mathbf{z}) \sum_{\mathbf{z}<\mathbf{y}\le\mathbf{x}} \varphi_{f_j}(\mathbf{y}-\mathbf{z}) f(\mathbf{x}-\mathbf{y}) \right) \\
&= \frac{1}{1-\tau_a(h(\mathbf{0}))} \\
&\quad \times \sum_{\mathbf{0}<\mathbf{z}\le\mathbf{x}} (z_\bullet((a+\Psi b)\vee h)(\mathbf{z}) f(\mathbf{x}-\mathbf{z}) + (a\vee h)(\mathbf{z})\sigma_j(\mathbf{x}-\mathbf{z})).
\end{aligned}$$

□

The following theorem extends Dhaene–Vandebroek's method to the multivariate case.

Theorem 16.9 *Let* $f_1, f_2, \dots, f_M \in \mathcal{P}_{m\underline{\mathbf{0}}}$ *and* $f = *_{j=1}^{M} f_j$. *Then*

$$\sigma_j(\mathbf{x}) = \frac{1}{f_j(\mathbf{0})} \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} (y_{\bullet} f(\mathbf{x}-\mathbf{y}) - \sigma_j(\mathbf{x}-\mathbf{y})) f_j(\mathbf{y}). \tag{16.42}$$

$$(\mathbf{x} \in \mathbb{N}_{m+};\ j = 1, 2, \dots, M)$$

Proof For all $\mathbf{x} \in \mathbb{N}_{m+}$, application of (16.1) and (16.40) gives

$$\begin{aligned}
\sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} y_{\bullet} f(\mathbf{x}-\mathbf{y}) f_j(\mathbf{y}) &= \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} f(\mathbf{x}-\mathbf{y}) \sum_{\mathbf{0}<\mathbf{z}\leq\mathbf{y}} \varphi_{f_j}(\mathbf{z}) f_j(\mathbf{y}-\mathbf{z}) \\
&= \sum_{\mathbf{0}<\mathbf{z}\leq\mathbf{x}} \sum_{\mathbf{z}\leq\mathbf{y}\leq\mathbf{x}} \varphi_{f_j}(\mathbf{z}) f(\mathbf{x}-\mathbf{y}) f_j(\mathbf{y}-\mathbf{z}) \\
&= \sum_{\mathbf{0}<\mathbf{z}\leq\mathbf{x}} \sum_{\mathbf{0}\leq\mathbf{u}\leq\mathbf{x}-\mathbf{z}} \varphi_{f_j}(\mathbf{z}) f(\mathbf{x}-\mathbf{u}-\mathbf{z}) f_j(\mathbf{u}) \\
&= \sum_{\mathbf{0}\leq\mathbf{u}\leq\mathbf{x}} f_j(\mathbf{u}) \sum_{\mathbf{0}<\mathbf{z}\leq\mathbf{x}-\mathbf{u}} \varphi_{f_j}(\mathbf{z}) f(\mathbf{x}-\mathbf{u}-\mathbf{z}) \\
&= \sum_{\mathbf{0}\leq\mathbf{u}\leq\mathbf{x}} f_j(\mathbf{u}) \sigma_j(\mathbf{x}-\mathbf{u}),
\end{aligned}$$

and by solving for $\sigma_j(\mathbf{x})$, we obtain (16.42). □

The presentation in Sect. 7.3 of De Pril's individual model and methodology for evaluation the aggregate claims distribution within that model easily extends to the multivariate case with $h_1, h_2, \dots, h_m \in \mathcal{P}_{m+}$. In particular, in this setting, (7.11) extends to

$$\varphi_f(\mathbf{x}) = -x_{\bullet} \sum_{n=\mathbf{1}}^{x_{\bullet}} \frac{1}{n} \sum_{j=1}^{J} \left(\frac{\pi_j}{\pi_j - 1} \right)^n \sum_{i=1}^{I} M_{ij} h_i^{n*}(\mathbf{x}). \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{16.43}$$

In Chap. 18, we shall discuss approximations to φ_f by approximating the counting distributions. As these are still the same Bernoulli distributions for which we deduced approximations for the univariate case in Chap. 10, we simply insert these same approximations in (16.43) in the multivariate case.

Further Remarks and References

This chapter is to a large extent based on Sundt (2000b).

Papageorgiou (1984) deduced a recursion similar to (16.1).

The recursion (16.33) was presented by Teicher (1954). By using moment generating functions, Ambagaspitiya (1999) showed that the distribution f in Example 16.1 can be expressed as a compound Poisson distribution with Poisson parameter $\lambda_\bullet$ and severity distribution h given by (16.30). For more information on multivariate Poisson distributions, see Johnson et al. (1997).

Walhin and Paris (2001c) considered De Pril's second method and the Dhaene–Vandebroek method in the bivariate case and presented a numerical application from reinsurance.

Chapter 17
Moments

Summary

The purpose of the present chapter is to extend to a multivariate setting results that were deduced in a univariate setting in Chap. 9. Like that chapter, the present chapter is divided into two sections, Sect. 17.1 on convolutions and Sect. 17.2 on compound distributions.

In Sect. 17.1.1, we develop recursions for ordinary moments of convolutions of a multivariate distribution. We apply one of these recursions to prove a characterisation of multinormal distributions in Sect. 17.1.2.

In Sect. 17.2.1, we deduce recursions for moments of multivariate compound distributions of Type 1 with counting distribution satisfying (5.6). The special case with counting distribution in the Panjer class is treated in Sect. 17.2.2.

17.1 Convolutions of a Distribution

17.1.1 Moments

In Sect. 17.1, we shall discuss recursions for moments of M-fold convolutions of multivariate distributions. As we do not need to restrict to distributions on vectors of integers, we shall identify the distributions by their cumulative distribution function.

The following lemma extends Lemma 9.1 to the multivariate case and is proved in the same way by conditioning on $\mathbf{X}+\mathbf{Y}$.

Lemma 17.1 *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be independent random* $m \times 1$ *vectors such that the distribution of* $\mathbf{X}$ *is the* M*-fold convolution of the distribution of* $\mathbf{Y}$*. Then*

$$\mathrm{E}(X_l - MY_l)r(\mathbf{X}+\mathbf{Y}) = 0$$

for any function r *and* $l = 1, 2, \ldots, m$*.*

We are now ready to extend Theorem 9.1 to the multivariate case.

Theorem 17.1 *Let* G *be an* m*-variate distribution and* $F = G^{M*}$ *for some positive integer* M*. Then, for any* $m \times 1$ *vector* $\mathbf{c}$*, we have*

$$\mu_F(\mathbf{j};\mathbf{c}) = \sum_{\mathbf{e}_l \le \mathbf{u} \le \mathbf{j}} \binom{j_l - 1}{u_l - 1}\left(\prod_{i \ne l}\binom{j_i}{u_i}\right)$$
$$\times \left(\left(M + 1 - \frac{j_l}{u_l}\right)\mu_G(\mathbf{u}) - c_l\mu_G(\mathbf{u}-\mathbf{e}_l)\right)\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c})$$

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_17, © Springer-Verlag Berlin Heidelberg 2009

$$
\begin{aligned}
&- \sum_{\mathbf{0}<\mathbf{u}\le\mathbf{j}-j_l\mathbf{e}_l}\left(\prod_{i=1}^{m}\binom{j_i}{u_i}\right)\mu_G(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c})\\
&= \sum_{\mathbf{e}_l\le\mathbf{u}\le\mathbf{j}}\left(\prod_{i=1}^{m}\binom{j_i}{u_i}\right)\\
&\quad\times\left(\left(\frac{u_l}{j_l}(M+1)-1\right)\mu_G(\mathbf{u})-\frac{u_l}{j_l}c_l\mu_G(\mathbf{u}-\mathbf{e}_l)\right)\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c})\\
&- \sum_{\mathbf{0}<\mathbf{u}\le\mathbf{j}-j_l\mathbf{e}_l}\left(\prod_{i=1}^{m}\binom{j_i}{u_i}\right)\mu_G(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c}). \qquad (17.1)
\end{aligned}
$$

$(\mathbf{j}\ge\mathbf{e}_l;\, l=1,2,\dots,m)$

Proof Let $\mathbf{X}$ and $\mathbf{Y}$ be independent random $m\times 1$ vectors with distribution F and G respectively, and let $\mathbf{j}\ge\mathbf{e}_l$ for some $l\in\{1,2,\dots,m\}$. By letting

$$
r(\mathbf{x})=(x_l-c_l)^{j_l-1}t(\mathbf{x}) \qquad (\mathbf{x}\in\mathbb{R}^m) \qquad (17.2)
$$

in Lemma 17.1 for some function t and proceeding like in the proof of Theorem 9.1, we obtain

$$
\begin{aligned}
&\mathrm{E}(X_l-c_l)^{j_l}t(\mathbf{X}+\mathbf{Y})\\
&=\sum_{u_l=1}^{j_l}\binom{j_l-1}{u_l-1}\mathrm{E}\left(\left(M+1-\frac{j_l}{u_l}\right)Y_l^{u_l}-c_lY_l^{u_l-1}\right)(X_l-c_l)^{j_l-u_l}t(\mathbf{X}+\mathbf{Y}).
\end{aligned} \qquad (17.3)
$$

Letting

$$
t(\mathbf{x})=\prod_{i\ne l}(x_i-c_i)^{j_i} \qquad (\mathbf{x}\in\mathbb{R}^m) \qquad (17.4)
$$

gives

$$
t(\mathbf{X}+\mathbf{Y})=\prod_{i\ne l}(X_i+Y_i-c_i)^{j_i}=\prod_{i\ne l}\sum_{u_i=0}^{j_i}\binom{j_i}{u_i}Y_i^{u_i}(X_i-c_i)^{j_i-u_i}. \qquad (17.5)
$$

Insertion in the left-hand side and the right-hand side of (17.3) gives

$$
\begin{aligned}
&\mathrm{E}(X_l-c_l)^{j_l}t(\mathbf{X}+\mathbf{Y})\\
&=\sum_{\mathbf{0}\le\mathbf{u}\le\mathbf{j}-j_l\mathbf{e}_l}\left(\prod_{i\ne l}\binom{j_i}{u_i}\right)\mu_G(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c})\\
&=\mu_F(\mathbf{j};\mathbf{c})+\sum_{\mathbf{0}<\mathbf{u}\le\mathbf{j}-j_l\mathbf{e}_l}\left(\prod_{i\ne l}\binom{j_i}{u_i}\right)\mu_G(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c})
\end{aligned}
$$

and

$$\sum_{u_l=1}^{j_l}\binom{j_l-1}{u_l-1}\mathrm{E}\left(\left(M+1-\frac{j_l}{u_l}\right)Y_l^{u_l}-c_lY_l^{u_l-1}\right)(X_l-c_l)^{j_l-u_l}t(\mathbf{X}+\mathbf{Y})$$
$$=\sum_{\mathbf{e}_l\le\mathbf{u}\le\mathbf{j}}\binom{j_l-1}{u_l-1}\left(\prod_{i\neq l}\binom{j_i}{u_i}\right)$$
$$\times\left(\left(M+1-\frac{j_l}{u_l}\right)\mu_G(\mathbf{u})-c_l\mu_G(\mathbf{u}-\mathbf{e}_l)\right)\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c}),$$

and by equating these two expressions and solving for $\mu_F(\mathbf{j};\mathbf{c})$, we obtain (17.1). □

For all the recursions that we deduce in Chap. 17, we use the initial condition $\mu_F(\mathbf{0};\mathbf{c})=1$.

We see that like the recursion (15.21), the recursion (17.1) depends on one of the dimensions in a special way and cannot be applied on the whole range of the recursion variable, so that we normally will have to apply it with more than one value of l. With (15.21), it was simple to combine the recursions to get the universal recursion (15.22). However, in (17.1), the dependence on the special dimensions seems too complex for doing something similar.

When $\mathbf{c}=\mathbf{0}$, the recursion (17.1) reduces to

$$\mu_F(\mathbf{j})=\sum_{\mathbf{e}_l\le\mathbf{u}\le\mathbf{j}}\binom{j_l-1}{u_l-1}\left(\prod_{i\neq l}\binom{j_i}{u_i}\right)\left(M+1-\frac{j_l}{u_l}\right)\mu_G(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u})$$
$$-\sum_{\mathbf{0}<\mathbf{u}\le\mathbf{j}-j_l\mathbf{e}_l}\left(\prod_{i=1}^{m}\binom{j_i}{u_i}\right)\mu_G(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u})$$
$$=\sum_{\mathbf{e}_l\le\mathbf{u}\le\mathbf{j}}\left(\prod_{i=1}^{m}\binom{j_i}{u_i}\right)\left(\frac{u_l}{j_l}(M+1)-1\right)\mu_G(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u})$$
$$-\sum_{\mathbf{0}<\mathbf{u}\le\mathbf{j}-j_l\mathbf{e}_l}\left(\prod_{i=1}^{m}\binom{j_i}{u_i}\right)\mu_G(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u}).\qquad(\mathbf{j}\ge\mathbf{e}_l;\ l=1,2,\dots,m)$$

Example 17.1 In the special case when $m=2$ and $l=1$, (17.1) gives

$$\mu_F(j_1,j_2;\mathbf{c})=\sum_{u_2=0}^{j_2}\binom{j_2}{u_2}\sum_{u_1=1}^{j_1}\binom{j_1-1}{u_1-1}$$
$$\times\left(\left(M+1-\frac{j_1}{u_1}\right)\mu_G(u_1,u_2)-c_1\mu_G(u_1-1,u_2)\right)$$
$$\times\mu_F(j_1-u_1,j_2-u_2;\mathbf{c})$$

$$-\sum_{u=1}^{j_2}\binom{j_2}{u}\mu_G(0,u)\mu_F(j_1,j_2-u;\mathbf{c}).$$

$$(j_1=1,2,\dots;\ j_2=0,1,2,\dots)$$ □

Example 17.2 Let G be the distribution of an $m\times 1$ random vector that we split into two subvectors, and let G_1 be the marginal distribution of the first one. Then, using the notation of Sect. 16.5,

$$\mu_{G_1}(\mathbf{j})=\mu_G\left(\begin{pmatrix}\mathbf{j}\\ \mathbf{0}\end{pmatrix}\right);\qquad \mu_{G_1^{M*}}(\mathbf{j};\mathbf{c})=\mu_{G^{M*}}\left(\begin{pmatrix}\mathbf{j}\\ \mathbf{0}\end{pmatrix},\begin{pmatrix}\mathbf{c}\\ \mathbf{0}\end{pmatrix}\right),$$

and insertion in (17.1) gives the corresponding recursion for $\mu_{G_1^{M*}}(\mathbf{j};\mathbf{c})$ when $l=1,2,\dots,m_1$. □

Example 17.3 Let $f\in\mathcal{P}_{m+}$ be the multinomial distribution $\mathrm{mnom}(M,\boldsymbol{\pi})$. Then, $f=g^{M*}$ where $g\in\mathcal{P}_{m+}$ is the m-variate Bernoulli distribution given by (14.3). In this case, we have

$$\mu_g(\mathbf{j})=\begin{cases}\pi_i & (\mathbf{j}=j\mathbf{e}_i;\ j=1,2,\dots;\ i=1,2,\dots,m)\\ 1 & (\mathbf{j}=\mathbf{0})\\ 0. & (\text{otherwise})\end{cases}$$

Insertion in (17.1) gives that for $\mathbf{j}\ge\mathbf{e}_l$ and $l=1,2,\dots,m$,

$$\begin{aligned}\mu_f(\mathbf{j};\mathbf{c})&=\pi_l\sum_{u=1}^{j_l}\binom{j_l-1}{u-1}\left(M+1-\frac{j_l}{u}-c_l\right)\mu_f(\mathbf{j}-\mathbf{e}_l u;\mathbf{c})\\&\quad-\sum_{i\neq l}\pi_i\sum_{u=1}^{j_i}\binom{j_i}{u}(\mu_f(\mathbf{j}-\mathbf{e}_i u;\mathbf{c})+c_l\mu_f(\mathbf{j}-\mathbf{e}_l-\mathbf{e}_i u;\mathbf{c}))\\&\quad-c_l(1-\pi_l)\mu_f(\mathbf{j}-\mathbf{e}_l;\mathbf{c}).\end{aligned}$$

When $\mathbf{c}=\mathbf{0}$, this reduces to

$$\mu_f(\mathbf{j})=\pi_l\sum_{u=1}^{j_l}\binom{j_l-1}{u-1}\left(M+1-\frac{j_l}{u}\right)\mu_f(\mathbf{j}-\mathbf{e}_l u)-\sum_{i\neq l}\pi_i\sum_{u=1}^{j_i}\binom{j_i}{u}\mu_f(\mathbf{j}-\mathbf{e}_i u).$$
□

Let $\mathbf{X}$ be a random $m\times 1$ vector with distribution F. With $\tilde{F}$ denoting the distribution of $(X_1/d_1, X_2/d_2,\dots,X_m/d_m)'$ with $d_i\neq 0$ for $i=1,2,\dots,m$, (17.1) gives

$$\mu_{\tilde{F}}(\mathbf{j};\mathbf{c})=\sum_{\mathbf{e}_l\le\mathbf{u}\le\mathbf{j}}\binom{j_l-1}{u_l-1}\left(\prod_{i\neq l}\binom{j_i}{u_i}\right)\frac{1}{\prod_{i=1}^m d_i^{u_i}}$$

$$\times\left(\left(M+1-\frac{j_l}{u_l}\right)\mu_G(\mathbf{u})-c_l d_l \mu_G(\mathbf{u}-\mathbf{e}_l)\right)\mu_{\tilde{F}}(\mathbf{j}-\mathbf{u};\mathbf{c})$$

$$-\sum_{\mathbf{0}<\mathbf{u}\leq\mathbf{j}-j_l\mathbf{e}_l}\left(\prod_{i=1}^{m}\binom{j_i}{u_i}\right)\frac{\mu_G(\mathbf{u})}{\prod_{i=1}^{m} d_i^{u_i}}\mu_{\tilde{F}}(\mathbf{j}-\mathbf{u};\mathbf{c}).$$

$$(\mathbf{j}\geq\mathbf{e}_l;\ l=1,2,\ldots,m)$$

In particular, with $d_i=\sqrt{M}$ for $i=1,2,\ldots,m$ and $\mathbf{c}=\mathrm{E}\,\mathbf{X}=\mathbf{0}$, we obtain

$$\mu_{\tilde{F}}(\mathbf{j})=\sum_{\mathbf{e}_l<\mathbf{u}\leq\mathbf{j}}\binom{j_l-1}{u_l-1}\left(\prod_{i\neq l}\binom{j_i}{u_i}\right)\left(M+1-\frac{j_l}{u_l}\right)\frac{\mu_G(\mathbf{u})}{M^{u_\bullet/2}}\mu_{\tilde{F}}(\mathbf{j}-\mathbf{u})$$

$$-\sum_{\mathbf{0}<\mathbf{u}\leq\mathbf{j}-j_l\mathbf{e}_l}\left(\prod_{i=1}^{m}\binom{j_i}{u_i}\right)\frac{\mu_G(\mathbf{u})}{M^{u_\bullet/2}}\mu_{\tilde{F}}(\mathbf{j}-\mathbf{u}). \tag{17.6}$$

$$(\mathbf{j}>\mathbf{e}_l;\ l=1,2,\ldots,m)$$

17.1.2 The Multinormal Distribution

Before extending Theorem 9.2 to the multivariate case, we shall recapitulate some properties of the multinormal distribution.

Let $\mathbf{Y}$ be an $m\times 1$ vector of independent and identically distributed random variables with the standard normal distribution N(0, 1), and let $\mathbf{\Omega}$ be a non-random $m\times m$ matrix and $\boldsymbol{\xi}$ a non-random $m\times 1$ vector. Then

$$\mathbf{X}=\mathbf{\Omega}\mathbf{Y}+\boldsymbol{\xi} \tag{17.7}$$

has the *multinormal distribution* $\mathrm{mN}(\boldsymbol{\xi},\mathbf{\Sigma})$ with $\mathbf{\Sigma}=\mathbf{\Omega}\mathbf{\Omega}'$. As for any $m\times 1$ vector $\mathbf{v}$,

$$\mathbf{v}'\mathbf{\Sigma}\mathbf{v}=\mathbf{v}'\mathbf{\Omega}\mathbf{\Omega}'\mathbf{v}=(\mathbf{\Omega}'\mathbf{v})'\mathbf{\Omega}'\mathbf{v}\geq 0,$$

the matrix $\mathbf{\Sigma}$ is symmetric and positive semi-definite. Furthermore, for any symmetric positive semi-definite $m\times m$ matrix $\mathbf{\Sigma}$, there exists an $m\times m$ matrix $\mathbf{\Omega}$ such that $\mathbf{\Sigma}=\mathbf{\Omega}\mathbf{\Omega}'$ so that our definition of the multinormal distribution covers all symmetric positive definite $m\times m$ matrices $\mathbf{\Sigma}$.

As the standard normal distribution has finite moments of all orders, (17.7) gives that this is also the case with multinormal distributions. In particular, as the standard normal distribution has mean zero, (17.7) gives that the multinormal distribution $\mathrm{mN}(\boldsymbol{\xi},\mathbf{\Sigma})$ has mean $\boldsymbol{\xi}$.

We are now ready to extend Theorem 9.2 to the multivariate case.

Theorem 17.2 *An m-variate distribution G with finite moments of first and second order is multinormal with mean $\mathbf{0}$ iff*

$$G^{M*}(\mathbf{x}) = G\big(\mathbf{x}/\sqrt{M}\big) \qquad (\mathbf{x} \in \mathbb{R}^m) \tag{17.8}$$

for some integer $M > 1$.

Proof We first assume that G is an m-variate multinormal distribution with mean $\mathbf{0}$. Then G is the distribution of $\mathbf{X} = \boldsymbol{\Omega}\mathbf{Y}$ for some non-random $m \times m$ matrix $\boldsymbol{\Omega}$ and $\mathbf{Y}$ being an $m \times 1$ vector of independent and identically distributed random variables with standard normal distribution. As the standard normal distribution satisfies (9.9), it follows that G satisfies (17.8).

Let us now assume that (17.8) holds for some integer $M > 1$. This means that if $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_M$ are independent and identically distributed random vectors with distribution G, then $\mathbf{Y}_{\bullet M}/\sqrt{M}$ also has distribution G. Then they must have the same mean, which must be equal to $\mathbf{0}$. From (17.6), we obtain

$$\mu_G(\mathbf{j}) = \sum_{\mathbf{e}_l < \mathbf{u} \le \mathbf{j}} \binom{j_l - 1}{u_l - 1}\left(\prod_{i \ne l}\binom{j_i}{u_i}\right)\left(M + 1 - \frac{j_l}{u_l}\right)\frac{\mu_G(\mathbf{u})}{M^{u_\bullet/2}}\mu_G(\mathbf{j}-\mathbf{u})$$
$$- \sum_{\mathbf{0} < \mathbf{u} \le \mathbf{j} - j_l\mathbf{e}_l}\left(\prod_{i=1}^{m}\binom{j_i}{u_i}\right)\frac{\mu_G(\mathbf{u})}{M^{u_\bullet/2}}\mu_G(\mathbf{j}-\mathbf{u}). \qquad (\mathbf{j} > \mathbf{e}_l;\, l = 1, 2, \ldots, m)$$

Solving for $\mu_G(\mathbf{j})$ gives

$$\mu_G(\mathbf{j}) = \frac{1}{1 - M^{1 - j_\bullet/2}}\left(\sum_{\mathbf{e}_l < \mathbf{u} < \mathbf{j}} \binom{j_l - 1}{u_l - 1}\left(\prod_{i \ne l}\binom{j_i}{u_i}\right)\left(M + 1 - \frac{j_l}{u_l}\right)\frac{\mu_G(\mathbf{u})}{M^{u_\bullet/2}}\right.$$
$$\times\, \mu_G(\mathbf{j}-\mathbf{u})$$
$$\left. - \sum_{\mathbf{0} < \mathbf{u} \le \mathbf{j} - j_l\mathbf{e}_l}\left(\prod_{i=1}^{m}\binom{j_i}{u_i}\right)\frac{\mu_G(\mathbf{u})}{M^{u_\bullet/2}}\mu_G(\mathbf{j}-\mathbf{u})\right).$$
$$(\mathbf{j} > \mathbf{e}_l;\, j_\bullet > 2;\, l = 1, 2, \ldots, m)$$

This recursion determines all moments of G from $\mu_G(\mathbf{j}) = 0$ for each $\mathbf{j}$ with $j_\bullet = 1$ and a given value of $\mu_G(\mathbf{j})$ for each $\mathbf{j}$ with $j_\bullet = 2$. As there exists a multinormal distribution with the same first and second order moments as G and this multinormal distribution satisfies (17.8) for all positive integers M, this multinormal distribution must have the same moments as G, and as the moment generating function of a multinormal distribution exists for all $m \times 1$ vectors with real elements, this implies that G must be equal to that multinormal distribution.

This completes the proof of Theorem 17.2. □

For $\mathbf{j} > \mathbf{e}_l$ and $l = 1, 2, \ldots, m$, letting $M \uparrow \infty$ in (17.6) gives the limiting expression

$$\mu_{\tilde{F}}(\mathbf{j}) = \sum_{\substack{\mathbf{e}_l < \mathbf{u} \leq \mathbf{j} \\ u_\bullet = 2}} \binom{j_l - 1}{u_l - 1} \left(\prod_{i \neq l} \binom{j_i}{u_i} \right) \mu_G(\mathbf{u}) \mu_{\tilde{F}}(\mathbf{j} - \mathbf{u}),$$

that is,

$$\begin{aligned} \mu_{\tilde{F}}(\mathbf{j}) = & \sum_{i=1}^{m} I(\mathbf{j} \geq \mathbf{e}_l + \mathbf{e}_i) j_i \mu_G(\mathbf{e}_l + \mathbf{e}_i) \mu_{\tilde{F}}(\mathbf{j} - \mathbf{e}_l - \mathbf{e}_i) \\ & - I(\mathbf{j} \geq 2\mathbf{e}_l) \mu_G(2\mathbf{e}_l) \mu_{\tilde{F}}(\mathbf{j} - 2\mathbf{e}_l), \end{aligned} \tag{17.9}$$

from which we see that the asymptotic moments of $\tilde{F}$ are determined by the second order moments of G. As there exists a multinormal distribution with the same first and second order moments as G and $\tilde{F}$ is multinormal when G is multinormal, the recursion (17.9) must generate the moments of a multinormal distribution, as expected. In particular, if G is multinormal, we obtain

$$\begin{aligned} \mu_G(\mathbf{j}) = & \sum_{i=1}^{m} I(\mathbf{j} \geq \mathbf{e}_l + \mathbf{e}_i) j_i \mu_G(\mathbf{e}_l + \mathbf{e}_i) \mu_G(\mathbf{j} - \mathbf{e}_l - \mathbf{e}_i) \\ & - I(\mathbf{j} \geq 2\mathbf{e}_l) \mu_G(2\mathbf{e}_l) \mu_G(\mathbf{j} - 2\mathbf{e}_l). \qquad (\mathbf{j} \geq \mathbf{e}_l;\ j_\bullet > 2;\ l = 1, 2, \ldots, m) \end{aligned}$$

17.2 Compound Distributions

17.2.1 General Results

In Sect. 17.2.1, we consider recursions for moments of a compound distribution $F = p \vee H$ with m-variate severity distribution H and counting distribution $p \in \mathcal{P}_{10}$ satisfying the recursion (5.6).

The following multivariate extension of Lemma 9.2 is proved in the same way as in the univariate case.

Lemma 17.2 *Let $\mathbf{X}, \mathbf{Y}_1, \mathbf{Y}_2, \ldots$ be independent random $m \times 1$ vectors, the $\mathbf{Y}_i$s with distribution H and $\mathbf{X}$ with distribution $p \vee H$ with $p \in \mathcal{P}_{10}$ satisfying the recursion (5.6) where k is a positive number or infinity. Then*

$$\begin{aligned} \mathrm{E}\, X_l r(\mathbf{X}) = & \sum_{n=1}^{\infty} q(n)\, \mathrm{E}\, Y_{\bullet nl} r(\mathbf{Y}_{\bullet n}) \\ & + \sum_{i=1}^{k} \mathrm{E}(a(i)(X_l + Y_{\bullet il}) + \Psi b(i) Y_{\bullet il}) r(\mathbf{X} + \mathbf{Y}_{\bullet i}) \end{aligned} \tag{17.10}$$

$$(l = 1, 2, \ldots, m)$$

for any function r for which these expectations exist.

We are now ready to prove the following multivariate extension of Theorem 9.3.

Theorem 17.3 *If $F = p \vee H$ with m-variate distribution H and $p \in \mathcal{P}_{10}$ satisfying* (5.6) *with $\mu_a(0) \neq 1$, then*

$$\begin{aligned}
\mu_F(\mathbf{j};\mathbf{c}) &= \frac{1}{1-\mu_a(0)}\Bigg(\mu_{q\vee H}(\mathbf{j};\mathbf{c}) + c_l\mu_{q\vee H}(\mathbf{j}-\mathbf{e}_l;\mathbf{c}) - c_l\mu_F(\mathbf{j}-\mathbf{e}_l;\mathbf{c}) \\
&\quad + \sum_{\mathbf{0}<\mathbf{u}\leq\mathbf{j}-j_l\mathbf{e}_l}\left(\prod_{i=1}^m\binom{j_i}{u_i}\right)\mu_{a\vee H}(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c}) \\
&\quad + \sum_{\mathbf{e}_l\leq\mathbf{u}\leq\mathbf{j}}\binom{j_l-1}{u_l-1}\left(\prod_{i\neq l}\binom{j_i}{u_i}\right) \\
&\quad \times\left(\frac{j_l}{u_l}\mu_{a\vee H}(\mathbf{u}) + \mu_{\Psi b\vee H}(\mathbf{u}) + c_l\mu_{a\vee H}(\mathbf{u}-\mathbf{e}_l)\right)\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c})\Bigg) \\
&= \frac{1}{1-\mu_a(0)}\Bigg(\mu_{q\vee H}(\mathbf{j};\mathbf{c}) + c_l\mu_{q\vee H}(\mathbf{j}-\mathbf{e}_l;\mathbf{c}) - c_l\mu_F(\mathbf{j}-\mathbf{e}_l;\mathbf{c}) \\
&\quad + \sum_{\mathbf{0}<\mathbf{u}\leq\mathbf{j}-j_l\mathbf{e}_l}\left(\prod_{i=1}^m\binom{j_i}{u_i}\right)\mu_{a\vee H}(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c}) \\
&\quad + \sum_{\mathbf{e}_l\leq\mathbf{u}\leq\mathbf{j}}\left(\prod_{i=1}^m\binom{j_i}{u_i}\right) \\
&\quad \times\left(\mu_{a\vee H}(\mathbf{u}) + \frac{u_l}{j_l}(\mu_{\Psi b\vee H}(\mathbf{u}) + c_l\mu_{a\vee H}(\mathbf{u}-\mathbf{e}_l))\right)\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c})\Bigg) \\
&(\mathbf{j}\geq\mathbf{e}_l;\ l=1,2,\dots,m) \qquad (17.11)
\end{aligned}$$

for any constant $m \times 1$ vector $\mathbf{c}$.

Proof Let $\mathbf{X}, \mathbf{Y}_1, \mathbf{Y}_2, \dots$ be independent random $m \times 1$ vectors, the $\mathbf{Y}_i$s with distribution H and $\mathbf{X}$ with distribution F.

With $\mathbf{c} \in \mathbb{R}^m$, $\mathbf{j} \geq \mathbf{e}_l$ for some $l \in \{1, 2, \dots, m\}$, and r given by (17.2), proceeding like in the deduction of (9.19), we obtain

$$\begin{aligned}
&\mathrm{E}(X_l - c_l)^{j_l} t(\mathbf{X}) \\
&= \sum_{n=1}^{\infty} q(n)\,\mathrm{E}(Y_{\bullet nl} - c_l)^{j_l} t(\mathbf{Y}_{\bullet n}) \\
&\quad + c_l \sum_{n=1}^{\infty} q(n)\,\mathrm{E}(Y_{\bullet nl} - c_l)^{j_l-1} t(\mathbf{Y}_{\bullet n}) - c_l\,\mathrm{E}(X_l - c_l)^{j_l-1} t(\mathbf{X})
\end{aligned}$$

$$+\sum_{i=1}^{k} a(i)\,\mathrm{E}(X_l-c_l)^{j_l} t(\mathbf{X}+\mathbf{Y}_{\bullet i})$$

$$+\sum_{u_l=1}^{j_l}\binom{j_l-1}{u_l-1}\sum_{i=1}^{k}\mathrm{E}\biggl(\Bigl(\frac{j_l}{u_l}a(i)+\Psi b(i)\Bigr)Y_{\bullet il}^{u_l}$$

$$+c_l a(i) Y_{\bullet il}^{u_l-1}\biggr)(X_l-c_l)^{j_l-u_l} t(\mathbf{X}+\mathbf{Y}_{\bullet i}).$$

With t given by (17.4) and using (17.5), this gives

$$\begin{aligned}\mu_F(\mathbf{j};\mathbf{c}) &= \mu_{q\vee H}(\mathbf{j};\mathbf{c})+c_l\mu_{q\vee H}(\mathbf{j}-\mathbf{e}_l;\mathbf{c})-c_l\mu_F(\mathbf{j}-\mathbf{e}_l;\mathbf{c})\\ &\quad+\sum_{\mathbf{0}\le\mathbf{u}\le\mathbf{j}-j_l\mathbf{e}_l}\Bigl(\prod_{i=1}^{m}\binom{j_i}{u_i}\Bigr)\mu_{a\vee H}(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c})\\ &\quad+\sum_{\mathbf{e}_l\le\mathbf{u}\le\mathbf{j}}\binom{j_l-1}{u_l-1}\Bigl(\prod_{i\ne l}\binom{j_i}{u_i}\Bigr)\\ &\quad\times\Bigl(\frac{j_l}{u_l}\mu_{a\vee H}(\mathbf{u})+\mu_{\Psi b\vee H}(\mathbf{u})+c_l\mu_{a\vee H}(\mathbf{u}-\mathbf{e}_l)\Bigr)\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c}),\end{aligned}$$

and by solving for $\mu_F(\mathbf{j};\mathbf{c})$, we obtain (17.11). □

In particular, for all $p\in\mathcal{P}_{1\underline{0}}$, we have

$$\mu_F(\mathbf{j};\mathbf{c})=\sum_{\mathbf{e}_l\le\mathbf{u}\le\mathbf{j}}\binom{j_l-1}{u_l-1}\Bigl(\prod_{i\ne l}\binom{j_i}{u_i}\Bigr)\mu_{\Psi\varphi_p\vee H}(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c})-c_l\mu_F(\mathbf{j}-\mathbf{e}_l;\mathbf{c}).$$

$(\mathbf{j}\ge\mathbf{e}_l;\ l=1,2,\dots,m)$

17.2.2 Compound Panjer Distributions

In Sect. 17.2.2, we consider the special case when p is $R_1[a,b]$. Under this assumption, (17.11) reduces to

$$\begin{aligned}\mu_F(\mathbf{j};\mathbf{c}) &= \frac{1}{1-a}\biggl(a\sum_{\mathbf{0}<\mathbf{u}\le\mathbf{j}-j_l\mathbf{e}_l}\Bigl(\prod_{i=1}^{m}\binom{j_i}{u_i}\Bigr)\mu_H(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c})\\ &\quad+\sum_{\mathbf{e}_l\le\mathbf{u}\le\mathbf{j}}\binom{j_l-1}{u_l-1}\Bigl(\prod_{i\ne l}\binom{j_i}{u_i}\Bigr)\Bigl(\Bigl(\frac{j_l}{u_l}a+b\Bigr)\mu_H(\mathbf{u})\\ &\quad+c_l a\mu_H(\mathbf{u}-\mathbf{e}_l)\Bigr)\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c})-c_l\mu_F(\mathbf{j}-\mathbf{e}_l;\mathbf{c})\biggr)\end{aligned}$$

$$= \frac{1}{1-a}\left(a \sum_{\mathbf{0}<\mathbf{u}\le\mathbf{j}-j_l\mathbf{e}_l} \left(\prod_{i=1}^{m}\binom{j_i}{u_i}\right)\mu_H(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c})\right.$$
$$+ \sum_{\mathbf{e}_l\le\mathbf{u}\le\mathbf{j}} \left(\prod_{i=1}^{m}\binom{j_i}{u_i}\right)\left(\left(a+\frac{u_l}{j_l}b\right)\mu_H(\mathbf{u})\right.$$
$$\left.\left. + \frac{u_l}{j_l}c_l a \mu_H(\mathbf{u}-\mathbf{e}_l)\right)\mu_F(\mathbf{j}-\mathbf{u};\mathbf{c}) - c_l\mu_F(\mathbf{j}-\mathbf{e}_l;\mathbf{c})\right). \qquad (17.12)$$

In particular, with $\mathbf{c} = \mathbf{0}$, we obtain

$$\mu_F(\mathbf{j}) = \frac{1}{1-a}\left(a \sum_{\mathbf{0}<\mathbf{u}\le\mathbf{j}-j_l\mathbf{e}_l} \left(\prod_{i=1}^{m}\binom{j_i}{u_i}\right)\mu_H(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u})\right.$$
$$\left. + \sum_{\mathbf{e}_l\le\mathbf{u}\le\mathbf{j}} \binom{j_l-1}{u_l-1}\left(\prod_{i\ne l}\binom{j_i}{u_i}\right)\left(\frac{j_l}{u_l}a+b\right)\mu_H(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u})\right)$$
$$= \frac{1}{1-a}\left(a \sum_{\mathbf{0}<\mathbf{u}\le\mathbf{j}-j_l\mathbf{e}_l} \left(\prod_{i=1}^{m}\binom{j_i}{u_i}\right)\mu_H(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u})\right.$$
$$\left. + \sum_{\mathbf{e}_l\le\mathbf{u}\le j} \left(\prod_{i=1}^{m}\binom{j_i}{u_i}\right)\left(a+\frac{u_l}{j_l}b\right)\mu_H(\mathbf{u})\mu_F(\mathbf{j}-\mathbf{u})\right). \qquad (17.13)$$

Example 17.4 In the special case when $m = 2$ and $l = 1$, (17.12) gives

$$\mu_F(j_1, j_2;\mathbf{c}) = \frac{1}{1-a}\left(a\sum_{u=1}^{j_2}\binom{j_2}{u}\mu_H(0,u)\mu_F(j_1, j_2-u;\mathbf{c})\right.$$
$$+ \sum_{u_1=1}^{j_1}\binom{j_1-1}{u_1-1}\sum_{u_2=0}^{j_2}\binom{j_2}{u_2}\left(\left(\frac{j_1}{u_1}a+b\right)\mu_H(u_1,u_2)\right.$$
$$\left. + c_1 a\mu_H(u_1-1,u_2)\right)\mu_F(j_1-u_1, j_2-u_2;\mathbf{c})$$
$$\left. - c_1\mu_F(j_1-1, j_2;\mathbf{c})\right).$$

□

Let us now consider the recursions (17.12) and (17.13) in the three main subclasses of $\mathcal{R}_1$. The values of a and b are found in Table 2.1.

1. *Poisson distribution* $\mathrm{Po}(\lambda)$.

$$\mu_F(\mathbf{j};\mathbf{c}) = \lambda \sum_{\mathbf{e}_l \le \mathbf{u} \le \mathbf{j}} \binom{j_l - 1}{u_l - 1} \left(\prod_{i \ne l} \binom{j_i}{u_i} \right) \mu_H(\mathbf{u}) \mu_F(\mathbf{j} - \mathbf{u}; \mathbf{c}) - c_l \mu_F(\mathbf{j} - \mathbf{e}_l; \mathbf{c})$$

$$\mu_F(\mathbf{j}) = \lambda \sum_{\mathbf{e}_l \le \mathbf{u} \le \mathbf{j}} \binom{j_l - 1}{u_l - 1} \left(\prod_{i \ne l} \binom{j_i}{u_i} \right) \mu_H(\mathbf{u}) \mu_F(\mathbf{j} - \mathbf{u}).$$

2. *Binomial distribution* $\mathrm{bin}(M, \pi)$.

$$\begin{aligned} \mu_F(\mathbf{j};\mathbf{c}) = \pi \Bigg(& \sum_{\mathbf{e}_l \le \mathbf{u} \le \mathbf{j}} \binom{j_l - 1}{u_l - 1} \left(\prod_{i \ne l} \binom{j_i}{u_i} \right) \\ & \times \left(\left(M + 1 - \frac{j_l}{u_l} \right) \mu_H(\mathbf{u}) - c_l \mu_H(\mathbf{u} - \mathbf{e}_l) \right) \mu_F(\mathbf{j} - \mathbf{u}; \mathbf{c}) \\ & - \sum_{\mathbf{0} < \mathbf{u} \le \mathbf{j} - j_l \mathbf{e}_l} \left(\prod_{i=1}^{m} \binom{j_i}{u_i} \right) \mu_H(\mathbf{u}) \mu_F(\mathbf{j} - \mathbf{u}; \mathbf{c}) \Bigg) \\ & - (1 - \pi) c_l \mu_F(\mathbf{j} - \mathbf{e}_l; \mathbf{c}) \end{aligned}$$

$$\begin{aligned} \mu_F(\mathbf{j}) = \pi \Bigg(& \sum_{\mathbf{e}_l \le \mathbf{u} \le \mathbf{j}} \binom{j_l - 1}{u_l - 1} \left(\prod_{i \ne l} \binom{j_i}{u_i} \right) \left(M + 1 - \frac{j_l}{u_l} \right) \mu_H(\mathbf{u}) \mu_F(\mathbf{j} - \mathbf{u}) \\ & - \sum_{\mathbf{0} < \mathbf{u} \le \mathbf{j} - j_l \mathbf{e}_l} \left(\prod_{i=1}^{m} \binom{j_i}{u_i} \right) \mu_H(\mathbf{u}) \mu_F(\mathbf{j} - \mathbf{u}) \Bigg). \end{aligned}$$

3. *Negative binomial distribution* $\mathrm{NB}(\alpha, \pi)$.

$$\begin{aligned} \mu_F(\mathbf{j};\mathbf{c}) = \frac{1}{1 - \pi} \Bigg(& \pi \sum_{\mathbf{0} < \mathbf{u} \le \mathbf{j} - j_l \mathbf{e}_l} \left(\prod_{i=1}^{m} \binom{j_i}{u_i} \right) \mu_H(\mathbf{u}) \mu_F(\mathbf{j} - \mathbf{u}; \mathbf{c}) \\ & + \pi \sum_{\mathbf{e}_l \le \mathbf{u} \le \mathbf{j}} \binom{j_l - 1}{u_l - 1} \left(\prod_{i \ne l} \binom{j_i}{u_i} \right) \\ & \times \left(\left(\frac{j_l}{u_l} + \alpha - 1 \right) \mu_H(\mathbf{u}) + c_l \mu_H(\mathbf{u} - \mathbf{e}_l) \right) \mu_F(\mathbf{j} - \mathbf{u}; \mathbf{c}) \\ & - c_l \mu_F(\mathbf{j} - \mathbf{e}_l; \mathbf{c}) \Bigg) \end{aligned}$$

$$\mu_F(\mathbf{j}) = \frac{\pi}{1 - \pi} \Bigg(\sum_{\mathbf{0} < \mathbf{u} \le \mathbf{j} - j_l \mathbf{e}_l} \left(\prod_{i=1}^{m} \binom{j_i}{u_i} \right) \mu_H(\mathbf{u}) \mu_F(\mathbf{j} - \mathbf{u})$$

$$+\sum_{\mathbf{e}_l \le \mathbf{u} \le \mathbf{j}} \binom{j_l - 1}{u_l - 1}\left(\prod_{i \ne l}\binom{j_i}{u_i}\right)\left(\frac{j_l}{u_l} + \alpha - 1\right)\mu_H(\mathbf{u})\mu_F(\mathbf{j} - \mathbf{u})\Bigg).$$

Further Remarks and References

Section 17.1 is based on Sundt (2003a) and Sect. 17.2 on Sundt (2003b). For more information on the multinormal distribution, see e.g. Johnson et al. (2000).

Chapter 18
Approximations Based on De Pril Transforms

Summary

The purpose of the present chapter is to extend to a multivariate setting the approximations of Chap. 10 and the corresponding error bounds.

In Sect. 18.1, we extend the approximations and the definition of the error measure ε.

Section 18.2 is devoted to error bounds. We extend Theorem 10.4 and indicate that some of its consequences are easily extended. Furthermore, we extend Theorems 10.5–10.7.

18.1 Approximations

For convolutions of compound distributions with counting distribution in $\mathcal{P}_{1\underline{0}}$ and severity distribution in $\mathcal{P}_{m+}$, we shall approximate the counting distribution with a function in $\mathcal{F}_{1\underline{0}}^{(r)}$.

To assess the quality of the approximations, we extend the definition of the error measure ε defined by (10.2) to the multivariate case by

$$\varepsilon(f, \hat{f}) = \mu_{|f-\hat{f}|}(\mathbf{0}) = \sum_{\mathbf{x}\in\mathbb{N}_m} |f(\mathbf{x}) - \hat{f}(\mathbf{x})|. \qquad (f, \hat{f} \in \mathcal{F}_{m\mathbf{0}})$$

We extend the generalised De Pril individual model as defined in Sect. 10.4 by assuming that $h_1, h_2, \ldots, h_I \in \mathcal{P}_{m+}$ and keeping the other assumptions, letting the approximation $f^{(r)}$ still be defined by (10.20).

In the special case when each p_j is the Bernoulli distribution $\mathrm{Bern}(\pi_j)$, we obtain a multivariate version of De Pril's individual model. The approximations of De Pril, Kornya, and Hipp immediately extend to this model. We simply replace the p_js with their approximations. Table 18.1 extends Table 10.1 to the multivariate case.

18.2 Error Bounds

In addition to the error measure ε, we shall also need the distance measure δ given by

$$\delta(f, \hat{f}) = \left|\ln \frac{\hat{f}(\mathbf{0})}{f(\mathbf{0})}\right| + \sum_{\mathbf{x}\in\mathbb{N}_{m+}} \frac{|\varphi_f(\mathbf{x}) - \varphi_{\hat{f}}(\mathbf{x})|}{x_\bullet}, \qquad (f, \hat{f} \in \mathcal{F}_{m\mathbf{0}})$$

that is, we extend the definition (10.6) to the multivariate case.

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_18, © Springer-Verlag Berlin Heidelberg 2009

Table 18.1 Approximations of order r

Approximation	$f^{(r)}(\mathbf{0})$	$\varphi_{f^{(r)}}(\mathbf{x})$
De Pril	$\sum_{j=1}^{J}(1-\pi_j)^{M_{\bullet j}}$	$-x_\bullet \sum_{n=1}^{r}\frac{1}{n}\sum_{j=1}^{J}(\frac{\pi_j}{\pi_j-1})^n \sum_{i=1}^{I} M_{ij}h_i^{n*}(\mathbf{x})$
Kornya	$\exp(\sum_{n=1}^{r}\frac{1}{n}\sum_{j=1}^{J} M_{\bullet j}(\frac{\pi_j}{\pi_j-1})^n)$	$-x_\bullet \sum_{n=1}^{r}\frac{1}{n}\sum_{j=1}^{J}(\frac{\pi_j}{\pi_j-1})^n$ $\times \sum_{i=1}^{I} M_{ij}h_i^{n*}(\mathbf{x})$
Hipp	$\exp(-\sum_{l=1}^{r}\sum_{j=1}^{J} M_{\bullet j}\frac{\pi_j^l}{l})$	$x_\bullet \sum_{l=1}^{r}\frac{1}{l}\sum_{n=1}^{l}(-1)^{n+1}\binom{l}{n}\sum_{j=1}^{J}\pi_j^l$ $\times \sum_{i=1}^{I} M_{ij}h_i^{n*}(\mathbf{x})$

The following multivariate version of Theorem 10.4 can be proved by a trivial modification of the proof of the univariate case.

Theorem 18.1 *If $f \in \mathcal{P}_{m\underline{0}}$ and $\hat{f} \in \mathcal{F}_{m\underline{0}}$, then*

$$\varepsilon(f,\hat{f}) \le e^{\delta(f,\hat{f})} - 1. \tag{18.1}$$

The results in Sect. 10.3.3 are trivially extended to the present multivariate setting.

The following multivariate extensions of Theorems 10.5–10.7 are proved in the same way as in the univariate case.

Theorem 18.2 *If $f_j, \hat{f}_j \in \mathcal{F}_{m\underline{0}}$ for $j = 1, 2, \dots, M$, then*

$$\delta(*_{j=1}^{M} f_j, *_{j=1}^{M} \hat{f}_j) \le \sum_{j=1}^{M} \delta(f_j, \hat{f}_j).$$

Theorem 18.3 *If $p, \hat{p} \in \mathcal{F}_{1\underline{0}}$ and $h \in \mathcal{P}_{m+}$, then*

$$\delta(p \vee h, \hat{p} \vee h) \le \delta(p, \hat{p}); \qquad \varepsilon(p \vee h, \hat{p} \vee h) \le \varepsilon(p, \hat{p}).$$

By application of Theorems 18.2 and 18.3, we easily show that (10.25) holds also for the multivariate version of the generalised De Pril individual model introduced in Sect. 18.1.

The error bounds of Table 10.2 still hold.

Further Remarks and References

This chapter is to a large extent based on Sundt (2000c). Other results on the error measure ε are given there and in Sundt and Vernic (2002).

Walhin and Paris (2001c) considered the De Pril approximation in the bivariate case and presented a numerical application from reinsurance.

Chapter 19
Multivariate Compound Distributions of Type 2

Summary

Till now, our treatment of multivariate compound distributions has been restricted to Type 1. As we have seen, in that setting, it was often rather easy to extend results from the univariate case, and, although the multivariate case might be more awkward to program, the formulae did not look too messy compared to the univariate case. Unfortunately, it is not that simple with Type 2, which is the topic of the present chapter. Although we can utilise some elements from the univariate case, it is not that obvious how to proceed, and the formulae look much more messy.

In Sect. 19.1, we present a setting that could be considered as a Type 2 extension of the framework of Chap. 2, and in Sect. 19.2 we indicate an extension comparable to the extension from Chap. 2 to Chap. 5. Finally, in Sect. 19.3, we briefly indicate how recursions developed for Type 2 can be extended to Type 3.

19.1 Main Framework

Let $f = p \vee \mathbf{h}$ with $p \in \mathcal{P}_{m\underline{\mathbf{0}}}$ and $h_1, h_2, \dots, h_m \in \mathcal{P}_{11}$. It will be convenient to relate these distributions to random variables, so let $\mathbf{N}$ be a random $m \times 1$ vector with distribution p and for $i = 1, 2, \dots$ and $j = 1, 2, \dots, m$, let Y_{ij} be a random variable with distribution h_j. We assume that the Y_{ij}s are mutually independent and independent of $\mathbf{N}$ and introduce $\mathbf{X} = (X_1, X_2, \dots, X_m)'$ with $X_j = \sum_{i=1}^{N_j} Y_{ij}$ for $j = 1, 2, \dots, m$.

Lemma 19.1 *If there exist functions t and v and a subset $\{j_1, j_2, \dots, j_s\}$ of $\{1, 2, \dots, m\}$ such that*

$$\mathrm{E}\left[t(Y_{1j_1}, Y_{1j_2}, \dots, Y_{1j_s}; \mathbf{x}) \,\middle|\, \bigcap_{j=1}^{m} \left(\sum_{i=1}^{n_j} Y_{ij} = x_j\right)\right] = v(\mathbf{n}) \qquad (19.1)$$

for all $\mathbf{x}, \mathbf{n} \in \mathbb{N}_{m+}$ for which $\prod_{j=1}^{m} h_j^{n_j *}(x_j) > 0$, then

$$\sum_{\mathbf{n} \in \mathbb{N}_{m+}} v(\mathbf{n}) p(\mathbf{n} - \mathbf{e}_{j_1 j_2 \dots j_s}) \prod_{j=1}^{m} h_j^{n_j *}(x_j) \qquad (\mathbf{x} \in \mathbb{N}_{m+})$$
$$= \sum_{y_{j_1}=1}^{x_{j_1}} \sum_{y_{j_2}=1}^{x_{j_2}} \cdots \sum_{y_{j_s}=1}^{x_{j_s}} t(y_{j_1}, y_{j_2}, \dots, y_{j_s}; \mathbf{x}) f\left(\mathbf{x} - \sum_{i=1}^{s} y_{j_i} \mathbf{e}_{j_i}\right) \prod_{i=1}^{s} h_{j_i}(y_{j_i}).$$

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_19, © Springer-Verlag Berlin Heidelberg 2009

Proof Extending the set $\{j_1, j_2, \ldots, j_s\}$ to a permutation $\{j_1, j_2, \ldots, j_m\}$ of $\{1, 2, \ldots, m\}$, we obtain that for all $\mathbf{x} \in \mathbb{N}_{m+}$,

$$\sum_{\mathbf{n}\in\mathbb{N}_{m+}} v(\mathbf{n})p(\mathbf{n}-\mathbf{e}_{j_1 j_2\ldots j_s})\prod_{j=1}^{m} h_j^{n_j *}(x_j)$$

$$= \sum_{\mathbf{n}\in\mathbb{N}_{m+}} p(\mathbf{n}-\mathbf{e}_{j_1 j_2\ldots j_s})\,\mathrm{E}\left[t(Y_{1j_1}, Y_{1j_2}, \ldots, Y_{1j_s};\mathbf{x})\middle|\bigcap_{j=1}^{m}\left(\sum_{i=1}^{n_j} Y_{ij}=x_j\right)\right]$$

$$\times\prod_{j=1}^{m} h_j^{n_j *}(x_j)$$

$$= \sum_{\mathbf{n}\in\mathbb{N}_{m+}} p(\mathbf{n}-\mathbf{e}_{j_1 j_2\ldots j_s})\sum_{y_{j_1}=1}^{x_{j_1}}\sum_{y_{j_2}=1}^{x_{j_2}}\cdots\sum_{y_{j_s}=1}^{x_{j_s}} t(y_{j_1}, y_{j_2}, \ldots, y_{j_s};\mathbf{x})$$

$$\times\left(\prod_{i=1}^{s} h_{j_i}(y_{j_i})h_{j_i}^{(n_{j_i}-1)*}(x_{j_i}-y_{j_i})\right)\prod_{i=s+1}^{m} h_{j_i}^{n_{j_i}*}(x_{j_i})$$

$$= \sum_{y_{j_1}=1}^{x_{j_1}}\sum_{y_{j_2}=1}^{x_{j_2}}\cdots\sum_{y_{j_s}=1}^{x_{j_s}} t(y_{j_1}, y_{j_2}, \ldots, y_{j_s};\mathbf{x})\left(\prod_{i=1}^{s} h_{j_i}(y_{j_i})\right)$$

$$\times\sum_{\mathbf{n}\in\mathbb{N}_{m+}} p(\mathbf{n}-\mathbf{e}_{j_1 j_2\ldots j_s})\left(\prod_{i=1}^{s} h_{j_i}^{(n_{j_i}-1)*}(x_{j_i}-y_{j_i})\right)\prod_{i=s+1}^{m} h_{j_i}^{n_{j_i}*}(x_{j_i})$$

$$= \sum_{y_{j_1}=1}^{x_{j_1}}\sum_{y_{j_2}=1}^{x_{j_2}}\cdots\sum_{y_{j_s}=1}^{x_{j_s}} t(y_{j_1}, y_{j_2}, \ldots, y_{j_s};\mathbf{x})\,f\left(\mathbf{x}-\sum_{i=1}^{s} y_{j_i}\mathbf{e}_{j_i}\right)\prod_{i=1}^{s} h_{j_i}(y_{j_i}). \qquad \square$$

The following theorem follows immediately from Lemma 19.1 and (14.2).

Theorem 19.1 *If for any subset* $\{j_1, j_2, \ldots, j_s\}$ *of* $\{1, 2, \ldots, m\}$ *for* $s = 1, 2, \ldots, m$, *the pair* $(t_{j_1 j_2\ldots j_s}, v_{j_1 j_2\ldots j_s})$ *satisfies* (19.1), *then*

$$f(\mathbf{x}) = \sum_{\mathbf{n}\in\mathbb{N}_{m+}}\left(p(\mathbf{n}) - \sum_{s=1}^{m}\sum_{1\le j_1<j_2<\cdots<j_s\le m} v_{j_1 j_2\cdots j_s}(\mathbf{n})p(\mathbf{n}-\mathbf{e}_{j_1 j_2\ldots j_s})\right)$$

$$\times\prod_{j=1}^{m} h_j^{n_j *}(x_j)$$

$$+\sum_{s=1}^{m}\sum_{1\le j_1<j_2<\cdots<j_s\le m}\sum_{y_{j_1}=1}^{x_{j_1}}\sum_{y_{j_2}=1}^{x_{j_2}}\cdots\sum_{y_{j_s}=1}^{x_{j_s}} t_{j_1 j_2\ldots j_s}(y_{j_1}, y_{j_2}, \ldots, y_{j_s};\mathbf{x})$$

$$\times f\left(\mathbf{x}-\sum_{i=1}^{s} y_{j_i}\mathbf{e}_{j_i}\right)\prod_{i=1}^{s} h_{j_i}(y_{j_i}). \qquad (\mathbf{x}\in\mathbb{N}_{m+})$$

In particular, if p satisfies the recursion

$$p(\mathbf{n})=\sum_{s=1}^{m}\sum_{1\le j_1<j_2<\cdots<j_s\le m} v_{j_1 j_2\ldots j_s}(\mathbf{n})p(\mathbf{n}-\mathbf{e}_{j_1 j_2\ldots j_s}), \qquad (\mathbf{n}\in\mathbb{N}_{m+})$$

then we obtain the recursion

$$f(\mathbf{x})=\sum_{s=1}^{m}\sum_{1\le j_1<j_2<\cdots<j_s\le m}\sum_{y_{j_1}=1}^{x_{j_1}}\sum_{y_{j_2}=1}^{x_{j_2}}\cdots\sum_{y_{j_s}=1}^{x_{j_s}} t_{j_1 j_2\ldots j_s}(y_{j_1},y_{j_2},\ldots,y_{j_s};\mathbf{x})$$
$$\times f\left(\mathbf{x}-\sum_{i=1}^{s} y_{j_i}\mathbf{e}_{j_i}\right)\prod_{i=1}^{s} h_{j_i}(y_{j_i}). \qquad (\mathbf{x}\in\mathbb{N}_{m+}) \tag{19.2}$$

Let us now have a look at what sort of pairs (t, v) satisfy (19.1).

When t and v are equal to the same constant, then (19.1) is obviously satisfied.

We easily see that if the couples (t_1, v_1) and (t_2, v_2) satisfy (19.1), then $(c_1t_1 + c_2t_2, c_1v_1 + c_2v_2)$ also satisfies that relation for any constants c_1 and c_2.

As for $j = 1, 2, \ldots, m$, $h_j \in \mathcal{P}_{11}$, we obviously have $X_j = 0$ when $N_j = 0$. This implies that if $\{k_1, k_2, \ldots, k_r\} \subseteq \{j_1, j_2, \ldots, j_s\}$ and (t_1, v_1) satisfies (19.1) when both $\prod_{l=1}^{r} n_{k_l}$ and $\prod_{l=1}^{r} x_{k_l}$ are positive, and (t_0, v_0) satisfies (19.1) when $\prod_{l=1}^{r} x_{k_l} = \prod_{l=1}^{r} n_{k_l} = 0$, then (19.1) holds for (t, v) given by

$$t(y_{j_1},y_{j_2},\ldots,y_{j_s};\mathbf{x})=\begin{cases} t_1(y_{j_1},y_{j_2},\ldots,y_{j_s};\mathbf{x}) & (\prod_{l=1}^{r} x_{k_l}>0)\\ t_0(y_{j_1},y_{j_2},\ldots,y_{j_s};\mathbf{x}) & (\prod_{l=1}^{r} x_{k_l}=0)\end{cases}$$

$$v(\mathbf{n})=\begin{cases} v_1(\mathbf{n}) & (\prod_{l=1}^{r} n_{k_l}>0)\\ v_0(\mathbf{n}). & (\prod_{l=1}^{r} n_{k_l}=0)\end{cases}$$

In particular, we see that when both $\prod_{l=1}^{r} n_{k_l}$ and $\prod_{l=1}^{r} x_{k_l}$ are positive, then (19.1) is satisfied for

$$t(y_{j_1},y_{j_2},\ldots,y_{j_s};\mathbf{x})=\prod_{l=1}^{r}\frac{y_{k_l}}{x_{k_l}}; \qquad v(\mathbf{n})=\frac{1}{\prod_{l=1}^{r} n_{k_l}}.$$

However, we obviously cannot apply this when $\prod_{l=1}^{r} x_{k_l} = \prod_{l=1}^{r} n_{k_l} = 0$.

Using the additivity property of (19.1), we obtain that when both $\prod_{l=1}^{r} n_{k_l}$ and $\prod_{l=1}^{r} x_{k_l}$ are positive, then (19.1) is satisfied for

$$t(y_{j_1},y_{j_2},\ldots,y_{j_s};\mathbf{x})=a+\sum_{u=1}^{r}\sum_{\{z_1,z_2,\ldots,z_u\}\subseteq\{k_1,k_2,\ldots,k_r\}} a_{z_1 z_2\ldots z_u}\prod_{l=1}^{u}\frac{y_{z_l}}{x_{z_l}}$$

$$v(\mathbf{n}) = a + \sum_{u=1}^{r} \sum_{\{z_1,z_2,\ldots,z_u\}\subseteq\{k_1,k_2,\ldots,k_r\}} \frac{a_{z_1 z_2 \ldots z_u}}{\prod_{l=1}^{u} n_{z_l}}.$$

Then we can construct a similar function when, say, $n_{k_w} = 0$ and the other n_{k_l}s are positive, and, so on.

Combining this with (19.2), we obtain the following theorem.

Theorem 19.2 *Let* $\{j_1, j_2, \ldots, j_s\} \subseteq \{1, 2, \ldots, m\}$. *If* p *satisfies the recursion*

$$p(\mathbf{n}) = \sum_{r=1}^{s} \sum_{\{k_1,k_2,\ldots,k_r\}\subseteq\{j_1,j_2,\ldots,j_s\}} \Bigg(a_{k_1 k_2 \ldots k_r} + \sum_{u=1}^{r} \sum_{\{z_1,z_2,\ldots,z_u\}\subseteq\{k_1,k_2,\ldots,k_r\}} \frac{a_{k_1 k_2 \ldots k_r | z_1 z_2 \ldots z_u}}{\prod_{l=1}^{u} n_{z_l}} \Bigg) p(\mathbf{n} - \mathbf{e}_{k_1 k_2 \ldots k_r}) \tag{19.3}$$

for all $\mathbf{n} \in \mathbb{N}_{m+}$ *with* $\prod_{i=1}^{s} n_{j_i} > 0$ *and* $\prod_{i=s+1}^{m} n_{j_i} = 0$, *then*

$$f(\mathbf{x}) = \sum_{r=1}^{s} \sum_{\{k_1,k_2,\ldots,k_r\}\subseteq\{j_1,j_2,\ldots,j_s\}} \sum_{y_{k_1}=1}^{x_{k_1}} \sum_{y_{k_2}=1}^{x_{k_2}} \cdots \sum_{y_{k_r}=1}^{x_{k_r}} \Bigg(a_{k_1 k_2 \ldots k_r} + \sum_{u=1}^{r} \sum_{\{z_1,z_2,\ldots,z_u\}\subseteq\{k_1,k_2,\ldots,k_r\}} a_{k_1 k_2 \ldots k_r | z_1 z_2 \ldots z_u} \prod_{l=1}^{u} \frac{y_{z_l}}{x_{z_l}} \Bigg) \times f\Bigg(\mathbf{x} - \sum_{i=1}^{r} y_{k_i} \mathbf{e}_{k_i}\Bigg) \prod_{i=1}^{r} h_{k_i}(y_{k_i}) \tag{19.4}$$

for all $\mathbf{x} \in \mathbb{N}_{m+}$ *with* $\prod_{i=1}^{s} x_{j_i} > 0$ *and* $\prod_{i=s+1}^{m} x_{j_i} = 0$.

To obtain more transparency, let us restrict to the bivariate case $m = 2$ in the following corollary.

Corollary 19.1 *Let* $p \in \mathcal{P}_{2\underline{0}}$ *and* $h_1, h_2 \in \mathcal{P}_{11}$.

i) *If*

$$p(n_1, n_2) = \left(a_1 + \frac{a_2}{n_1}\right) p(n_1 - 1, n_2) + \left(a_3 + \frac{a_4}{n_2}\right) p(n_1, n_2 - 1) + \left(a_5 + \frac{a_6}{n_1} + \frac{a_7}{n_2} + \frac{a_8}{n_1 n_2}\right) p(n_1 - 1, n_2 - 1),$$

$$(n_1, n_2 = 1, 2, \ldots)$$

then

$$f(x_1,x_2)=\sum_{y_1=1}^{x_1}\left(a_1+a_2\frac{y_1}{x_1}\right)h_1(y_1)f(x_1-y_1,x_2)$$
$$+\sum_{y_2=1}^{x_2}\left(a_3+a_4\frac{y_2}{x_2}\right)h_2(y_2)f(x_1,x_2-y_2)$$
$$+\sum_{y_1=1}^{x_1}\sum_{y_2=1}^{x_2}\left(a_5+a_6\frac{y_1}{x_1}+a_7\frac{y_2}{x_2}+a_8\frac{y_1y_2}{x_1x_2}\right)h_1(y_1)h_2(y_2)$$
$$\times f(x_1-y_1,x_2-y_2). \qquad (x_1,x_2=1,2,\dots)$$

ii) *If*

$$p(n_1,0)=\left(b_1+\frac{b_2}{n_1}\right)p(n_1-1,0), \qquad (n_1=1,2,\dots)$$

then

$$f(x_1,0)=\sum_{y_1=1}^{x_1}\left(b_1+b_2\frac{y_1}{x_1}\right)h_1(y_1)f(x_1-y_1,0). \qquad (x_1=1,2,\dots)$$

iii) *If*

$$p(0,n_2)=\left(c_1+\frac{c_2}{n_2}\right)p(0,n_2-1), \qquad (n_2=1,2,\dots)$$

then

$$f(0,x_2)=\sum_{y_2=1}^{x_2}\left(c_1+c_2\frac{y_2}{x_2}\right)h_2(y_2)f(0,x_2-y_2). \qquad (x_2=1,2,\dots)$$

Let us now give two examples of multivariate distributions that satisfy (19.3).

Example 19.1 Let the distribution q of $N_\bullet$ be $R_1[a,b]$ and the conditional distribution of $\mathbf{N}$ given $N_\bullet$ the multinomial distribution $\mathrm{mnom}(N_\bullet,\boldsymbol{\pi})$. Then p is the compound distribution with counting distribution q and the severity distribution is the multivariate Bernoulli distribution given by (14.3). In Example 15.1, we showed that in this case, p satisfies the recursions (15.8). Hence, it follows from Theorem 19.2 that f satisfies the recursions (15.6). □

Example 19.2 From (16.35), we see that (19.3) is satisfied for multivariate Poisson distributions. Hence, we can apply (19.4) to obtain recursions for compound multivariate Poisson distributions. This will be discussed more thoroughly in Sect. 20.1. □

19.2 Recursions of Higher Order

For the univariate case $m = 1$, we extended the recursion of Theorem 2.4 for compound distributions with counting distribution that satisfies a recursion of order one, to the recursion of Theorem 5.4 for compound distributions with counting distribution that satisfies a recursion of higher order. Analogously, we could extend the recursions that we have deduced for multivariate distributions in the present chapter to allow for counting distributions that satisfy higher order recursions. In that connection, we would need the following extension of Lemma 19.1.

Lemma 19.2 *If there exist functions t and v, a subset $\{j_1, j_2, \dots, j_s\}$ of $\{1, 2, \dots, m\}$, and positive integers $k_1, k_2, \dots, k_s$ such that*

$$\mathrm{E}\left[t\left(\sum_{i=1}^{k_1} Y_{ij_1}, \sum_{i=1}^{k_2} Y_{ij_2}, \dots, \sum_{i=1}^{k_s} Y_{ij_s}; \mathbf{x}\right) \Bigg| \bigcap_{j=1}^{m}\left(\sum_{i=1}^{n_j} Y_{ij} = x_j\right)\right] = v(\mathbf{n})$$

for all $\mathbf{x}, \mathbf{n} \in \mathbb{N}_{m+}$ for which $\prod_{j=1}^{m} h_j^{n_j *}(x_j) > 0$, then

$$\sum_{\mathbf{n}\in\mathbb{N}_{m+}} v(\mathbf{n}) p\left(\mathbf{n} - \sum_{i=1}^{s} k_i \mathbf{e}_{j_i}\right) \prod_{j=1}^{m} h_j^{n_j *}(x_j)$$
$$= \sum_{y_{j_1}=1}^{x_{j_1}} \sum_{y_{j_2}=1}^{x_{j_2}} \cdots \sum_{y_{j_s}=1}^{x_{j_s}} t(y_{j_1}, y_{j_2}, \dots, y_{j_s}; \mathbf{x}) f\left(\mathbf{x} - \sum_{i=1}^{s} y_{j_i} \mathbf{e}_{j_i}\right) \prod_{i=1}^{s} h_{j_i}^{k_i *}(y_{j_i}).$$
$$(\mathbf{x} \in \mathbb{N}_{m+})$$

Theorem 19.2 can be extended correspondingly, although more messy. As a special case, we give the following extension of Corollary 19.1.

Corollary 19.2 *Let $p \in \mathcal{P}_{2\underline{0}}$ and $h_1, h_2 \in \mathcal{P}_{11}$. If*

$$\begin{aligned} p(n_1, n_2) &= \sum_{i_1=1}^{k_1}\left(a_1(i_1) + \frac{a_2(i_1)}{n_1}\right) p(n_1 - i_1, n_2) \\ &\quad + \sum_{i_2=1}^{k_2}\left(a_3(i_2) + \frac{a_4(i_2)}{n_2}\right) p(n_1, n_2 - i_2) \\ &\quad + \sum_{i_1=1}^{k_1}\sum_{i_2=1}^{k_2}\left(a_5(i_1, i_2) + \frac{a_6(i_1, i_2)}{n_1} + \frac{a_7(i_1, i_2)}{n_2} + \frac{a_8(i_1, i_2)}{n_1 n_2}\right) \\ &\quad \times p(n_1 - i_1, n_2 - i_2), \qquad (n_1, n_2 = 1, 2, \dots) \end{aligned} \tag{19.5}$$

then

$$\begin{aligned}
f(x_1,x_2) &= \sum_{y_1=1}^{x_1} f(x_1-y_1,x_2)\sum_{i_1=1}^{k_1}\left(a_1(i_1)+\frac{a_2(i_1)}{i_1}\frac{y_1}{x_1}\right)h_1^{i_1*}(y_1)\\
&\quad+\sum_{y_2=1}^{x_2} f(x_1,x_2-y_2)\sum_{i_2=1}^{k_2}\left(a_3(i_2)+\frac{a_4(i_2)}{i_2}\frac{y_2}{x_2}\right)h_2^{i_2*}(y_2)\\
&\quad+\sum_{y_1=1}^{x_1}\sum_{y_2=1}^{x_2} f(x_1-y_1,x_2-y_2)\\
&\quad\times\sum_{i_1=1}^{k_1}\sum_{i_2=1}^{k_2}\left(a_5(i_1,i_2)+\frac{a_6(i_1,i_2)}{i_1}\frac{y_1}{x_1}+\frac{a_7(i_1,i_2)}{i_2}\frac{y_2}{x_2}\right.\\
&\quad\left.+\frac{a_8(i_1,i_2)}{i_1 i_2}\frac{y_1 y_2}{x_1 x_2}\right)h_1^{i_1*}(y_1)h_2^{i_2*}(y_2). \qquad (x_1,x_2=1,2,\dots)
\end{aligned}$$

19.3 Multivariate Compound Distributions of Type 3

In Chap. 15, we extended the theory of compound univariate distributions to compound distributions with multivariate severity distribution. Analogously, we can extend the theory of the present chapter to the case when the severity distributions $h_1, h_2, \dots, h_m$ are multivariate.

Further Remarks and References

The present chapter is primarily based on Sundt (2000d).

Vernic (1999) proved Corollary 19.1 and discussed some special cases. Hesselager (1996b) discussed Example 19.1 in the bivariate case $m = 2$. He also presented some other recursions for bivariate distributions and compound distributions with bivariate counting distribution and univariate severity distributions.

Vernic (2004) proved Corollary 19.2 and discussed the properties of bivariate distributions that satisfy a recursion of the type (19.5), along the line of what we did for univariate distributions in Sect. 5.3.

Eisele (2008) extended the recursion (5.41) to bivariate counting distributions.

Vernic (1997) deduced a recursion for the bivariate generalised Poisson distribution. She extended this to multivariate distributions in Vernic (2000).

Recursions for compound distributions with bivariate counting distribution and multivariate severity distributions have also been studied by Ambagaspitiya (1998) and Walhin and Paris (2000b, 2001b).

Chapter 20
Compound Mixed Multivariate Poisson Distributions

Summary

The purpose of the present chapter is to extend the theory of Chap. 3 to a class of compound mixed multivariate Poisson distributions.

Multivariate Poisson distributions were introduced in Example 16.1. We give a more extensive presentation in Sect. 20.1. In particular, we emphasise that, whereas a multivariate Poisson distribution is multivariate so that when compounding it, it becomes a compound distribution of Type 2, it can also be expressed as a compound distribution of Type 1, so that we can apply the theory of compound multivariate distributions of Type 1. When extending the class of counting distributions to mixed distributions, we want to restrict the class of mixing distributions in such a way that this aspect is preserved. Such an extension is the topic of Sect. 20.2.

Like in the univariate case, the Gamma mixing distribution is a rather simple case, so we warm up with that in Sect. 20.3.

Then we turn to compound distributions of Type 1 in Sect. 20.4. Like in the univariate case, we first treat a general case in Sect. 20.4.1 before restricting to the Willmot class in Sect. 20.4.2.

Recursions for the univariate counting distribution are deduced in Sect. 20.5.

In a short Sect. 20.6, we describe how the theory of Sect. 20.4 can be used to evaluate compound mixed multivariate Poisson distributions by using the Type 1 representation of such distributions. The special case of the multivariate counting distribution is the topic of Sect. 20.7.

Finally, in Sect. 20.8, we consider some specific parametric classes of mixing distributions within the Willmot class.

20.1 Multivariate Poisson Distributions

Let $\boldsymbol{\lambda}$ be an $s \times 1$ vector of positive numbers, $\mathbf{A}$ an $m \times s$ matrix with elements in $\{0, 1\}$, and $\nu = \lambda_\bullet$. Furthermore, let $f = p \vee \mathbf{h}$ with $p \in \mathcal{P}_{m\underline{\mathbf{0}}}$ and $h_1, h_2, \dots, h_m \in \mathcal{P}_{10}$.

In Example 16.1, we showed that if p is the m-variate Poisson distribution $\mathrm{mPo}(\mathbf{A}, \boldsymbol{\lambda})$, then p can be written as a compound distribution $p = q \vee g$ where the counting distribution q is the Poisson distribution $\mathrm{Po}(\nu)$, that is, $R_1[0, \nu]$, and the severity distribution g the m-variate Bernoulli severity distribution given by (16.34), that is,

$$g(\mathbf{y}) = \frac{\lambda_i}{\nu}. \qquad (\mathbf{y} = \mathbf{a}_i;\, i = 1, 2, \dots, s) \tag{20.1}$$

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0_20, © Springer-Verlag Berlin Heidelberg 2009

Insertion in (15.4) and (15.5) with $a=0$ and $b=\nu$ gives

$$p(\mathbf{n})=\frac{1}{n_l}\sum_{\{i:a_{li}=1\}}\lambda_i p(\mathbf{n}-\mathbf{a}_i) \qquad (\mathbf{n}\geq\mathbf{e}_l;\, l=1,2,\dots,m)$$

$$p(\mathbf{n})=\frac{1}{n_\bullet}\sum_{i=1}^{s}a_{\bullet i}\lambda_i p(\mathbf{n}-\mathbf{a}_i) \qquad (\mathbf{n}\in\mathbb{N}_{m+})$$

with initial condition $p(\mathbf{0})=\mathrm{e}^{-\nu}$; these recursions also follow from (16.35) and (16.32).

We now have $f=q\vee h$ with $h=g\vee\mathbf{h}$. We see that $h\in\mathcal{P}_{m\mathbf{0}}$. By application of (14.2), we obtain

$$h(\mathbf{y})=\frac{1}{\nu}\sum_{i=1}^{s}\lambda_i\prod_{j=1}^{m}h_j^{a_{ji}*}(y_j); \qquad (\mathbf{y}\in\mathbb{N}_m) \tag{20.2}$$

in particular, we have

$$h(\mathbf{0})=\frac{1}{\nu}\sum_{i=1}^{s}\lambda_i\prod_{j=1}^{m}h_j^{a_{ji}}(0). \tag{20.3}$$

Insertion of (20.2) in (15.4) and (15.5) with $a=0$ and $b=\nu$ gives

$$f(\mathbf{x})=\frac{1}{x_l}\sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}}f(\mathbf{x}-\mathbf{y})y_l\sum_{i=1}^{s}\lambda_i\prod_{j=1}^{m}h_j^{a_{ji}*}(y_j) \qquad (\mathbf{x}\geq\mathbf{e}_l;\, l=1,2,\dots,m) \tag{20.4}$$

$$f(\mathbf{x})=\frac{1}{x_\bullet}\sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}}f(\mathbf{x}-\mathbf{y})y_\bullet\sum_{i=1}^{s}\lambda_i\prod_{j=1}^{m}h_j^{a_{ji}*}(y_j) \qquad (\mathbf{x}\in\mathbb{N}_{m+}) \tag{20.5}$$

with initial condition

$$f(\mathbf{0})=\mathrm{e}^{-\nu(1-h(\mathbf{0}))}=\exp\left(-\nu+\sum_{i=1}^{s}\lambda_i\prod_{j=1}^{m}h_j^{a_{ji}}(0)\right).$$

Let us consider the special case studied in Example 16.2, that is,

$$s=m+1; \qquad \mathbf{A}=(\mathbf{I},\mathbf{e}); \qquad \boldsymbol{\lambda}=(\lambda_1,\lambda_2,\dots,\lambda_m,\mu)', \tag{20.6}$$

with some abuse of notation; here $\nu=\mu+\sum_{j=1}^{m}\lambda_j$. We shall refer to these assumptions as *the special design*. Under these assumptions, we obtain

$$h(\mathbf{y})=\frac{1}{\nu}\left(\sum_{j=1}^{m}\lambda_j h_j(y_j)I(\mathbf{y}=y_j\mathbf{e}_j)+\mu\prod_{j=1}^{m}h_j(y_j)\right) \qquad (\mathbf{y}\in\mathbb{N}_m)$$

$$f(\mathbf{x}) = \frac{1}{x_l}\left(\lambda_l \sum_{y=1}^{x_l} y h_l(y) f(\mathbf{x} - y\mathbf{e}_l) + \mu \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} y_l f(\mathbf{x}-\mathbf{y}) \prod_{j=1}^{m} h_j(y_j)\right)$$
$$(\mathbf{x} \geq \mathbf{e}_l;\, l = 1, 2, \ldots, m) \qquad (20.7)$$

$$f(\mathbf{x}) = \frac{1}{x_\bullet}\left(\sum_{i=1}^{m} \lambda_i \sum_{y=1}^{x_i} y h_i(y) f(\mathbf{x} - y\mathbf{e}_i) + \mu \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} y_\bullet f(\mathbf{x}-\mathbf{y}) \prod_{j=1}^{m} h_j(y_j)\right)$$
$$(\mathbf{x} \in \mathbb{N}_{m+}) \qquad (20.8)$$

$$f(\mathbf{0}) = \exp\left(-\nu + \sum_{j=1}^{m} \lambda_j h_j(0) + \mu \prod_{j=1}^{m} h_j(0)\right).$$

When comparing (20.7) and (20.8), we see that in (20.8), we have to sum over i whereas in (20.7), we need the corresponding term only for $i = l$. At first glance, this seems to be an advantage with (20.7). However, in (20.5), with (20.8) as a special case, one would typically do the summation over i in advance or, even better, evaluate the De Pril transform

$$\varphi_f(\mathbf{y}) = \nu y_\bullet h(\mathbf{y}) = y_\bullet \sum_{i=1}^{s} \lambda_i \prod_{j=1}^{m} h_j^{a_{ji}*}(y_j) \qquad (\mathbf{y} \in \mathbb{N}_{m+})$$

and then f by (16.1).

As a special case under the assumptions (20.6), we obtain that the counting distribution p satisfies the recursions

$$p(\mathbf{n}) = \frac{1}{n_l}(\lambda_l p(\mathbf{n} - \mathbf{e}_l) + \mu p(\mathbf{n} - \mathbf{e})) \qquad (\mathbf{n} \geq \mathbf{e}_l;\, l = 1, 2, \ldots, m)$$
$$p(\mathbf{n}) = \frac{1}{n_\bullet}\left(\sum_{i=1}^{m} \lambda_i p(\mathbf{n} - \mathbf{e}_i) + m\mu p(\mathbf{n} - \mathbf{e})\right). \qquad (\mathbf{n} \in \mathbb{N}_{m+})$$

Now let $\mu = 0$. Then f is the distribution of a vector of independent random variables with compound Poisson distribution, and we have $f(\mathbf{x}) = \prod_{j=1}^{m} f_j(x_j)$ for all $\mathbf{x} \in \mathbb{N}_{m+}$. Application of this and (20.7) gives that for all $\mathbf{x} \in \mathbb{N}_{m+}$ and $l = 1, 2, \ldots, m$, we have

$$\prod_{j=1}^{m} f_j(x_j) = f(\mathbf{x}) = \frac{\lambda_l}{x_l} \sum_{y=1}^{x_l} y h_l(y) f(\mathbf{x} - y\mathbf{e}_l)$$
$$= \frac{\lambda_l}{x_l} \sum_{y=1}^{x_l} y h_l(y) f_l(x_l - y) \prod_{j \neq l} f_j(x_j).$$

When $\prod_{j \neq l} f_j(x_j) > 0$, this gives

$$f_l(x_l) = \frac{\lambda_l}{x_l} \sum_{y=1}^{x_l} y h_l(y) f_l(x_l - y),$$

which is the univariate recursion (2.7) for f_l.

20.2 Extension to Mixed Distributions

When deducing the recursions (20.4) and (20.5) for f, we expressed our compound distribution f with m-variate Poisson counting distribution p and univariate severity distributions $h_1, h_2, \ldots, h_m$ as a compound distribution with univariate Poisson counting distribution q and m-variate severity distribution. We then evaluated f by known recursions for compound distributions with univariate Poisson counting distribution and multivariate severity distribution. Could we use such a procedure more generally when p is a mixed multivariate Poisson distribution?

Let $\mathbf{\Theta}$ be a random $s \times 1$ vector with non-negative elements and distribution V. Then a random $m \times 1$ vector $\mathbf{N}$ has a mixed multivariate Poisson distribution if the conditional distribution of $\mathbf{N}$ given $\mathbf{\Theta} = \boldsymbol{\theta}$ has the distribution $\mathrm{mPo}(\mathbf{A}, \boldsymbol{\theta})$ for any value of $\boldsymbol{\theta}$ in the range of $\mathbf{\Theta}$. Let p denote the unconditional distribution of $\mathbf{N}$ and $p_{\boldsymbol{\theta}}$ the conditional distribution given $\mathbf{\Theta} = \boldsymbol{\theta}$ for all values of $\boldsymbol{\theta}$ in the range of $\mathbf{\Theta}$. Proceeding like in Sect. 20.1, we obtain that $p_{\boldsymbol{\theta}}$ can be expressed in the form $p_{\boldsymbol{\theta}} = q_{\boldsymbol{\theta}} \vee g_{\boldsymbol{\theta}}$ where $q_{\boldsymbol{\theta}}$ is the Poisson distribution $\mathrm{Po}(\theta_\bullet)$ and

$$g_{\boldsymbol{\theta}}(\mathbf{y}) = \frac{\theta_i}{\theta_\bullet}, \qquad (\mathbf{y} = \mathbf{a}_i;\, i = 1, 2, \ldots, s)$$

and we can express $f_{\boldsymbol{\theta}} = p_{\boldsymbol{\theta}} \vee \mathbf{h}$ in the form $f_{\boldsymbol{\theta}} = q_{\boldsymbol{\theta}} \vee h_{\boldsymbol{\theta}}$ with $h_{\boldsymbol{\theta}} = g_{\boldsymbol{\theta}} \vee \mathbf{h}$. However, as $h_{\boldsymbol{\theta}}$ depends on $\boldsymbol{\theta}$, we cannot in general express f as a compound distribution with mixed Poisson counting distribution q given by

$$q = \int_{\mathbb{R}_+^s} q_{\boldsymbol{\theta}}\, \mathrm{d}V(\boldsymbol{\theta}). \tag{20.9}$$

Hence, we have to restrict V such that $g_{\boldsymbol{\theta}}$ does not depend on $\boldsymbol{\theta}$, that is, we should have $\theta_i/\theta_\bullet$ independent of $\boldsymbol{\theta}$ for all values of $\boldsymbol{\theta}$ in the range of $\mathbf{\Theta}$ and $i = 1, 2, \ldots, s$. This condition is fulfilled if $\mathbf{\Theta} \equiv \Theta \boldsymbol{\lambda}$ for some non-negative random variable Θ; to avoid the degenerate case that $\Theta = 0$, we assume that $\Theta > 0$. We now get $g_{\boldsymbol{\theta}} = g$ given by (20.1) for all values of $\boldsymbol{\theta}$ in the range of $\mathbf{\Theta}$, so that $f = q \vee h$ with q and h given by (20.9) and (20.2). In particular, this gives

$$p = q \vee g \tag{20.10}$$

with g given by (20.1). Letting U denote the distribution of Θ and q_θ the Poisson distribution $\mathrm{Po}(\theta\nu)$ for $0 < \theta < \infty$, we have

$$q(n) = \int_{(0,\infty)} q_\theta(n)\,\mathrm{d}U(\theta) = \int_{(0,\infty)} \frac{(\theta\nu)^n}{n!}\mathrm{e}^{-\theta\nu}\,\mathrm{d}U(\theta) = \frac{(-\nu)^n}{n!}\gamma_U^{(n)}(\nu).$$
$$(n = 0, 1, 2, \dots) \tag{20.11}$$

In the special case when $s = m$ and $\mathbf{A} = \mathbf{I}$, insertion of (14.5) and (20.1) in (20.10) gives

$$p(\mathbf{n}) = q(n_\bullet)g^{n_\bullet *}(\mathbf{n}) = q(n_\bullet)\frac{n_\bullet!}{\nu^{n_\bullet}}\prod_{j=1}^{m}\frac{\lambda_j^{n_j}}{n_j!}. \qquad (\mathbf{n} \in \mathbb{N}_{m+}) \tag{20.12}$$

As the elements of $\mathbf{N}$ are now mutually dependent through Θ, p is no longer the distribution of a vector of independent random variables unless the mixing distribution U is degenerate.

20.3 Gamma Mixing Distribution

Before continuing with developing the general theory, let us consider the special case where the mixing distribution U is the Gamma distribution $\mathrm{Gamma}(\alpha, \beta)$ with density u given by (3.5). Then, for $n = 1, 2, \dots,$ we have

$$\begin{aligned} q(n) &= \int_0^\infty \frac{(\theta\nu)^n}{n!}\mathrm{e}^{-\theta\nu}\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}\mathrm{e}^{-\beta\theta}\,\mathrm{d}\theta \\ &= \frac{\beta^\alpha\nu^n}{n!\Gamma(\alpha)}\int_0^\infty \theta^{\alpha+n-1}\mathrm{e}^{-(\beta+\nu)\theta}\,\mathrm{d}\theta = \frac{\beta^\alpha\nu^n}{n!\Gamma(\alpha)}\frac{\Gamma(\alpha+n)}{(\beta+\nu)^{\alpha+n}}, \end{aligned}$$

which gives

$$q(n) = \binom{\alpha+n-1}{n}\left(\frac{\nu}{\beta+\nu}\right)^n\left(\frac{\beta}{\beta+\nu}\right)^\alpha. \tag{20.13}$$

This is the negative binomial distribution $\mathrm{NB}(\alpha, \nu/(\beta+\nu))$, that is, $R_1[a, b]$ with

$$a = \frac{\nu}{\beta+\nu}; \qquad b = (\alpha-1)\frac{\nu}{\beta+\nu}.$$

Insertion of this and (20.2) in (15.4) and (15.5) gives

$$f(\mathbf{x}) = \frac{1}{\beta+\nu-\sum_{i=1}^{s}\lambda_i\prod_{j=1}^{m}h_j^{a_{ji}}(0)}$$
$$\times \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}}\left(1+(\alpha-1)\frac{y_l}{x_l}\right)f(\mathbf{x}-\mathbf{y})\sum_{i=1}^{s}\lambda_i\prod_{j=1}^{m}h_j^{a_{ji}*}(y_j)$$
$$(\mathbf{x} \ge \mathbf{e}_l;\, l = 1, 2, \dots, m)$$

$$f(\mathbf{x}) = \frac{1}{\beta + \nu - \sum_{i=1}^{s} \lambda_i \prod_{j=1}^{m} h_j^{a_{ji}}(0)} \times \sum_{\mathbf{0} < \mathbf{y} \le \mathbf{x}} \left(1 + (\alpha - 1)\frac{y_\bullet}{x_\bullet}\right) f(\mathbf{x} - \mathbf{y}) \sum_{i=1}^{s} \lambda_i \prod_{j=1}^{m} h_j^{a_{ji}*}(y_j). \qquad (\mathbf{x} \in \mathbb{N}_{m+})$$

By a trivial extension of the expression in Table 2.3, we obtain

$$f(\mathbf{0}) = \left(\frac{\beta}{\beta + \nu - \sum_{i=1}^{s} \lambda_i \prod_{j=1}^{m} h_j^{a_{ji}}(0)}\right)^{\alpha}.$$

Letting all the h_js be concentrated in one gives

$$p(\mathbf{n}) = \frac{1}{\beta + \nu} \sum_{i=1}^{s} \lambda_i \left(1 + (\alpha - 1)\frac{a_{li}}{n_l}\right) p(\mathbf{n} - \mathbf{a}_i) \qquad (\mathbf{n} \ge \mathbf{e}_l;\, l = 1, 2, \ldots, m)$$

$$p(\mathbf{n}) = \frac{1}{\beta + \nu} \sum_{i=1}^{s} \lambda_i \left(1 + (\alpha - 1)\frac{a_{\bullet i}}{n_\bullet}\right) p(\mathbf{n} - \mathbf{a}_i) \qquad (\mathbf{n} \in \mathbb{N}_{m+})$$

with initial condition

$$p(\mathbf{0}) = \left(\frac{\beta}{\beta + \nu}\right)^{\alpha}.$$

Under the special design (20.6), we have

$$f(\mathbf{x}) = \frac{1}{\beta + \nu - \sum_{j=1}^{m} \lambda_j h_j(0) - \mu \prod_{j=1}^{m} h_j(0)} \times \left(\frac{\lambda_l}{x_l}(\alpha - 1) \sum_{y=1}^{x_l} y h_l(y) f(\mathbf{x} - \mathbf{e}_l y) + \sum_{j=1}^{m} \lambda_j \sum_{y=1}^{x_j} h_j(y) f(\mathbf{x} - \mathbf{e}_j y) + \mu \sum_{\mathbf{0} < \mathbf{y} \le \mathbf{x}} \left(1 + (\alpha - 1)\frac{y_l}{x_l}\right) f(\mathbf{x} - \mathbf{y}) \prod_{j=1}^{m} h_j(y_j)\right)$$

$$(\mathbf{x} \ge \mathbf{e}_l;\, l = 1, 2, \ldots, m)$$

$$f(\mathbf{x}) = \frac{1}{\beta + \nu - \sum_{j=1}^{m} \lambda_j h_j(0) - \mu \prod_{j=1}^{m} h_j(0)} \times \left(\sum_{j=1}^{m} \lambda_j \sum_{y=1}^{x_j} \left(1 + (\alpha - 1)\frac{y}{x_\bullet}\right) h_j(y) f(\mathbf{x} - \mathbf{e}_j y) + \mu \sum_{\mathbf{0} < \mathbf{y} \le \mathbf{x}} \left(1 + (\alpha - 1)\frac{y_\bullet}{x_\bullet}\right) f(\mathbf{x} - \mathbf{y}) \prod_{j=1}^{m} h_j(y_j)\right) \qquad (\mathbf{x} \in \mathbb{N}_{m+})$$

$$f(\mathbf{0}) = \left(\frac{\beta}{\beta + \nu - \sum_{j=1}^{m} \lambda_j h_j(0) - \mu \prod_{j=1}^{m} h_j(0)}\right)^{\alpha}$$

$$p(\mathbf{n}) = \frac{1}{\beta + \nu}\left(\frac{\lambda_l}{n_l}(\alpha - 1)p(\mathbf{n} - \mathbf{e}_l) + \sum_{j=1}^{m} \lambda_j p(\mathbf{n} - \mathbf{e}_j) + \mu\left(1 + \frac{\alpha - 1}{n_l}\right)p(\mathbf{n} - \mathbf{e})\right) \qquad (\mathbf{n} \geq \mathbf{e}_l;\, l = 1, 2, \ldots, m) \tag{20.14}$$

$$p(\mathbf{n}) = \frac{1}{\beta + \nu}\left(\left(1 + \frac{\alpha - 1}{n_{\bullet}}\right)\sum_{j=1}^{m} \lambda_j p(\mathbf{n} - \mathbf{e}_j) + \mu\left(1 + (\alpha - 1)\frac{m}{n_{\bullet}}\right)p(\mathbf{n} - \mathbf{e})\right). \qquad (\mathbf{n} \in \mathbb{N}_{m+})$$

By equating the two expressions for $p(\mathbf{n})$ for $\mathbf{n} \geq \mathbf{e}_l$ and solving for $\sum_{j=1}^{m} \lambda_j p(\mathbf{n} - \mathbf{e}_j)$, we obtain

$$\sum_{j=1}^{m} \lambda_j p(\mathbf{n} - \mathbf{e}_j) = \frac{\lambda_l}{n_l} n_{\bullet} p(\mathbf{n} - \mathbf{e}_l) + \left(\frac{n_{\bullet}}{n_l} - m\right)\mu p(\mathbf{n} - \mathbf{e}),$$

and insertion in (20.14) gives

$$p(\mathbf{n}) = \frac{1}{\beta + \nu}\left(\frac{\lambda_l}{n_l}(\alpha + n_{\bullet} - 1)p(\mathbf{n} - \mathbf{e}_l) + \mu\left(\frac{\alpha + n_{\bullet} - 1}{n_l} - m + 1\right)p(\mathbf{n} - \mathbf{e})\right), \qquad (\mathbf{n} \geq \mathbf{e}_l;\, l = 1, 2, \ldots, m) \tag{20.15}$$

which also follows from (20.65).

In the special case when $\mu = 0$, insertion of (20.13) in (20.12) gives

$$p(\mathbf{n}) = (\alpha + n_{\bullet} - 1)^{(n_{\bullet})}\left(\frac{\beta}{\beta + \nu}\right)^{\alpha} \prod_{j=1}^{m} \frac{1}{n_j!}\left(\frac{\lambda_j}{\beta + \nu}\right)^{n_j}. \qquad (\mathbf{n} \in \mathbb{N}_{m+}) \tag{20.16}$$

This is a *negative multinomial distribution*. In this case, (20.15) reduces to

$$p(\mathbf{n}) = \frac{1}{\beta + \nu}\frac{\lambda_l}{n_l}(\alpha + n_{\bullet} - 1)p(\mathbf{n} - \mathbf{e}_l), \qquad (\mathbf{n} \geq \mathbf{e}_l;\, l = 1, 2, \ldots, m)$$

which also follows immediately from (20.16).

20.4 Compound Distributions with Univariate Counting Distribution

20.4.1 General Recursion

In Sect. 20.4, we shall extend the theory of Sects. 3.3 and 3.5 to the case with multivariate severity distribution $h \in \mathcal{P}_{m\mathbf{0}}$. Let q be the mixed Poisson distribution given by (20.11) with q_θ being the Poisson distribution $\text{Po}(\theta\nu)$ for all $\theta \in (0, \infty)$. We want to evaluate $f = q \vee h$.

Let

$$v_i(\mathbf{x}) = \int_{(0,\infty)} \theta^i f_\theta(\mathbf{x})\,\mathrm{d}U(\theta) \qquad (\mathbf{x} \in \mathbb{N}_m;\ i = 0, 1, 2, \dots) \tag{20.17}$$

with $f_\theta = q_\theta \vee h$. In particular, we have $v_0 = f$. Furthermore, for $i = 0, 1, 2, \dots,$ we have

$$\begin{aligned} v_i(\mathbf{0}) &= \int_{(0,\infty)} \theta^i f_\theta(\mathbf{0})\,\mathrm{d}U(\theta) = \int_{(0,\infty)} \theta^i \mathrm{e}^{-\theta\nu(1-h(\mathbf{0}))}\,\mathrm{d}U(\theta) \\ &= (-1)^i \gamma_U^{(i)}(\nu(1-h(\mathbf{0}))); \qquad (i = 0, 1, 2, \dots) \end{aligned} \tag{20.18}$$

in particular, we get

$$f(\mathbf{0}) = \int_{(0,\infty)} \mathrm{e}^{-\theta\nu(1-h(\mathbf{0}))}\,\mathrm{d}U(\theta) = \gamma_U(\nu(1-h(\mathbf{0}))). \tag{20.19}$$

By letting $a = 0$ and $b = \theta\nu$ in (15.4) and (15.5), we obtain

$$f_\theta(\mathbf{x}) = \frac{\theta\nu}{x_l} \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}} y_l h(\mathbf{y}) f_\theta(\mathbf{x}-\mathbf{y}) \qquad (\mathbf{x} \ge \mathbf{e}_l;\ l = 1, 2, \dots, m) \tag{20.20}$$

$$f_\theta(\mathbf{x}) = \frac{\theta\nu}{x_\bullet} \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}} y_\bullet h(\mathbf{y}) f_\theta(\mathbf{x}-\mathbf{y}), \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{20.21}$$

and multiplication by $\theta^i\,\mathrm{d}U(\theta)$ and integration over θ gives that for $i = 0, 1, 2, \dots,$

$$v_i(\mathbf{x}) = \frac{\nu}{x_l} \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}} y_l h(\mathbf{y}) v_{i+1}(\mathbf{x}-\mathbf{y}) \qquad (\mathbf{x} \ge \mathbf{e}_l;\ l = 1, 2, \dots, m) \tag{20.22}$$

$$v_i(\mathbf{x}) = \frac{\nu}{x_\bullet} \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}} y_\bullet h(\mathbf{y}) v_{i+1}(\mathbf{x}-\mathbf{y}). \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{20.23}$$

We can now evaluate $f(\mathbf{y})$ for all $\mathbf{y} \in \mathbb{N}_m$ such that $\mathbf{0} \le \mathbf{y} \le \mathbf{x}$ by the following algorithm:

Evaluate $f(\mathbf{0})$ by (20.19).
For $i = 1, 2, \dots, x_\bullet$:
  Evaluate $v_i(\mathbf{0})$ by (20.18).
  For $j = 1, 2, \dots, i$:
    For all $\mathbf{y} \in \mathbb{N}_{m+}$ such that $\mathbf{y} \le \mathbf{x}$ and $y_\bullet = j$:
      Evaluate $v_{i-j}(\mathbf{y})$ by (20.22) or (20.23).
  For all $\mathbf{y} \in \mathbb{N}_{m+}$ such that $\mathbf{y} \le \mathbf{x}$ and $y_\bullet = i$:
    Let $f(\mathbf{y}) = v_0(\mathbf{y})$.

The discussion in Sect. 3.4 on finite mixtures also applies to the multivariate case.

20.4.2 Willmot Mixing Distribution

Like in the univariate case, we shall show that if the mixing distribution U is continuous with density u on the interval (γ, δ) with $0 \le \gamma < \delta \le \infty$ that satisfies the condition (3.15), then we need to evaluate v_i only for $i = 0, 1, 2, \dots, k$. We shall need the auxiliary functions

$$w_\theta(\mathbf{x}) = f_\theta(\mathbf{x}) u(\theta) \sum_{i=0}^{k} \chi(i)\theta^i \qquad (\mathbf{x} \in \mathbb{N}_m;\ \gamma < \theta < \delta) \tag{20.24}$$

$$\rho(i) = (1 - h(\mathbf{0}))\nu\chi(i) - \eta(i) - (i+1)\chi(i+1) \qquad (i = -1, 0, 1, \dots, k) \tag{20.25}$$

with $\chi(-1) = \eta(-1) = \chi(k+1) = 0$.

Theorem 20.1 *If $f = q \vee h$ with $h \in \mathcal{P}_{m+}$ and q given by* (20.11) *with continuous mixing distribution U on the interval (γ, δ) with differentiable density u that satisfies* (3.15), *and $w_{\gamma+}(\mathbf{x})$ and $w_{\delta-}(\mathbf{x})$ exist and are finite for all $\mathbf{x} \in \mathbb{N}_{m+}$, then*

$$\rho(k) v_k(\mathbf{x}) = \nu \sum_{\mathbf{0} < \mathbf{y} \le \mathbf{x}} h(\mathbf{y}) \sum_{i=0}^{k} \chi(i) v_i(\mathbf{x} - \mathbf{y}) - \sum_{i=0}^{k-1} \rho(i) v_i(\mathbf{x}) + w_{\gamma+}(\mathbf{x}) - w_{\delta-}(\mathbf{x}). \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{20.26}$$

Proof Application of (20.17), (3.15), and partial integration gives that for all $\mathbf{x} \in \mathbb{N}_{m+}$

$$\sum_{i=0}^{k} \eta(i) v_i(\mathbf{x}) = \sum_{i=0}^{k} \eta(i) \int_\gamma^\delta \theta^i f_\theta(\mathbf{x}) u(\theta)\,\mathrm{d}\theta$$
$$= \sum_{i=0}^{k} \chi(i) \int_\gamma^\delta \theta^i f_\theta(\mathbf{x}) u'(\theta)\,\mathrm{d}\theta$$

$$= w_{\delta-}(\mathbf{x}) - w_{\gamma+}(\mathbf{x}) - \sum_{i=0}^{k} \chi(i) \int_{\gamma}^{\delta} \left(\frac{\mathrm{d}}{\mathrm{d}\theta} \theta^i f_\theta(\mathbf{x}) \right) u(\theta)\, \mathrm{d}\theta$$

$$= w_{\delta-}(\mathbf{x}) - w_{\gamma+}(\mathbf{x})$$

$$- \sum_{i=0}^{k} \chi(i) \int_{\gamma}^{\delta} \left(i\theta^{i-1} f_\theta(\mathbf{x}) + \theta^i \left(\frac{\mathrm{d}}{\mathrm{d}\theta} f_\theta(\mathbf{x}) \right) \right) u(\theta)\, \mathrm{d}\theta,$$

that is,

$$\sum_{i=0}^{k} \eta(i) v_i(\mathbf{x}) = w_{\delta-}(\mathbf{x}) - w_{\gamma+}(\mathbf{x}) - \sum_{i=1}^{k} i\chi(i) v_{i-1}(\mathbf{x})$$

$$- \sum_{i=0}^{k} \chi(i) \int_{\gamma}^{\delta} \theta^i \left(\frac{\mathrm{d}}{\mathrm{d}\theta} f_\theta(\mathbf{x}) \right) u(\theta)\, \mathrm{d}\theta. \tag{20.27}$$

We have

$$\frac{\mathrm{d}}{\mathrm{d}\theta} f_\theta(\mathbf{x}) = \sum_{n=1}^{\infty} h^{n*}(\mathbf{x}) \frac{\mathrm{d}}{\mathrm{d}\theta} \frac{(\theta\nu)^n}{n!} \mathrm{e}^{-\theta\nu}$$

$$= \sum_{n=1}^{\infty} h^{n*}(\mathbf{x}) \frac{\nu^n}{n!} (n\theta^{n-1} \mathrm{e}^{-\theta\nu} - \nu\theta^n \mathrm{e}^{-\theta\nu})$$

$$= \nu \sum_{n=1}^{\infty} h^{n*}(\mathbf{x})(q_\theta(n-1) - q_\theta(n)) = \nu((f_\theta * h)(\mathbf{x}) - f_\theta(\mathbf{x}))$$

$$= \nu \left(\sum_{\mathbf{0} \le \mathbf{y} \le \mathbf{x}} h(\mathbf{y}) f_\theta(\mathbf{x} - \mathbf{y}) - f_\theta(\mathbf{x}) \right)$$

$$= \nu \left(\sum_{\mathbf{0} < \mathbf{y} \le \mathbf{x}} h(\mathbf{y}) f_\theta(\mathbf{x} - \mathbf{y}) - (1 - h(\mathbf{0})) f_\theta(\mathbf{x}) \right),$$

and insertion in (20.27) gives

$$\sum_{i=0}^{k} \eta(i) v_i(\mathbf{x}) = w_{\delta-}(\mathbf{x}) - w_{\gamma+}(\mathbf{x}) - \sum_{i=0}^{k-1} (i+1)\chi(i+1) v_i(\mathbf{x})$$

$$- \nu \sum_{i=0}^{k} \chi(i) \left(\sum_{\mathbf{0} < \mathbf{y} \le \mathbf{x}} h(\mathbf{y}) v_i(\mathbf{x} - \mathbf{y}) - (1 - h(\mathbf{0})) v_i(\mathbf{x}) \right).$$

After some rearranging, we obtain

$$\sum_{i=0}^{k}\rho(i)v_i(\mathbf{x}) = \nu\sum_{i=0}^{k}\chi(i)\sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}} h(\mathbf{y})v_i(\mathbf{x}-\mathbf{y}) + w_{\gamma+}(\mathbf{x}) - w_{\delta-}(\mathbf{x}),$$

from which (20.26) follows. □

If $\rho(k)\neq 0$, then (20.26) gives

$$v_k(\mathbf{x}) = \frac{1}{\rho(k)}\Bigg(\nu\sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}} h(\mathbf{y})\sum_{i=0}^{k}\chi(i)v_i(\mathbf{x}-\mathbf{y}) - \sum_{i=0}^{k-1}\rho(i)v_i(\mathbf{x}) + w_{\gamma+}(\mathbf{x}) - w_{\delta-}(\mathbf{x})\Bigg). \qquad (\mathbf{x}\in\mathbb{N}_{m+}) \qquad (20.28)$$

In this case, we can evaluate $f(\mathbf{y})$ for all $\mathbf{y}\in\mathbb{N}_m$ such that $\mathbf{0}\le\mathbf{y}\le\mathbf{x}$ for some $\mathbf{x}\in\mathbb{N}_{m+}$ such that $x_\bullet > k$ by the following algorithm:

Evaluate $f(\mathbf{0})$ by (20.19).
For $i = 1, 2, \dots, k$:
  Evaluate $v_i(\mathbf{0})$ by (20.18).
  For $j = 1, 2, \dots, i$:
    For all $\mathbf{y}\in\mathbb{N}_{m+}$ such that $\mathbf{y}\le\mathbf{x}$ and $y_\bullet = j$:
      Evaluate $v_{i-j}(\mathbf{y})$ by (20.22) or (20.23).
  For all $\mathbf{y}\in\mathbb{N}_{m+}$ such that $\mathbf{y}\le\mathbf{x}$ and $y_\bullet = i$:
    Let $f(\mathbf{y}) = v_0(\mathbf{y})$.
For $i = k+1, k+2, \dots, x_\bullet$:
  For all $\mathbf{y}\in\mathbb{N}_{m+}$ such that $\mathbf{y}\le\mathbf{x}$ and $y_\bullet = i-k$:
    Evaluate $v_k(\mathbf{y})$ by (20.28).
  For $j = 1, 2, \dots, k$:
    For all $\mathbf{y}\in\mathbb{N}_{m+}$ such that $\mathbf{y}\le\mathbf{x}$ and $y_\bullet = i-k+j$:
      Evaluate $v_{k-j}(\mathbf{y})$ by (20.22) or (20.23).
  For all $\mathbf{y}\in\mathbb{N}_{m+}$ such that $\mathbf{y}\le\mathbf{x}$ and $y_\bullet = i$:
  Let $f(\mathbf{y}) = v_0(\mathbf{y})$.

Let us now consider the condition that $w_{\gamma+}(\mathbf{x})$ and $w_{\delta-}(\mathbf{x})$ should exist and be finite for all $\mathbf{x}\in\mathbb{N}_{m+}$. For finite γ and δ, this condition holds when $u(\gamma+)$ and $u(\delta-)$ exist and are finite. In particular, we have

$$w_{0+}(\mathbf{x}) = f_0(\mathbf{x})u(0+)\chi(0) = 0 \qquad (\mathbf{x}\in\mathbb{N}_{m+})$$

as the distribution f_0 is concentrated in zero.

For $\theta \in (\gamma, \delta)$, w_θ is proportional to f_θ, and (20.20) and (20.21) give the recursions

$$w_\theta(\mathbf{x}) = \frac{\theta\nu}{x_l} \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} y_l h(\mathbf{y}) w_\theta(\mathbf{x}-\mathbf{y}) \qquad (\mathbf{x} \geq \mathbf{e}_l;\, l = 1, 2, \ldots, m) \tag{20.29}$$

$$w_\theta(\mathbf{x}) = \frac{\theta\nu}{x_\bullet} \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} y_\bullet h(\mathbf{y}) w_\theta(\mathbf{x}-\mathbf{y}), \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \tag{20.30}$$

and from (20.24), we obtain the initial condition

$$w_\theta(\mathbf{0}) = \mathrm{e}^{-\theta\nu(1-h(\mathbf{0}))} u(\theta) \sum_{i=0}^{k} \chi(i)\theta^i. \tag{20.31}$$

If $\gamma > 0$ and $u(\gamma+)$ exists and is finite, then we can evaluate $w_{\gamma+}$ recursively in the same way. Analogous for $w_{\delta-}$ when $\delta < \infty$ and $u(\delta-)$ exists and is finite.

20.5 The Univariate Mixed Counting Distribution

For evaluation of the univariate mixed counting distribution, we have to slightly extend the theory of Sect. 3.6 as there we had $\nu = 1$.

We introduce

$$\begin{aligned} \dot{v}_i(n) &= \int_{(0,\infty)} \theta^i q_\theta(n)\,\mathrm{d}U(\theta) = \frac{\nu^n}{n!} \int_{(0,\infty)} \theta^{n+i} \mathrm{e}^{-\theta\nu}\,\mathrm{d}U(\theta) \\ &= \frac{(n+i)^{(i)}}{\nu^i} q(i+n) \qquad (n, i = 0, 1, 2, \ldots) \end{aligned} \tag{20.32}$$

$$\dot{w}_\theta(n) = q_\theta(n) u(\theta) \sum_{i=0}^{k} \chi(i)\theta^i = \frac{(\theta\nu)^n}{n!} \mathrm{e}^{-\theta\nu} u(\theta) \sum_{i=0}^{k} \chi(i)\theta^i$$
$$(n = 0, 1, \ldots;\, \gamma < \theta < \delta) \tag{20.33}$$

$$\dot{\rho}(i) = \nu\chi(i) - \eta(i) - (i+1)\chi(i+1). \qquad (i = -1, 0, 1, \ldots, k) \tag{20.34}$$

The following corollary to Theorem 3.3 extends that theorem to general ν.

Corollary 20.1 *If q is the mixed Poisson distribution given by* (20.11) *with continuous mixing distribution U on the interval (γ, δ) with differentiable density u that satisfies* (3.15), *and $\dot{w}_{\gamma+}(n)$ and $\dot{w}_{\delta-}(n)$ exist and are finite for all non-negative integers n, then*

$$\begin{aligned} \dot{\rho}(k) q(n) &= \sum_{i=1}^{k+1} ((n-k)\chi(k-i+1) - \dot{\rho}(k-i)) \frac{\nu^i}{n^{(i)}} q(n-i) \\ &\quad + (\dot{w}_{\gamma+}(n-k) - \dot{w}_{\delta-}(n-k)) \frac{\nu^k}{n^{(k)}}. \quad (n = k+1, k+2, \ldots) \end{aligned} \tag{20.35}$$

Proof From the discussion in Sect. 3.7.2, we have that q can be expressed in the form

$$q(n) = \int_{\tilde{\gamma}}^{\tilde{\delta}} \frac{\theta^n}{n!} \mathrm{e}^{-\theta} \tilde{u}(\theta)\, \mathrm{d}\theta \qquad (n = 0, 1, 2, \dots; \tilde{\gamma} < \theta < \tilde{\delta})$$

with $\tilde{\gamma} = \nu\gamma$, $\tilde{\delta} = \nu\delta$ and the mixing density $\tilde{u}$ satisfying

$$\frac{\mathrm{d}}{\mathrm{d}\theta} \ln \tilde{u}(\theta) = \frac{\sum_{i=0}^{k} \tilde{\eta}(i)\theta^i}{\sum_{i=0}^{k} \tilde{\chi}(i)\theta^i} \qquad (\tilde{\gamma} < \theta < \tilde{\delta})$$

with $\tilde{\eta}(i) = \eta(i)\nu^{-i}$ and $\tilde{\chi}(i) = \chi(i)\nu^{1-i}$ for $i = 0, 1, 2, \dots, k$. From Theorem 3.3, we obtain

$$\tilde{\dot{\rho}}(k)q(n) = \sum_{i=1}^{k+1} ((n-k)\tilde{\chi}(k-i+1) - \tilde{\dot{\rho}}(k-i)) \frac{q(n-i)}{n^{(i)}} + \frac{\tilde{\dot{w}}_{\tilde{\gamma}+}(n-k) - \tilde{\dot{w}}_{\tilde{\delta}-}(n-k)}{n^{(k)}} \qquad (n = k+1, k+2, \dots) \tag{20.36}$$

with

$$\tilde{\dot{\rho}}(i) = \tilde{\chi}(i) - \tilde{\eta}(i) - (i+1)\tilde{\chi}(i+1) = \frac{\dot{\rho}(i)}{\nu^i}. \qquad (i = -1, 0, 1, 2, \dots, k)$$

$$\tilde{\dot{w}}_\theta(n) = \frac{\theta^n}{n!} \mathrm{e}^{-\theta} \tilde{u}(\theta) \sum_{i=0}^{k} \tilde{\chi}(i)\theta^i = \dot{w}_{\theta/\nu}(n). \qquad (n = 0, 1, \dots; \tilde{\gamma} < \theta < \tilde{\delta})$$

Insertion of the expressions for $\tilde{\gamma}$, $\tilde{\delta}$, $\tilde{\dot{\rho}}$, $\tilde{\chi}$, and $\tilde{\dot{w}}_\theta$ in (20.36) gives

$$\frac{\dot{\rho}(k)}{\nu^k} q(n) = \sum_{i=1}^{k+1} \left((n-k) \frac{\chi(k-i+1)}{\nu^{k-i}} - \frac{\dot{\rho}(k-i)}{\nu^{k-i}} \right) \frac{q(n-i)}{n^{(i)}} + \frac{\dot{w}_{\gamma+}(n-k) - \dot{w}_{\delta-}(n-k)}{n^{(k)}}, \qquad (n = k+1, k+2, \dots)$$

and by multiplication by ν^k, we obtain (20.35). □

When $\dot{\rho}(k) \neq 0$, (20.35) gives

$$q(n) = \frac{1}{\dot{\rho}(k)} \left(\sum_{i=1}^{k+1} ((n-k)\chi(k-i+1) - \dot{\rho}(k-i)) \frac{\nu^i}{n^{(i)}} q(n-i) + (\dot{w}_{\gamma+}(n-k) - \dot{w}_{\delta-}(n-k)) \frac{\nu^k}{n^{(k)}} \right). \qquad (n = k+1, k+2, \dots)$$

Let us consider the special case $k=1$. Then (20.35) gives

$$q(n)=\frac{1}{\dot\rho(1)}\biggl(b_{\gamma+}(n)-b_{\delta-}(n)+\biggl(\chi(1)-\frac{\chi(1)+\dot\rho(0)}{n}\biggr)\nu q(n-1)\\ +\frac{\chi(0)}{n}\nu^2 q(n-2)\biggr)\qquad (n=2,3,\dots) \tag{20.37}$$

with

$$b_\theta(n)=\frac{\dot w_\theta(n-1)}{n}\nu=\frac{\dot w_\theta(n)}{\theta}.\qquad (n=1,2,\dots;\ \gamma<\theta<\delta) \tag{20.38}$$

By application of (15.25) and (15.26), we obtain

$$f(\mathbf{x})=\frac{1}{\dot\rho(1)-\chi(1)\nu h(\mathbf{0})}\biggl((\dot\rho(1)q(1)+\dot\rho(0)\nu q(0)+b_{\delta-}(1)-b_{\gamma+}(1))h(\mathbf{x})\\ +(b_{\gamma+}\vee h)(\mathbf{x})-(b_{\delta-}\vee h)(\mathbf{x})\\ +\nu\sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}}\biggl(\biggl(\chi(1)-(\chi(1)+\dot\rho(0))\frac{y_l}{x_l}\biggr)h(\mathbf{y})+\nu\frac{\chi(0)}{2}\frac{y_l}{x_l}h^{2*}(\mathbf{y})\biggr)f(\mathbf{x}-\mathbf{y})\biggr)\\ (\mathbf{x}\ge\mathbf{e}_l;\ l=1,2,\dots,m) \tag{20.39}$$

$$f(\mathbf{x})=\frac{1}{\dot\rho(1)-\chi(1)\nu h(\mathbf{0})}\biggl((\dot\rho(1)q(1)+\dot\rho(0)\nu q(0)+b_{\delta-}(1)-b_{\gamma+}(1))h(\mathbf{x})\\ +(b_{\gamma+}\vee h)(\mathbf{x})-(b_{\delta-}\vee h)(\mathbf{x})\\ +\nu\sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}}\biggl(\biggl(\chi(1)-(\chi(1)+\dot\rho(0))\frac{y_\bullet}{x_\bullet}\biggr)h(\mathbf{y})+\nu\frac{\chi(0)}{2}\frac{y_\bullet}{x_\bullet}h^{2*}(\mathbf{y})\biggr)\\ \times f(\mathbf{x}-\mathbf{y})\biggr).\qquad (\mathbf{x}\in\mathbb{N}_{m+}) \tag{20.40}$$

We can evaluate $b_\theta\vee h$ recursively by

$$(b_\theta\vee h)(\mathbf{x})=\frac{\theta\nu}{x_l}\sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}}y_l h(\mathbf{y})(b_\theta\vee h)(\mathbf{x}-\mathbf{y})\qquad (\mathbf{x}\ge\mathbf{e}_l;\ l=1,2,\dots,m) \tag{20.41}$$

$$(b_\theta\vee h)(\mathbf{x})=\frac{\theta\nu}{x_\bullet}\sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}}y_\bullet h(\mathbf{y})(b_\theta\vee h)(\mathbf{x}-\mathbf{y}),\qquad (\mathbf{x}\in\mathbb{N}_{m+}) \tag{20.42}$$

and by (20.33) and (20.38), we obtain the initial condition

$$(b_\theta\vee h)(\mathbf{0})=\biggl(\frac{\chi(0)}{\theta}+\chi(1)\biggr)u(\theta)\mathrm{e}^{-\theta\nu(1-h(\mathbf{0}))}. \tag{20.43}$$

When $k = 1$, these recursions for f are more efficient than the algorithm given after formula (20.28).

20.6 Compound Distributions with Multivariate Counting Distribution

Let us now return to the setting of Sect. 20.2. In that setting, we had that the compound mixed multivariate Poisson distribution $f = p \vee \mathbf{h}$ could also be represented in the form $f = q \vee h$ with $h \in \mathcal{P}_{m\mathbf{0}}$ given by (20.2). Hence, we can evaluate f by the recursions of Sect. 20.4 with h given by (20.2). As one would typically evaluate this h in advance, we will not display the formulae obtained by insertion of (20.2).

20.7 The Multivariate Counting Distribution

20.7.1 General Design

For evaluation of the multivariate counting distribution p by the recursions for f, we replace v_i and w_θ with $\ddot{v}_i$ and $\ddot{w}_\theta$ given by

$$\ddot{v}_i(\mathbf{n}) = \int_{(0,\infty)} \theta^i p_\theta(\mathbf{n})\,\mathrm{d}U(\theta) \qquad (\mathbf{n} \in \mathbb{N}_m;\ i = 0, 1, 2, \dots)$$

$$\ddot{w}_\theta(\mathbf{n}) = p_\theta(\mathbf{n})u(\theta)\sum_{i=0}^{k}\chi(i)\theta^i \qquad (\mathbf{n} \in \mathbb{N}_m;\ \gamma < \theta < \delta) \tag{20.44}$$

whereas ρ is replaced with $\dot{\rho}$ given by (20.34). Insertion of (20.1) in (20.22) and (20.23) gives that for $i = 0, 1, 2, \dots,$

$$\ddot{v}_i(\mathbf{n}) = \frac{1}{n_l}\sum_{\{t:a_{lt}=1\}} \lambda_t \ddot{v}_{i+1}(\mathbf{n} - \mathbf{a}_t) \qquad (\mathbf{n} \geq \mathbf{e}_l;\ l = 1, 2, \dots, m) \tag{20.45}$$

$$\ddot{v}_i(\mathbf{n}) = \frac{1}{n_\bullet}\sum_{t=1}^{s} \lambda_t a_{\bullet t} \ddot{v}_{i+1}(\mathbf{n} - \mathbf{a}_t) \qquad (\mathbf{n} \in \mathbb{N}_{m+}) \tag{20.46}$$

with initial condition

$$\ddot{v}_i(\mathbf{0}) = \int_{(0,\infty)} \theta^i \mathrm{e}^{-\theta\nu}\,\mathrm{d}U(\theta) = (-1)^i \gamma_U^{(i)}(\nu).$$

In particular, we have

$$p(\mathbf{0}) = \ddot{v}_0(\mathbf{0}) = \int_{(0,\infty)} \mathrm{e}^{-\theta\nu}\,\mathrm{d}U(\theta) = \gamma_U(\nu).$$

If the mixing distribution U satisfies the conditions of Theorem 20.1, then insertion of (20.1) in (20.26) and (20.29)–(20.31) gives

$$\dot{\rho}(k)\ddot{v}_k(\mathbf{n}) = \sum_{i=0}^{k}\chi(i)\sum_{t=1}^{s}\lambda_t\ddot{v}_i(\mathbf{n}-\mathbf{a}_t) - \sum_{i=0}^{k-1}\dot{\rho}(i)\ddot{v}_i(\mathbf{n}) + \ddot{w}_{\gamma+}(\mathbf{n}) - \ddot{w}_{\delta-}(\mathbf{n}) \qquad (\mathbf{n}\in\mathbb{N}_{m+}) \tag{20.47}$$

$$\ddot{w}_\theta(\mathbf{n}) = \frac{\theta}{n_l}\sum_{\{t:a_{lt}=1\}}\lambda_t\ddot{w}_\theta(\mathbf{n}-\mathbf{a}_t) \qquad (\mathbf{n}\geq\mathbf{e}_l;\, l=1,2,\dots,m) \tag{20.48}$$

$$\ddot{w}_\theta(\mathbf{n}) = \frac{\theta}{n_\bullet}\sum_{t=1}^{s}\lambda_t a_{\bullet t}\ddot{w}_\theta(\mathbf{n}-\mathbf{a}_t) \qquad (\mathbf{n}\in\mathbb{N}_{m+}) \tag{20.49}$$

$$\ddot{w}_\theta(\mathbf{0}) = \mathrm{e}^{-\theta\nu}u(\theta)\sum_{i=0}^{k}\chi(i)\theta^i.$$

20.7.2 The Special Design

In Sect. 20.7.2, we shall concentrate on the special design (20.6). Then (20.45) and (20.46) give that for $i = 1, 2, \dots,$

$$\ddot{v}_i(\mathbf{n}) = \frac{\lambda_l}{n_l}\ddot{v}_{i+1}(\mathbf{n}-\mathbf{e}_l) + \frac{\mu}{n_l}\ddot{v}_{i+1}(\mathbf{n}-\mathbf{e}) \quad (\mathbf{n}\geq\mathbf{e}_l;\, l=1,2,\dots,m) \tag{20.50}$$

$$\ddot{v}_i(\mathbf{n}) = \frac{1}{n_\bullet}\left(\sum_{l=1}^{m}\lambda_l\ddot{v}_{i+1}(\mathbf{n}-\mathbf{e}_l) + m\mu\ddot{v}_{i+1}(\mathbf{n}-\mathbf{e})\right), \qquad (\mathbf{n}\in\mathbb{N}_{m+})$$

and if the mixing distribution U satisfies the conditions of Theorem 20.1, then (20.47)–(20.49) give

$$\dot{\rho}(k)\ddot{v}_k(\mathbf{n}) = \sum_{i=0}^{k}\chi(i)\left(\sum_{j=1}^{m}\lambda_j\ddot{v}_i(\mathbf{n}-\mathbf{e}_j) + \mu\ddot{v}_i(\mathbf{n}-\mathbf{e})\right) - \sum_{i=0}^{k-1}\dot{\rho}(i)\ddot{v}_i(\mathbf{n}) + \ddot{w}_{\gamma+}(\mathbf{n}) - \ddot{w}_{\delta-}(\mathbf{n}) \quad (\mathbf{n}\in\mathbb{N}_{m+}) \tag{20.51}$$

$$\ddot{w}_\theta(\mathbf{n}) = \frac{\theta}{n_l}(\lambda_l\ddot{w}_\theta(\mathbf{n}-\mathbf{e}_l) + \mu\ddot{w}_\theta(\mathbf{n}-\mathbf{e})) \qquad (\mathbf{n}\geq\mathbf{e}_l;\, l=1,2,\dots,m) \tag{20.52}$$

$$\ddot{w}_\theta(\mathbf{n}) = \frac{\theta}{n_\bullet}\left(\sum_{l=1}^{m}\lambda_l\ddot{w}_\theta(\mathbf{n}-\mathbf{e}_l) + m\mu\ddot{w}_\theta(\mathbf{n}-\mathbf{e})\right). \qquad (\mathbf{n}\in\mathbb{N}_{m+})$$

From (20.50), we get

$$n_l \ddot{v}_i(\mathbf{n}) = \lambda_l \ddot{v}_{i+1}(\mathbf{n} - \mathbf{e}_l) + \mu \ddot{v}_{i+1}(\mathbf{n} - \mathbf{e}), \qquad (20.53)$$
$$(l = 1, 2, \ldots, m;\ i = 0, 1, 2, \ldots)$$

which holds for all $\mathbf{n} \in \mathbb{Z}_m$. As this also goes for many of the other formulae in the following, we shall often drop giving the range explicitly. By summation over l and some manipulation, we obtain

$$\sum_{l=1}^{m} \lambda_l \ddot{v}_i(\mathbf{n} - \mathbf{e}_l) = n_\bullet \ddot{v}_{i-1}(\mathbf{n}) - m\mu \ddot{v}_i(\mathbf{n} - \mathbf{e}). \qquad (i = 1, 2, \ldots)$$

Insertion in (20.51) gives

$$\dot{\rho}(k)\ddot{v}_k(\mathbf{n}) = \chi(0)\left(\sum_{j=1}^{m} \lambda_j p(\mathbf{n} - \mathbf{e}_j) + \mu p(\mathbf{n} - \mathbf{e})\right)$$
$$+ \sum_{i=1}^{k} \chi(i)(n_\bullet \ddot{v}_{i-1}(\mathbf{n}) - (m-1)\mu \ddot{v}_i(\mathbf{n} - \mathbf{e}))$$
$$- \sum_{i=0}^{k-1} \dot{\rho}(i)\ddot{v}_i(\mathbf{n}) + \ddot{w}_{\gamma+}(\mathbf{n}) - \ddot{w}_{\delta-}(\mathbf{n}). \qquad (\mathbf{n} \in \mathbb{N}_{m+}) \qquad (20.54)$$

To get rid of the summation over j, we multiply the numerator and denominator in (3.15) by θ, so that for $\gamma < \theta < \delta$

$$\frac{\mathrm{d}}{\mathrm{d}\theta} \ln u(\theta) = \frac{\sum_{i=0}^{k} \eta(i)\theta^{i+1}}{\sum_{i=0}^{k} \chi(i)\theta^{i+1}} = \frac{\sum_{i=1}^{k+1} \eta(i-1)\theta^{i}}{\sum_{i=1}^{k+1} \chi(i-1)\theta^{i}} = \frac{\sum_{i=0}^{k^*} \eta^*(i)\theta^{i}}{\sum_{i=0}^{k^*} \chi^*(i)\theta^{i}}$$

with $k^* = k + 1$ and

$$\eta^*(i) = \eta(i-1); \qquad \chi^*(i) = \chi(i-1) \qquad (i = 1, 2, \ldots, k^*)$$
$$\eta^*(0) = \chi^*(0) = 0.$$

Correspondingly, we introduce

$$\dot{\rho}^*(i) = \dot{\rho}(i-1) - \chi(i) \qquad (i = -1, 0, 1, 2, \ldots, k^*)$$
$$\ddot{w}^*_\theta = \theta \ddot{w}_\theta. \qquad (\gamma < \theta < \delta)$$

From (20.54), we obtain

$$\dot{\rho}^*(k^*)\ddot{v}_{k^*}(\mathbf{n}) = \sum_{i=0}^{k^*} \chi^*(i)(n_\bullet \ddot{v}_{i-1}(\mathbf{n}) - (m-1)\mu \ddot{v}_i(\mathbf{n} - \mathbf{e}))$$
$$- \sum_{i=0}^{k^*-1} \dot{\rho}^*(i)\ddot{v}_i(\mathbf{n}) + \ddot{w}^*_{\gamma+}(\mathbf{n}) - \ddot{w}^*_{\delta-}(\mathbf{n}).$$

Insertion of the expressions for k^*, $\dot{\rho}^*$, χ^*, and $\ddot{w}^*_\theta$ and some manipulation give

$$\dot{\rho}(k)\ddot{v}_{k+1}(\mathbf{n}) = \sum_{i=0}^{k}(((n_\bullet + 1)\chi(i) - \dot{\rho}(i-1))\ddot{v}_i(\mathbf{n}) - (m-1)\chi(i)\mu\ddot{v}_{i+1}(\mathbf{n}-\mathbf{e})) + \ddot{w}^*_{\gamma+}(\mathbf{n}) - \ddot{w}^*_{\delta-}(\mathbf{n}).$$

We shall need the operator Ξ_l defined by

$$\Xi_l c(\mathbf{n}) = \lambda_l c(\mathbf{n} - \mathbf{e}_l) + \mu c(\mathbf{n} - \mathbf{e})$$

where c is a function on $\mathbb{Z}_m$.

Lemma 20.1 *For any function c on $\mathbb{Z}_m$ and $l = 1, 2, \ldots, m$, we have*

$$\Xi_l^i c(\mathbf{n}) = \sum_{z=0}^{i}\binom{i}{z}\mu^z\lambda_l^{i-z}c(\mathbf{n} - (i-z)\mathbf{e}_l - z\mathbf{e}). \quad (i = 0, 1, 2, \ldots) \tag{20.55}$$

Proof For $i = 0$, (20.55) obviously holds. Let us now assume that it holds for i equal to some non-negative integer j. Then

$$\begin{aligned}
\Xi_l^{j+1} c(\mathbf{n}) &= \lambda_l \Xi_l^j c(\mathbf{n} - \mathbf{e}_l) + \mu \Xi_l^j c(\mathbf{n} - \mathbf{e}) \\
&= \sum_{z=0}^{j}\binom{j}{z}\mu^z\lambda_l^{j-z+1}c(\mathbf{n} - (j-z+1)\mathbf{e}_l - z\mathbf{e}) \\
&\quad + \sum_{z=0}^{j}\binom{j}{z}\mu^{z+1}\lambda_l^{j-z}c(\mathbf{n} - (j-z)\mathbf{e}_l - (z+1)\mathbf{e}) \\
&= \lambda_l^{j+1}c(\mathbf{n} - (j+1)\mathbf{e}_l) \\
&\quad + \sum_{z=1}^{j}\left(\binom{j}{z} + \binom{j}{z-1}\right)\mu^z\lambda_l^{j+1-z}c(\mathbf{n} - (j+1-z)\mathbf{e}_l - z\mathbf{e}) \\
&\quad + \mu^{j+1}c(\mathbf{n} - (j+1)\mathbf{e}) \\
&= \sum_{z=0}^{j+1}\binom{j+1}{z}\mu^z\lambda_l^{j+1-z}c(\mathbf{n} - (j+1-z)\mathbf{e}_l - z\mathbf{e}).
\end{aligned}$$

Hence, (20.55) holds also for $i = j+1$, and by induction it holds for all non-negative integers i. □

Lemma 20.2 *For $l = 1, 2, \ldots, m$, we have*

$$\Xi_l^i \ddot{w}^*_\theta(\mathbf{n}) = \frac{n_l^{(i)}}{\theta^{i-1}}\ddot{w}_\theta(\mathbf{n}). \qquad (i = 0, 1, 2, \ldots; \gamma < \theta < \delta) \tag{20.56}$$

Proof For $i = 0$, (20.56) obviously holds. Let us now assume that it holds for i equal to some non-negative integer j. Then

$$\Xi_l^{j+1} \ddot{w}_\theta^*(\mathbf{n}) = \lambda_l \Xi_l^j \ddot{w}_\theta^*(\mathbf{n}-\mathbf{e}_l) + \mu \Xi_l^j \ddot{w}_\theta^*(\mathbf{n}-\mathbf{e})$$
$$= \lambda_l \frac{(n_l-1)^{(j)}}{\theta^{j-1}} \ddot{w}_\theta(\mathbf{n}-\mathbf{e}_l) + \mu \frac{(n_l-1)^{(j)}}{\theta^{j-1}} \ddot{w}_\theta(\mathbf{n}-\mathbf{e})$$
$$= \frac{n_l^{(j+1)}}{\theta^j} \frac{\theta}{n_l} (\lambda_l \ddot{w}_\theta(\mathbf{n}-\mathbf{e}_l) + \mu \ddot{w}_\theta(\mathbf{n}-\mathbf{e})) = \frac{n_l^{(j+1)}}{\theta^j} \ddot{w}_\theta(\mathbf{n}),$$

using (20.52) for the last step. Thus, (20.56) holds also for $i = j+1$, and by induction it holds for all non-negative integers i. □

Lemma 20.3 *We have*

$$n_l^{(j)} \ddot{v}_{i-j}(\mathbf{n}) = \Xi_l^j \ddot{v}_i(\mathbf{n}) = \sum_{z=0}^{j} \binom{j}{z} \mu^z \lambda_l^{j-z} \ddot{v}_i(\mathbf{n} - (j-z)\mathbf{e}_l - z\mathbf{e}).$$
$$(j = 0, 1, 2, \dots, i;\ i = 0, 1, 2, \dots) \tag{20.57}$$

Proof The second equality in (20.57) follows immediately from Lemma 20.1. The first equality obviously holds for $j = 0$. Let us now assume that it holds for $j = t-1$ for some positive integer $t \le i$. Then application of (20.53) gives

$$n_l^{(t)} \ddot{v}_{i-t}(\mathbf{n}) = (n_l - 1)^{(t-1)} n_l \ddot{v}_{i-t}(\mathbf{n})$$
$$= (n_l - 1)^{(t-1)} (\lambda_l \ddot{v}_{i-t+1}(\mathbf{n}-\mathbf{e}_l) + \mu \ddot{v}_{i-t+1}(\mathbf{n}-\mathbf{e}))$$
$$= \lambda_l \Xi_l^{t-1} \ddot{v}_i(\mathbf{n}-\mathbf{e}_l) + \mu \Xi_l^{t-1} \ddot{v}_i(\mathbf{n}-\mathbf{e}) = \Xi_l^t \ddot{v}_i(\mathbf{n}),$$

that is, (20.57) holds also for $j = t$, and by induction it holds $j = 0, 1, 2, \dots, i$.

This completes the proof of Lemma 20.3. □

By letting $j = i$ in (20.57), we obtain

$$p(\mathbf{n}) = \frac{\Xi_l^i \ddot{v}_i(\mathbf{n})}{n_l^{(i)}} = \frac{1}{n_l^{(i)}} \sum_{z=0}^{i} \binom{i}{z} \mu^z \lambda_l^{i-z} \ddot{v}_i(\mathbf{n} - (i-z)\mathbf{e}_l - z\mathbf{e}).$$
$$(i = 0, 1, 2, \dots) \tag{20.58}$$

As indicated above, we can always convert a situation with $\chi(0) \neq 0$ to a situation with $\chi(0) = 0$. Hence, for simplicity, we shall assume that $\chi(0) = 0$ in the following.

By application of (20.53) and (20.54), we obtain

$$
\begin{aligned}
\dot\rho(k)n_l\ddot v_{k-1}(\mathbf{n}) &= \dot\rho(k)(\lambda_l\ddot v_k(\mathbf{n}-\mathbf{e}_l)+\mu\ddot v_k(\mathbf{n}-\mathbf{e}))\\
&=\sum_{i=1}^k\chi(i)\big(\lambda_l(n_\bullet-1)\ddot v_{i-1}(\mathbf{n}-\mathbf{e}_l)+\mu(n_\bullet-m)\ddot v_{i-1}(\mathbf{n}-\mathbf{e})\\
&\quad-(m-1)\mu(\lambda_l\ddot v_i(\mathbf{n}-\mathbf{e}_l-\mathbf{e})+\mu\ddot v_i(\mathbf{n}-2\mathbf{e}))\big)\\
&\quad-\sum_{i=0}^{k-1}\dot\rho(i)(\lambda_l\ddot v_i(\mathbf{n}-\mathbf{e}_l)+\mu\ddot v_i(\mathbf{n}-\mathbf{e}))+\Xi_l\ddot w_{\gamma+}(\mathbf{n})-\Xi_l\ddot w_{\delta-}(\mathbf{n})\\
&=\chi(1)(\lambda_l(n_\bullet-1)p(\mathbf{n}-\mathbf{e}_l)+\mu(n_\bullet-m)p(\mathbf{n}-\mathbf{e}))\\
&\quad+\sum_{i=2}^k\chi(i)(n_l(n_\bullet-1)\ddot v_{i-2}(\mathbf{n})-\mu(m-1)\ddot v_{i-1}(\mathbf{n}-\mathbf{e}))\\
&\quad-(m-1)\mu(n_l-1)\sum_{i=1}^k\chi(i)\ddot v_{i-1}(\mathbf{n}-\mathbf{e})\\
&\quad-\dot\rho(0)(\lambda_l p(\mathbf{n}-\mathbf{e}_l)+\mu p(\mathbf{n}-\mathbf{e}))-n_l\sum_{i=1}^{k-1}\dot\rho(i)v_{i-1}(\mathbf{n})\\
&\quad+\Xi_l\ddot w_{\gamma+}(\mathbf{n})-\Xi_l\ddot w_{\delta-}(\mathbf{n}),
\end{aligned}
$$

that is,

$$
\begin{aligned}
\dot\rho(k)n_l\ddot v_{k-1}(\mathbf{n}) &= \chi(1)(\lambda_l(n_\bullet-1)p(\mathbf{n}-\mathbf{e}_l)+\mu(n_\bullet-(m-1)n_l-1)p(\mathbf{n}-\mathbf{e}))\\
&\quad+n_l\sum_{i=2}^k\chi(i)((n_\bullet-1)\ddot v_{i-2}(\mathbf{n})-\mu(m-1)\ddot v_{i-1}(\mathbf{n}-\mathbf{e}))\\
&\quad-\dot\rho(0)(\lambda_l p(\mathbf{n}-\mathbf{e}_l)+\mu p(\mathbf{n}-\mathbf{e}))-n_l\sum_{i=1}^{k-1}\dot\rho(i)v_{i-1}(\mathbf{n})\\
&\quad+\Xi_l\ddot w_{\gamma+}(\mathbf{n})-\Xi_l\ddot w_{\delta-}(\mathbf{n}).
\end{aligned}
\tag{20.59}
$$

We have now got rid of $\ddot v_k$. From (20.53), we obtain

$$
\begin{aligned}
\dot\rho(k)n_l^{(2)}\ddot v_{k-2}(\mathbf{n}) &= \lambda_l\dot\rho(k)(n_l-1)\ddot v_{k-1}(\mathbf{n}-\mathbf{e}_l)\\
&\quad+\mu\dot\rho(k)(n_l-1)\ddot v_{k-1}(\mathbf{n}-\mathbf{e}),
\end{aligned}
\tag{20.60}
$$

and by insertion of (20.59), we get rid of $\ddot v_{k-1}$ too. In this way, we can continue until we have got rid of all the $\ddot v_i$s apart from $\ddot v_0 = p$.

Let us consider the special case $k = 2$. Then (20.59) reduces to

$$\begin{aligned}\dot{\rho}(2)n_l\ddot{v}_1(\mathbf{n}) = {} & n_l((n_\bullet - 1)\chi(2) - \dot{\rho}(1))p(\mathbf{n}) \\ & + ((n_\bullet - 1)\chi(1) - \dot{\rho}(0))\lambda_l p(\mathbf{n} - \mathbf{e}_l) \\ & + (\chi(1)(n_\bullet - (m-1)n_l - 1) - \dot{\rho}(0))\mu p(\mathbf{n} - \mathbf{e}) \\ & - n_l(m-1)\chi(2)\mu\ddot{v}_1(\mathbf{n} - \mathbf{e}) + \Xi_l\ddot{w}_{\gamma+}(\mathbf{n}) - \Xi_l\ddot{w}_{\delta-}(\mathbf{n}).\end{aligned}$$

Application of (20.60) gives

$$\begin{aligned}\dot{\rho}(2)n_l^{(2)}p(\mathbf{n}) &= \lambda_l\dot{\rho}(2)(n_l - 1)\ddot{v}_1(\mathbf{n} - \mathbf{e}_l) + \mu\dot{\rho}(2)(n_l - 1)\ddot{v}_1(\mathbf{n} - \mathbf{e}) \\ &= \lambda_l(n_l - 1)((n_\bullet - 2)\chi(2) - \dot{\rho}(1))p(\mathbf{n} - \mathbf{e}_l) \\ &\quad + \mu(n_l - 1)((n_\bullet - m - 1)\chi(2) - \dot{\rho}(1))p(\mathbf{n} - \mathbf{e}) \\ &\quad + \lambda_l((n_\bullet - 2)\chi(1) - \dot{\rho}(0))\lambda_l p(\mathbf{n} - 2\mathbf{e}_l) \\ &\quad + \mu((n_\bullet - m - 1)\chi(1) - \dot{\rho}(0))\lambda_l p(\mathbf{n} - \mathbf{e} - \mathbf{e}_l) \\ &\quad + \lambda_l(\chi(1)(n_\bullet - (m-1)(n_l - 1) - 2) - \dot{\rho}(0))\mu p(\mathbf{n} - \mathbf{e}_l - \mathbf{e}) \\ &\quad + \mu(\chi(1)(n_\bullet - m - (m-1)(n_l - 1) - 1) - \dot{\rho}(0))\mu p(\mathbf{n} - 2\mathbf{e}) \\ &\quad - (n_l - 1)(m-1)\chi(2)\mu(\lambda_l\ddot{v}_1(\mathbf{n} - \mathbf{e}_l - \mathbf{e}) + \mu\ddot{v}_1(\mathbf{n} - 2\mathbf{e})) \\ &\quad + \Xi_l^2\ddot{w}_{\gamma+}(\mathbf{n}) - \Xi_l^2\ddot{w}_{\delta-}(\mathbf{n}).\end{aligned}$$

Using that from (20.53) we have

$$\lambda_l\ddot{v}_1(\mathbf{n} - \mathbf{e}_l - \mathbf{e}) + \mu\ddot{v}_1(\mathbf{n} - 2\mathbf{e}) = (n_l - 1)p(\mathbf{n} - \mathbf{e}),$$

some reorganisation and division by $n_l^{(2)}$ gives

$$\begin{aligned}\dot{\rho}(2)p(\mathbf{n}) = {} & \frac{(n_\bullet - 2)\chi(2) - \dot{\rho}(1)}{n_l}\lambda_l p(\mathbf{n} - \mathbf{e}_l) \\ & + \left(\frac{(n_\bullet - 2)\chi(2) - \dot{\rho}(1)}{n_l} - (m-1)\chi(2)\right)\mu p(\mathbf{n} - \mathbf{e}) \\ & + ((n_\bullet - 2)\chi(1) - \dot{\rho}(0))\frac{\lambda_l^2}{n_l^{(2)}}p(\mathbf{n} - 2\mathbf{e}_l) \\ & + \frac{1}{n_l - 1}\left(\frac{(n_\bullet - 2)\chi(1) - \dot{\rho}(0)}{n_l} - (m-1)\chi(1)\right)\mu^2 p(\mathbf{n} - 2\mathbf{e}) \\ & + \frac{1}{n_l - 1}\left(2\frac{(n_\bullet - 2)\chi(1) - \dot{\rho}(0)}{n_l} - (m-1)\chi(1)\right)\mu\lambda_l p(\mathbf{n} - \mathbf{e} - \mathbf{e}_l) \\ & + \frac{\Xi_l^2\ddot{w}_{\gamma+}(\mathbf{n}) - \Xi_l^2\ddot{w}_{\delta-}(\mathbf{n})}{n_l^{(2)}}. \qquad (\mathbf{n} \geq 2\mathbf{e}_l)\end{aligned} \tag{20.61}$$

When $k = 1$ and not necessarily $\chi(0) = 0$, insertion of the expressions for k^*, χ^*, and $\dot{\rho}^*$ gives

$$\begin{aligned}\dot{\rho}(1)p(\mathbf{n}) &= \frac{(n_\bullet - 1)\chi(1) - \dot{\rho}(0)}{n_l}\lambda_l p(\mathbf{n} - \mathbf{e}_l) \\ &+ \left(\frac{(n_\bullet - 1)\chi(1) - \dot{\rho}(0)}{n_l} - (m-1)\chi(1)\right)\mu p(\mathbf{n} - \mathbf{e}) \\ &+ \frac{\chi(0)}{n_l - 1}\left(\frac{n_\bullet - 1}{n_l}\lambda_l^2 p(\mathbf{n} - 2\mathbf{e}_l) + \left(\frac{n_\bullet - 1}{n_l} - m + 1\right)\mu^2 p(\mathbf{n} - 2\mathbf{e})\right. \\ &+ \left.\left(2\frac{n_\bullet - 1}{n_l} - m + 1\right)\mu\lambda_l p(\mathbf{n} - \mathbf{e} - \mathbf{e}_l)\right) \\ &+ \frac{\Xi_l^2 \ddot{w}^*_{\gamma+}(\mathbf{n}) - \Xi_l^2 \ddot{w}^*_{\delta-}(\mathbf{n})}{n_l^{(2)}}. \qquad (\mathbf{n} \geq 2\mathbf{e}_l) \end{aligned} \tag{20.62}$$

20.7.3 The Special Case $\mu = 0$

For the rest of this section, we restrict to the special case $\mu = 0$, but general k. In this case, insertion of (20.12) in (20.44) gives that for all $\mathbf{n} \in \mathbb{N}_m$ and $\gamma < \theta < \delta$

$$\ddot{w}_\theta(\mathbf{n}) = q_\theta(n_\bullet)g^{n_\bullet *}(\mathbf{n})u(\theta)\sum_{i=0}^{k}\chi(i)\theta^i,$$

and by insertion of (20.33) we get

$$\ddot{w}_\theta(\mathbf{n}) = \dot{w}_\theta(n_\bullet)g^{n_\bullet *}(\mathbf{n}). \tag{20.63}$$

We shall now prove a multivariate corollary to Corollary 20.1.

Corollary 20.2 *For $i = 1, 2, \ldots, k+1$, let ${}_i\mathbf{y} \in \mathbb{N}_{m+}$ with ${}_iy_\bullet = i$ such that ${}_1\mathbf{y} < {}_2\mathbf{y} < \cdots < {}_{k+1}\mathbf{y}$. Then*

$$\begin{aligned}\dot{\rho}(k)p(\mathbf{n}) &= \sum_{i=1}^{k+1}((n_\bullet - k)\chi(k-i+1) - \dot{\rho}(k-i))\left(\prod_{j=1}^{m}\frac{\lambda_j^{{}_iy_j}}{n_j^{({}_iy_j)}}\right)p(\mathbf{n} - {}_i\mathbf{y}) \\ &+ (\ddot{w}_{\gamma+}(\mathbf{n} - {}_k\mathbf{y}) - \ddot{w}_{\delta-}(\mathbf{n} - {}_k\mathbf{y}))\left(\prod_{j=1}^{m}\frac{\lambda_j^{{}_ky_j}}{n_j^{({}_ky_j)}}\right). \qquad (\mathbf{n} \geq {}_{k+1}\mathbf{y}) \end{aligned} \tag{20.64}$$

Proof Application of (20.10), Theorem 15.8, Corollary 20.1, and (20.63) gives that for all $\mathbf{n} \in \mathbb{N}_{m+}$ such that $\mathbf{n} \geq {}_{k+1}\mathbf{y}$, we have

$$\begin{aligned}
\dot{\rho}(k)p(\mathbf{n}) &= (\dot{w}_{\gamma+}(n_\bullet - k) - \dot{w}_{\delta-}(n_\bullet - k))\frac{\nu^k}{n_\bullet^{(k)}} g^{n_\bullet *}(\mathbf{n}) \\
&\quad + \sum_{i=1}^{k+1}((n_\bullet - k)\chi(k-i+1) - \dot{\rho}(k-i))\frac{\nu^i}{n_\bullet^{(i)}} n_\bullet^{(i)} \\
&\quad \times \left(\prod_{j=1}^{m} \frac{(\lambda_j/\nu)^{i y_j}}{n_j^{(i y_j)}}\right) p(\mathbf{n} - {}_i\mathbf{y}) \\
&= (\ddot{w}_{\gamma+}(\mathbf{n} - {}_k\mathbf{y}) - \ddot{w}_{\delta-}(\mathbf{n} - {}_k\mathbf{y}))\frac{\nu^k}{n_\bullet^{(k)}} \frac{g^{n_\bullet *}(\mathbf{n})}{g^{(n_\bullet - k)*}(\mathbf{n} - {}_k\mathbf{y})} \\
&\quad + \sum_{i=1}^{k+1}((n_\bullet - k)\chi(k-i+1) - \dot{\rho}(k-i))\left(\prod_{j=1}^{m} \frac{\lambda_j^{i y_j}}{n_j^{(i y_j)}}\right) p(\mathbf{n} - {}_i\mathbf{y}),
\end{aligned}$$

and insertion of (14.4) gives (20.64). □

When letting ${}_i\mathbf{y} = i\mathbf{e}_l$ for $i = 1, 2, \ldots, k+1$ and $l \in \{1, 2, \ldots, m\}$, (20.64) reduces to

$$\begin{aligned}
\dot{\rho}(k)p(\mathbf{n}) &= \sum_{i=1}^{k+1}((n_\bullet - k)\chi(k-i+1) - \dot{\rho}(k-i))\frac{\lambda_l^i}{n_l^{(i)}} p(\mathbf{n} - i\mathbf{e}_l) \\
&\quad + (\ddot{w}_{\gamma+}(\mathbf{n} - k\mathbf{e}_l) - \ddot{w}_{\delta-}(\mathbf{n} - k\mathbf{e}_l))\frac{\lambda_l^k}{n_l^{(k)}}. \qquad (20.65)
\end{aligned}$$

$$(\mathbf{n} \geq (k+1)\mathbf{e}_l;\, l = 1, 2, \ldots, m)$$

When $k = 1$, this reduces to

$$\begin{aligned}
\dot{\rho}(1)p(\mathbf{n}) &= ((n_\bullet - 1)\chi(1) - \dot{\rho}(0))\frac{\lambda_l}{n_l} p(\mathbf{n} - \mathbf{e}_l) + (n_\bullet - 1)\chi(0)\frac{\lambda_l^2}{n_l^{(2)}} p(\mathbf{n} - 2\mathbf{e}_l) \\
&\quad + (\ddot{w}_{\gamma+}(\mathbf{n} - \mathbf{e}_l) - \ddot{w}_{\delta-}(\mathbf{n} - \mathbf{e}_l))\frac{\lambda_l}{n_l}, \qquad (\mathbf{n} \geq 2\mathbf{e}_l;\, l = 1, 2, \ldots, m)
\end{aligned}$$

which is a special case of (20.62), and for $k = 2$, we obtain

$$\begin{aligned}
\dot{\rho}(2)p(\mathbf{n}) &= ((n_\bullet - 2)\chi(2) - \dot{\rho}(1))\frac{\lambda_l}{n_l} p(\mathbf{n} - \mathbf{e}_l) \\
&\quad + ((n_\bullet - 2)\chi(1) - \dot{\rho}(0))\frac{\lambda_l^2}{n_l^{(2)}} p(\mathbf{n} - 2\mathbf{e}_l)
\end{aligned}$$

$$+ (n_\bullet - 2)\chi(0)\frac{\lambda_l^3}{n_l^{(3)}}p(\mathbf{n} - 3\mathbf{e}_l)$$

$$+ (\ddot{w}_{\gamma+}(\mathbf{n} - 2\mathbf{e}_l) - \ddot{w}_{\delta-}(\mathbf{n} - 2\mathbf{e}_l))\frac{\lambda_l^2}{n_l^{(2)}}, \quad (\mathbf{n} \geq 3\mathbf{e}_l;\, l = 1, 2, \dots, m)$$

which follows from (20.61) when $\chi(0) = 0$.

20.8 Special Classes of Mixing Distributions

20.8.1 Shifted Pareto distribution

Let U be the shifted Pareto distribution $\mathrm{SPar}(\alpha, \gamma)$ with density u given by (3.33). Application of (20.11) gives

$$q(n) = \alpha\gamma^\alpha \frac{\nu^n}{n!}\int_\gamma^\infty \theta^{n-\alpha-1}\mathrm{e}^{-\theta\nu}\,\mathrm{d}\theta. \qquad (n = 0, 1, 2, \dots)$$

By partial integration, we obtain the recursion

$$q(n) = \frac{\alpha(\gamma\nu)^{n-1}}{n!}\mathrm{e}^{-\gamma\nu} + \left(1 - \frac{\alpha+1}{n}\right)q(n-1) \qquad (n = 1, 2, \dots)$$

with initial value

$$q(0) = \alpha\gamma^\alpha \int_\gamma^\infty \frac{\mathrm{e}^{-\theta\nu}}{\theta^{\alpha+1}}\,\mathrm{d}\theta.$$

Application of (15.25) and (15.26) gives the recursions

$$f(\mathbf{x}) = \frac{1}{1 - h(\mathbf{0})}\left((b_\gamma \vee h)(\mathbf{x}) + \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}}\left(1 - (1+\alpha)\frac{y_l}{x_l}\right)h(\mathbf{y})f(\mathbf{x} - \mathbf{y})\right)$$

$$(\mathbf{x} \geq \mathbf{e}_l;\, l = 1, 2, \dots, m)$$

$$f(\mathbf{x}) = \frac{1}{1 - h(\mathbf{0})}\left((b_\gamma \vee h)(\mathbf{x}) + \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}}\left(1 - (1+\alpha)\frac{y_\bullet}{x_\bullet}\right)h(\mathbf{y})f(\mathbf{x} - \mathbf{y})\right)$$

$$(\mathbf{x} \in \mathbb{N}_{m+})$$

with

$$b_\gamma(n) = \frac{\alpha(\gamma\nu)^{n-1}}{n!}\mathrm{e}^{-\gamma\nu}, \qquad (n = 0, 1, 2, \dots) \tag{20.66}$$

and insertion of (3.33) in (20.19) gives the initial value

$$f(\mathbf{0}) = \alpha\gamma^{\alpha} \int_{\gamma}^{\infty} \frac{\mathrm{e}^{-\theta\nu(1-h(\mathbf{0}))}}{\theta^{\alpha+1}}\,\mathrm{d}\theta.$$

We could also have obtained these recursions from (20.37)–(20.40).

We can evaluate $b_\gamma \vee h$ recursively by (20.41) or (20.42) with initial condition

$$(b_\gamma \vee h)(0) = \frac{\alpha}{\gamma\nu}\mathrm{e}^{-\gamma\nu(1-h(\mathbf{0}))}$$

obtained by insertion of (20.66) in (1.7).

20.8.2 Pareto Distribution

Let U be the Pareto distribution $\mathrm{Par}(\alpha, \beta)$ with density u given by (3.34). From (3.35), we obtain that u satisfies (3.15) with $k = 1$, $\gamma = 0$, and $\delta = \infty$, and η and χ are given by (3.36) and (3.37) respectively. Insertion in (20.34) gives

$$\dot\rho(0) = \alpha + \beta\nu; \qquad \dot\rho(1) = \nu.$$

By insertion in (20.39) and (20.40), we obtain the recursions

$$f(\mathbf{x}) = \frac{1}{1-h(\mathbf{0})}\Bigg((q(1) + (\alpha+\beta\nu)q(0))h(\mathbf{x}) + \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}}\Bigg(\bigg(1-(\alpha+\beta\nu+1)\frac{y_l}{x_l}\bigg)h(\mathbf{y}) + \nu\frac{\beta}{2}\frac{y_l}{x_l}h^{2*}(\mathbf{y})\Bigg)f(\mathbf{x}-\mathbf{y})\Bigg)$$

$$(\mathbf{x} \ge \mathbf{e}_l;\, l = 1, 2, \ldots, m)$$

$$f(\mathbf{x}) = \frac{1}{1-h(\mathbf{0})}\Bigg((q(1) + (\alpha+\beta\nu)q(0))h(\mathbf{x}) + \sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}}\Bigg(\bigg(1-(\alpha+\beta\nu+1)\frac{y_\bullet}{x_\bullet}\bigg)h(\mathbf{y}) + \nu\frac{\beta}{2}\frac{y_\bullet}{x_\bullet}h^{2*}(\mathbf{y})\Bigg)f(\mathbf{x}-\mathbf{y})\Bigg),$$

$$(\mathbf{x} \in \mathbb{N}_{m+})$$

and insertion of (3.34) in (20.19) gives the initial value

$$f(\mathbf{0}) = \alpha\beta^{\alpha} \int_{0}^{\infty} \frac{\mathrm{e}^{-\theta\nu(1-h(\mathbf{0}))}}{(\beta+\theta)^{\alpha+1}}\,\mathrm{d}\theta.$$

By insertion of (3.34) in (20.11), we get

$$q(n) = \frac{\nu^n}{n!}\alpha\beta^\alpha \int_0^\infty \theta^n \frac{\mathrm{e}^{-\theta\nu}}{(\beta+\theta)^{\alpha+1}}\,\mathrm{d}\theta, \qquad (n = 0, 1, 2, \dots)$$

which gives in particular $q(0)$ and $q(1)$.

From (20.62), we obtain that under the special design (20.6) we have the recursion

$$\begin{aligned} p(\mathbf{n}) = \frac{1}{\nu}\Bigg(&\frac{n_\bullet - \alpha - \beta\nu - 1}{n_l}\lambda_l p(\mathbf{n} - \mathbf{e}_l) \\ &+ \left(\frac{n_\bullet - \alpha - \beta\nu - 1}{n_l} - m + 1\right)\mu p(\mathbf{n} - \mathbf{e}) \\ &+ \frac{\beta}{n_l - 1}\Bigg(\frac{n_\bullet - 1}{n_l}\lambda_l^2 p(\mathbf{n} - 2\mathbf{e}_l) + \left(\frac{n_\bullet - 1}{n_l} - m + 1\right)\mu^2 p(\mathbf{n} - 2\mathbf{e}) \\ &+ \left(2\frac{n_\bullet - 1}{n_l} - m + 1\right)\mu\lambda_l p(\mathbf{n} - \mathbf{e} - \mathbf{e}_l)\Bigg)\Bigg). \quad (\mathbf{n} \geq 2\mathbf{e}_l;\, l = 1, 2, \dots, m) \end{aligned}$$

20.8.3 Truncated Normal Distribution

Let U be the truncated normal distribution $\mathrm{TN}(\xi, \sigma)$ with density u given by (3.38). From (3.39), we obtain that u satisfies (3.15) with $k = 1$, $\gamma = 0$, and $\delta = \infty$, and η and χ are given by (3.40) and (3.41) respectively. Insertion in (20.34) gives

$$\dot{\rho}(0) = \nu\sigma^2 - \xi; \qquad \dot{\rho}(1) = 1.$$

By insertion in (20.39) and (20.40), we obtain the recursions

$$\begin{aligned} f(\mathbf{x}) &= (q(1) + (\nu\sigma^2 - \xi)\nu q(0))h(\mathbf{x}) \\ &\quad + \frac{\nu}{x_l}\sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} y_l\left((\xi - \nu\sigma^2)h(\mathbf{y}) + \nu\frac{\sigma^2}{2}h^{2*}(\mathbf{y})\right)f(\mathbf{x} - \mathbf{y}) \\ &(\mathbf{x} \geq \mathbf{e}_l;\, l = 1, 2, \dots, m) \end{aligned}$$

$$\begin{aligned} f(\mathbf{x}) &= (q(1) + (\nu\sigma^2 - \xi)\nu q(0))h(\mathbf{x}) \\ &\quad + \frac{\nu}{x_\bullet}\sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} y_\bullet\left((\xi - \nu\sigma^2)h(\mathbf{y}) + \nu\frac{\sigma^2}{2}h^{2*}(\mathbf{y})\right)f(\mathbf{x} - \mathbf{y}), \qquad (\mathbf{x} \in \mathbb{N}_{m+}) \end{aligned}$$

and insertion of (3.38) in (20.19) gives the initial condition

$$f(\mathbf{0}) = \frac{\int_0^\infty \mathrm{e}^{-\theta\nu(1-h(\mathbf{0})) - \frac{1}{2\sigma^2}(\theta-\xi)^2}\mathrm{d}\theta}{\int_0^\infty \mathrm{e}^{-\frac{1}{2\sigma^2}(\theta-\xi)^2}\mathrm{d}\theta}.$$

By insertion of (3.38) in (20.11), we get

$$q(n) = \frac{\nu^n}{n!} \frac{\int_0^\infty \theta^n \exp(-\theta\nu - \frac{1}{2\sigma^2}(\theta-\xi)^2)\mathrm{d}\theta}{\int_0^\infty \exp(-\frac{1}{2\sigma^2}(\theta-\xi)^2)\mathrm{d}\theta}, \qquad (n = 0, 1, 2, \dots)$$

which gives in particular $q(0)$ and $q(1)$.

From (20.62), we obtain that under the special design (20.6) we have the recursion

$$p(\mathbf{n}) = \frac{\xi - \nu\sigma^2}{n_l}(\lambda_l p(\mathbf{n}-\mathbf{e}_l) + \mu p(\mathbf{n}-\mathbf{e})) + \frac{\sigma^2}{n_l - 1}\left(\frac{n_\bullet - 1}{n_l}\lambda_l^2 p(\mathbf{n}-2\mathbf{e}_l) + \left(\frac{n_\bullet - 1}{n_l} - m + 1\right)\mu^2 p(\mathbf{n}-2\mathbf{e}) + \left(2\frac{n_\bullet - 1}{n_l} - m + 1\right)\mu\lambda_l p(\mathbf{n}-\mathbf{e}-\mathbf{e}_l)\right). \qquad (\mathbf{n} \geq 2\mathbf{e}_l;\ l = 1, 2, \dots, m)$$

20.8.4 Inverse Gauss Distribution

Let U be the inverse Gauss distribution $\mathrm{IGauss}(\xi, \beta)$ with density u given by (3.42). From (3.43), we obtain that u satisfies (3.15) with $k = 2$, $\gamma = 0$, and $\delta = \infty$, and η and χ are given by (3.44) and (3.45) respectively. Insertion in (20.25) gives

$$\rho(0) = -\xi^2; \qquad \rho(1) = -\beta; \qquad \rho(2) = 2(1 - h(\mathbf{0}))\nu\beta + 1,$$

and by letting $h(\mathbf{0}) = 0$, we obtain

$$\dot\rho(0) = -\xi^2; \qquad \dot\rho(1) = -\beta; \qquad \dot\rho(2) = 2\nu\beta + 1.$$

Insertion in (20.28) gives

$$v_2(\mathbf{x}) = \frac{1}{2(1 - h(\mathbf{0}))\nu\beta + 1}\left(2\nu\beta \sum_{\mathbf{0}<\mathbf{y}\leq\mathbf{x}} h(\mathbf{y})v_2(\mathbf{x}-\mathbf{y}) + \xi^2 f(\mathbf{x}) + \beta v_1(\mathbf{x})\right). \qquad (\mathbf{x} \in \mathbb{N}_{m+})$$

From (20.61), we obtain that under the special design (20.6) we have the recursion

$$p(\mathbf{n}) = \frac{1}{2\nu\beta + 1}\left(\frac{2n_\bullet - 3}{n_l}\beta\lambda_l p(\mathbf{n}-\mathbf{e}_l) + \left(\frac{2n_\bullet - 3}{n_l} - 2m + 2\right)\beta\mu p(\mathbf{n}-\mathbf{e}) + \frac{\xi^2}{n_l^{(2)}}(\lambda_l^2 p(\mathbf{n}-2\mathbf{e}_l) + \mu^2 p(\mathbf{n}-2\mathbf{e}) + 2\mu\lambda_l p(\mathbf{n}-\mathbf{e}-\mathbf{e}_l))\right). \qquad (\mathbf{n} \geq 2\mathbf{e}_l;\ l = 1, 2, \dots, m)$$

20.8.5 Transformed Gamma Distribution

Let U be the transformed Gamma distribution $\mathrm{TGamma}(\alpha, \beta, k)$ with density u given by (3.46) with k being a positive integer greater than one. From (3.47), we obtain that u satisfies (3.15) with $\gamma = 0$ and $\delta = \infty$, and η and χ are given by (3.48) and (3.49) respectively. Insertion in (20.25) gives

$$\rho(0) = -k\alpha; \qquad \rho(1) = (1 - h(\mathbf{0}))\nu; \qquad \rho(k) = k\beta,$$

and by letting $h(\mathbf{0}) = 0$, we obtain

$$\dot{\rho}(0) = -k\alpha; \qquad \dot{\rho}(1) = \nu; \qquad \dot{\rho}(k) = k\beta.$$

Insertion in (20.28) gives

$$v_k(\mathbf{x}) = \frac{1}{\beta}\left(\frac{\nu}{k}\left(\sum_{\mathbf{0}<\mathbf{y}\le\mathbf{x}} h(\mathbf{y})v_1(\mathbf{x}-\mathbf{y}) - (1-h(\mathbf{0}))v_1(\mathbf{x})\right) + \alpha f(\mathbf{x})\right).$$
$$(\mathbf{x} \in \mathbb{N}_{m+})$$

Let us now restrict to the special design (20.6). Then (20.59) gives

$$\ddot{v}_{k-1}(\mathbf{n}) = \frac{1}{k\beta}\left(-\nu p(\mathbf{n}) + \frac{n_\bullet + k\alpha - 1}{n_l}\lambda_l p(\mathbf{n}-\mathbf{e}_l) + \left(\frac{n_\bullet + k\alpha - 1}{n_l} - m + 1\right)\mu p(\mathbf{n}-\mathbf{e})\right). \tag{20.67}$$
$$(\mathbf{n} \ge \mathbf{e}_l;\, l = 1, 2, \dots, m)$$

Letting $i = k - 1$ in (20.58) gives

$$p(\mathbf{n}) = \frac{1}{n_l^{(k-1)}} \sum_{z=0}^{k-1} \binom{k-1}{z} \mu^z \lambda_l^{k-z-1} \ddot{v}_{k-1}(\mathbf{n} - (k-z-1)\mathbf{e}_l - z\mathbf{e}),$$
$$(\mathbf{n} \ge (k-1)\mathbf{e}_l;\, l = 1, 2, \dots, m)$$

and by insertion of (20.67) we obtain

$$p(\mathbf{n}) = \frac{1}{k\beta n_l^{(k-1)}} \sum_{z=0}^{k-1} \binom{k-1}{z} \mu^z \lambda_l^{k-z-1} \Bigg(-\nu p(\mathbf{n} - (k-z-1)\mathbf{e}_l - z\mathbf{e})$$
$$+ \frac{n_\bullet + k(\alpha-1) - z(m-1)}{n_l - k + 1}\lambda_l p(\mathbf{n} - (k-z)\mathbf{e}_l - z\mathbf{e})$$
$$+ \left(\frac{n_\bullet + k(\alpha-1) - z(m-1)}{n_l - k + 1} - m + 1\right)$$

$$\times \mu p(\mathbf{n} - (k - z - 1)\mathbf{e}_l - (z + 1)\mathbf{e})\Bigg).$$

$$(\mathbf{n} \geq k\mathbf{e}_l;\, l = 1, 2, \ldots, m) \tag{20.68}$$

When $k = 2$, (20.61) gives

$$p(\mathbf{n}) = \frac{1}{2\beta}\left(-\frac{\nu}{n_l}(\lambda_l p(\mathbf{n} - \mathbf{e}_l) + \mu p(\mathbf{n} - \mathbf{e})) + \frac{1}{n_l - 1}\left(\frac{n_\bullet + 2\alpha - 2}{n_l}\lambda_l^2 p(\mathbf{n} - 2\mathbf{e}_l)\right.\right.$$
$$+ \left(\frac{n_\bullet + 2\alpha - 2}{n_l} - m + 1\right)\mu^2 p(\mathbf{n} - 2\mathbf{e})$$
$$\left.\left. + \left(2\frac{n_\bullet + 2\alpha - 2}{n_l} - m + 1\right)\mu\lambda_l p(\mathbf{n} - \mathbf{e} - \mathbf{e}_l)\right)\right).$$
$$(\mathbf{n} \geq 2\mathbf{e}_l;\, l = 1, 2, \ldots, m)$$

When $\mu = 0$ and $k = 2, 3, \ldots$, (20.65) gives

$$p(\mathbf{n}) = \frac{1}{k\beta}\left(-\nu\frac{\lambda_l^{k-1}}{n_l^{(k-1)}} p(\mathbf{n} - (k - 1)\mathbf{e}_l) + (n_\bullet + k(\alpha - 1))\frac{\lambda_l^k}{n_l^{(k)}} p(\mathbf{n} - k\mathbf{e}_l)\right).$$
$$(\mathbf{n} \geq k\mathbf{e}_l;\, l = 1, 2, \ldots, m)$$

Both these recursions can also be deduced from (20.68).

Further Remarks and References

This chapter is to a large extent based on Sundt and Vernic (2004).

In the bivariate case $k = 2$, Hesselager (1996b) deduced the algorithm after (20.28), and Vernic (2002) discussed the algorithm after (20.23) without making the restriction (3.15).

Binnenhei (2004) developed recursions for bivariate distributions in connection with a modification of the CreditRisk+ model.

References

Adelson, R.M. (1966). Compound Poisson distributions. Oper. Res. Q. **17**, 73–75

Albrecht, P. (1981). Dynamische statistische Entscheidungsverfahren für Schadenzahlprozesse. Verlag Versicherungswirtschaft e.V., Karlsruhe

Albrecht, P. (1984). Laplace transforms, Mellin transforms and mixed Poisson processes. Scand. Actuar. J., 58–64

Ambagaspitiya, R.S. (1995). A family of discrete distributions. Insur. Math. Econ. **16**, 107–127

Ambagaspitiya, R.S. (1998). Compound bivariate Lagrangian Poisson distributions. Insur. Math. Econ. **23**, 21–31

Ambagaspitiya, R.S. (1999). On the distributions of two classes of correlated aggregate claims. Insur. Math. Econ. **24**, 301–308

Ambagaspitiya, R.S., Balakrishnan, N. (1994). On the compound generalized Poisson distribution. ASTIN Bull. **24**, 255–263

Ammeter, H. (1948). A generalization of the collective theory of risk in regard to fluctuating basic-probabilities. Skand. Aktuarietidskr. **31**, 171–198

Ammeter, H. (1949). Die Elemente der kollektiven Risikotheorie von festen und zufallsartig schwankenden Grundwahrscheinlichkeiten. Mitt. Ver. Schweiz. Versicher.math. **49**, 35–95

Antzoulakos, D.L., Chadjiconstantinidis, S. (2004). On mixed and compound mixed Poisson distributions. Scand. Actuar. J., 161–188

Barbour, A.D., Chryssaphinou, O. (2001). Compound Poisson approximation: a user's guide. Ann. Appl. Probab. **11**, 964–1002

Beall, G., Rescia, R.R. (1953). A generalization of Neyman's contagious distributions. Biometrics **9**, 354–386

Bertram, J. (1981). Numerische Berechnung von Gesamtschadenverteilungen. Blätter Dtsch. Ges. Versicher.math. **15**, 175–194

Binnenhei, C. (2004). An analytic approach to rating transitions. In: Gundlach, M., Lehrbass, F. (eds.) CreditRisk+ in the Banking Industry, pp. 175–202. Springer, Berlin

Bowers, N.L., Gerber, H.U., Hickman, J.C., Nesbitt, C.J. (1997). Actuarial Mathematics, 2nd edn. Society of Actuaries, Itasca

Buchmann, B., Grübel, R. (2003). Decompounding: an estimation problem for Poisson random sums. Ann. Statist. **31**, 1054–1074

Bühlmann, H. (1984). Numerical evaluation of the compound Poisson distribution: recursion or Fast Fourier Transform? Scand. Actuar. J., 116–126

Bühlmann, H., Buzzi, R. (1971). On a transformation of the weighted compound Poisson process. ASTIN Bull. **6**, 42–46

Bühlmann, H., Gagliardi, B., Gerber, H.U., Straub, E. (1977). Some inequalities for stop-loss premiums. ASTIN Bull. **9**, 75–83

Bürgisser, P., Kurth, A., Wagner, A. (2001). Incorporating severity variations into credit risk. J. Risk **3**, 5–31

Carver, H.C. (1919). On the graduation of frequency distributions. Proc. Casualty Actuar. Soc. **6**, 52–72

Chadjiconstantinidis, S., Antzoulakos, D.L. (2002). Moments of compound mixed Poisson distributions. Scand. Actuar. J., 138–161

Chadjiconstantinidis, S., Pitselis, G. (2008, in press). Further improved recursions for a class of compound Poisson distributions. Insur. Math. Econ.

Chan, B. (1982a). Recursive formulas for aggregate claims. Scand. Actuar. J., 38–40

Chan, B. (1982b). Recursive formulas for discrete distributions. Insur. Math. Econ. **1**, 241–243

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0,

Chan, F.-Y. (1984). On a family of aggregate claims distributions. Insur. Math. Econ. **3**, 151–155
Cossette, H., Gaillardetz, P., Marceau, É., Rioux, J. (2002). On two dependent individual risk models. Insur. Math. Econ. **30**, 153–166
Cossette, H., Landriault, D., Marceau, É. (2004). Exact expressions and upper bound for ruin probabilities in the compound Markov binomial model. Insur. Math. Econ. **34**, 449–466
Credit Suisse Financial Products (1997). CreditRisk+: A Credit Risk Management Framework. Credit Suisse Financial Products, London
De Pril, N. (1985). Recursions for convolutions of arithmetic distributions. ASTIN Bull. **15**, 135–139
De Pril, N. (1986a). Improved recursions for some compound Poisson distributions. Insur. Math. Econ. **5**, 129–132
De Pril, N. (1986b). On the exact computation of the aggregate claims distribution in the individual life model. ASTIN Bull. **16**, 109–112
De Pril, N. (1986c). Moments of a class of compound distributions. Scand. Actuar. J., 117–120
De Pril, N. (1988). Improved approximations for the aggregate claims distribution of a life insurance portfolio. Scand. Actuar. J., 61–68
De Pril, N. (1989). The aggregate claims distribution in the individual model with arbitrary positive claims. ASTIN Bull. **19**, 9–24
De Pril, N., Dhaene, J. (1992). Error bounds for compound Poisson approximations of the individual risk model. ASTIN Bull. **22**, 135–148
De Vylder, F., Goovaerts, M.J. (1988). Recursive calculation of finite time ruin probabilities. Insur. Math. Econ. **7**, 1–7
DeGroot, M.H. (2004). Optimal Statistical Decisions, 2nd edn. Wiley, New York
Delaporte, P. (1959). Quelques problèmes de statistique mathématique posés par l'assurance automobile et le bonus pour non sinistre. Bull. Trimest. Inst. Actuar. Fr. **227**, 87–102
Delaporte, P. (1960). Un problème de tarification de l'assurance accidents d'automobiles examiné par la statistique mathématique. Trans. Int. Congr. Actuar. **16.2**, 121–135
Delaporte, P. (1972a). Les mathématiques de l'assurance automobile. ASTIN Bull. **6**, 185–190
Delaporte, P. (1972b). Construction d'un tarif d'assurance automobile basé sur le princip de la prime modelée sur le risque. Mitt. Ver. Schweiz. Versicher.math. **72**, 101–113
den Iseger, P.W., Smith, M.A.J., Dekker, R. (1997). Computing compound distributions faster! Insur. Math. Econ. **20**, 23–34
Denuit, M., Dhaene, J., Goovaerts, M.J., Kaas, R. (2005). Actuarial Theory for Dependent Risks: Measures, Orders and Models. Wiley, New York
Dhaene, J., De Pril, N. (1994). On a class of approximative computation methods in the individual risk model. Insur. Math. Econ. **14**, 181–196
Dhaene, J., Goovaerts, M.J. (1997). On the dependency of risks in the individual life model. Insur. Math. Econ. **19**, 243–253
Dhaene, J., Sundt, B. (1997). On error bounds for approximations to aggregate claims distributions. ASTIN Bull. **27**, 243–262
Dhaene, J., Sundt, B. (1998). On approximating distributions by approximating their De Pril transforms. Scand. Actuar. J., 1–23
Dhaene, J., Vandebroek, M. (1995). Recursions for the individual model. Insur. Math. Econ. **16**, 31–38
Dhaene, J., Sundt, B., De Pril, N. (1996). Some moment relations for the Hipp approximation. ASTIN Bull. **26**, 117–121
Dhaene, J., Willmot, G.E., Sundt, B. (1999). Recursions for distribution functions and stop-loss transforms. Scand. Actuar. J., 52–65. Corrigendum: Scand. Actuar. J., 226 (2007)
Dhaene, J., Ribas, C., Vernic, R. (2006). Recursions for the individual risk model. Acta Math. Appl. Sin. Engl. Ser. **22**, 543–564
Dickson, D.C.M. (1995). A review of Panjer's recursion formula and its applications. Br. Actuar. J. **1**, 107–124
Dickson, D.C.M. (2005). Insurance Risk and Ruin. Cambridge University Press, Cambridge
Dickson, D.C.M., Sundt, B. (2001). Comparison of methods for evaluation of the convolution of two compound $\mathcal{R}_1$ distributions. Scand. Actuar. J., 40–54

Doray, L.G., Haziza, A. (2004). Minimum distance inference for Sundt's distribution. In: Antoch, J. (ed.) COMPSTAT 2004—Proceedings in Computational Statistics, 16th Symposium Held in Prague, Czech Republic, 2004. Springer, Heidelberg

Douligeris, C., Iakovou, E., Englehardt, J.D., Li, H., Ip, C.M., Mooers, C.N.K. (1997). Development of a national marine oil transportation system model. Spill Sci. Technol. Bull. **4**, 113–121

Dufresne, F. (1996). An extension of Kornya's method with application to pension funds. Bull. Swiss Assoc. Actuar., 171–181

Eisele, K.-T. (2006). Recursions for compound phase distributions. Insur. Math. Econ. **38**, 149–156

Eisele, K.-T. (2008). Recursions for multivariate compound phase variables. Insur. Math. Econ. **42**, 65–72

Embrechts, P., Maejima, M., Teugels, J.L. (1985). Asymptotic behaviour of compound distributions. ASTIN Bull. **15**, 45–48

Euler, L. (1748). Introductio in analysin infinitorum. Bousquet, Lausanne

Fang, Y. (2003a). GMM tests for the Katz family of distributions. J. Stat. Plan. Inference **110**, 55–73

Fang, Y. (2003b). Semi-parametric specification tests for discrete probability models. J. Risk Insur. **70**, 73–84

Feilmeier, M., Bertram, J. (1987). Anwendung numerischer Methoden in der Risikotheorie. Verlag Versicherungswirtschaft e.V., Karlsruhe

Feller, W. (1968). An Introduction to Probability Theory and Its Applications, vol. 1, 3rd edn. Wiley, New York

Gerathewohl, K. (1980). Reinsurance – Principles and Practice, vol. 1. Verlag Versicherungswirtschaft GmbH., Karlsruhe

Gerathewohl, K. (1983). Reinsurance – Principles and Practice, vol. 2. Verlag Versicherungswirtschaft GmbH., Karlsruhe

Gerber, H.U. (1979). An Introduction to Mathematical Risk Theory. Huebner Foundation Monographs, vol. 8. Richard D. Irwin, Homewood

Gerber, H.U. (1982). On the numerical evaluation of the distribution of aggregate claims and its stop-loss premiums. Insur. Math. Econ. **1**, 13–18

Gerber, H.U. (1984). Error bounds for the compound Poisson approximation. Insur. Math. Econ. **3**, 191–194

Gerber, H.U., Jones, D.A. (1976). Some practical considerations in connection with the calculation of stop-loss premiums. Trans. Soc. Actuar. **28**, 215–231. Discussion: Trans. Soc. Actuar. **28**, 233–235 (1976)

Gerhold, S., Schmock, U., Warnung, R. (2008, in press). A generalization of Panjer's recursion and numerically stable risk aggregation. Finance Stoch.

Giese, G. (2003). Enhancing CreditRisk+. Risk **16**, 73–77

Godambe, A.V., Patil, G.P. (1975). Some characterisations involving additivity and infinite divisibility and their applications to Poisson mixtures and Poisson sums. In: Patil, G.P., Katz, S., Ord, J.K. (eds.) Statistical Distributions in Scientific Work, vol. 3, Characterisations and Applications, pp. 339–351

Goovaerts, M.J., De Vylder, F. (1984). A stable recursive algorithm for evaluation of ultimate ruin probabilities. ASTIN Bull. **14**, 53–59

Goovaerts, M.J., Dhaene, J. (1996). The compound Poisson approximation for a portfolio of dependent risks. Insur. Math. Econ. **18**, 81–85

Goovaerts, M.J., Kaas, R. (1991). Evaluating compound generalized Poisson distributions recursively. ASTIN Bull. **21**, 193–198

Goovaerts, M.J., De Vylder, F., Haezendonck, J. (1984). Insurance Premiums. North-Holland, Amsterdam

Gould, H.W. (1974). Coefficient densities for powers of Taylor and Dirichlet series. Am. Math. Mon. **81**, 3–14

Grandell, J. (1991). Aspects of Risk Theory. Springer, New York

Grandell, J. (1997). Mixed Poisson Processes. Chapman & Hall, London

Grübel, R., Hermesmeier, R. (1999). Computation of compound distributions I: Aliasing errors and exponential tilting. ASTIN Bull. **29**, 197–214

Grübel, R., Hermesmeier, R. (2000). Computation of compound distributions II: Discretization errors and Richardson extrapolation. ASTIN Bull. **30**, 309–331

Guldberg, A. (1931). On discontinuous frequency-functions and statistical series. Skand. Aktuarietidskr. **14**, 167–187

Haaf, H., Reiss, O., Schoenmakers, J. (2004). Numerically stable computations of CreditRisk+. J. Risk **6**, 1–10

Hansen, M.B. (2005). Nonparametric estimation of the stationary M/G/1 workload distribution function. In: Proceedings of the 37th Conference on Winter Simulation, Orlando, Florida, pp. 869–877

Hansen, M.B., Pitts, S.M. (2006). Nonparametric inference from the M/G/1 workload. Bernoulli **12**, 737–759

Hansen, M.B., Pitts, S.M. (2009, in press). Decompounding random sums: A nonparametric approach. Ann. Inst. Stat. Math.

Hess, K.Th., Liewald, A., Schmidt, K.D. (2002). An extension of Panjer's recursion. ASTIN Bull. **32**, 283–297

Hesselager, O. (1994). A recursive procedure for calculation of some compound distributions. ASTIN Bull. **24**, 19–32

Hesselager, O. (1996a). A recursive procedure for calculation of some mixed compound Poisson distributions. Scand. Actuar. J., 54–63

Hesselager, O. (1996b). Recursions for certain bivariate counting distributions and their compound distributions. ASTIN Bull. **26**, 35–52

Hesselager, O. (1997). Recursions for a class of compound Lagrangian distributions. Bull. Swiss Assoc. Actuar., 95–101. Correction: Bull. Swiss Assoc. Actuar., 135 (1998)

Hipp, C. (1985). Approximation of aggregate claims distributions by compound Poisson distribution. Insur. Math. Econ. **4**, 227–232

Hipp, C. (1986). Improved approximations for the aggregate claims distribution in the individual model. ASTIN Bull. **16**, 89–100

Hipp, C. (2006). Speedy convolution algorithms and Panjer recursions for phase-type distributions. Insur. Math. Econ. **38**, 176–188

Hipp, C., Michel, R. (1990). Risikotheorie: Stochastische Modelle und statistische Methoden. Schriftenreihe angewandte Versicherungsmathematik, vol. 24. Verlag Versicherungswirtschaft e.V., Karlsruhe

Horn, R.A., Steutel, F.W. (1978). On multivariate infinitely divisible distributions. Stoch. Process. Their Appl. **6**, 139–151

Hürlimann, W. (1988). An elementary proof of the Adelson–Panjer recursion formula. Insur. Math. Econ. **7**, 39–40

Hürlimann, W. (1991). Negative claim amounts, Bessel functions, linear programming and Miller's algorithm. Insur. Math. Econ. **10**, 9–20

Jewell, W.S. (1984). Approximating the distribution of a dynamic risk portfolio. ASTIN Bull. **14**, 135–148

Jewell, W.S. (2004). Bayesian statistics. In: Teugels, J.L., Sundt, B. (eds.) Encyclopedia of Actuarial Science, vol. 1, pp. 153–166. Wiley, Chichester

Jewell, W.S., Milidiu, R.L. (1986). Strategies for computation of compound distributions with two-sided severities. Insur. Math. Econ. **5**, 119–127

Jewell, W.S., Sundt, B. (1981). Improved approximations for the distribution of a heterogeneous risk portfolio. Bull. Assoc. Swiss Actuar., 221–240

Johnson, N.L., Kotz, S., Balakrishnan, N. (1994). Continuous Univariate Distributions, vol. 1, 2nd edn. Wiley, New York

Johnson, N.L., Kotz, S., Balakrishnan, N. (1995). Continuous Univariate Distributions, vol. 2, 2nd edn. Wiley, New York

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

Johnson, N.L., Kotz, S., Balakrishnan, N. (2000). Continuous Multivariate Distributions, vol. 1, 2nd edn. Wiley, New York

Johnson, N.L., Kemp, A.W., Kotz, S. (2005). Univariate Discrete Distributions, 3rd edn. Wiley–Interscience, Hoboken

Kaas, R., Goovaerts, M.J. (1985). Computing moments of compound distributions. Scand. Actuar. J., 35–38

Kaas, R., van Heerwaarden, A.E., Goovaerts, M.J. (1988a). On stop-loss premiums for the individual model. ASTIN Bull. **18**, 91–97

Kaas, R., van Heerwaarden, A.E., Goovaerts, M.J. (1988b). Between individual and collective model for the total claims. ASTIN Bull. **18**, 169–174

Kaas, R., van Heerwaarden, A.E., Goovaerts, M.J. (1989). Combining Panjer's recursion with convolution. Insur. Math. Econ. **8**, 19–21

Kaas, R., van Heerwaarden, A.E., Goovaerts, M.J. (1994). Ordering of Actuarial Risks. CAIRE, Brussels

Kaas, R., Goovaerts, M.J., Dhaene, J., Denuit, M. (2001). Modern Actuarial Risk Theory. Kluwer, Boston

Karlis, D., Xekalaki, E. (2005). Mixed Poisson distributions. Int. Stat. Rev. **73**, 35–58

Katti, S.K. (1967). Infinite divisibility of integer-valued random variables. Ann. Math. Stat. **38**, 1306–1308

Katti, S.K., Gurland, J. (1958). Some families of contagious distributions. Air Force Technical Note, Air Force Office of Scientific Research Contract AF (49)(638)-43, Technical Note No. TN 59-4 AD 208184

Katz, L. (1945). Characteristics of frequency functions defined by first-order difference equations. Dissertation, University of Michigan, Ann Arbor

Katz, L. (1965). Unified treatment of a broad class of discrete probability distributions. In: Patil, G.P. (ed.) Classical and Contagious Discrete Distributions, pp. 175–182. Pergamon Press, Oxford

Kemp, C.D. (1967). "Stuttering-Poisson" distributions. J. Stat. Soc. Inq. Soc. Irel. **21**, 151–157

Kestemont, R.M., Paris, J. (1985). Sur l'ajustement du nombre des sinstres. Bull. Assoc. Swiss Actuar., 157–164

Khatri, C.G., Patel, I.R. (1961). Three classes of univariate discrete distributions. Biometrics **17**, 567–575

Kitano, M., Shimizu, K., Ong, S.H. (2005). The generalized Charlier series distribution as a distribution with two-step recursion. Stat. Probab. Lett. **75**, 280–290

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

Knopp, K. (1990). Theory and Application of Infinite Series. Dover, New York

Kocherlakota, S., Kocherlakota, K. (1992). Bivariate Discrete Distributions. Marcel Dekker, New York

Kornya, P.S. (1983). Distribution of aggregate claims in the individual risk theory model. Trans. Soc. Actuar. **35**, 823–836. Discussion: Trans. Soc. Actuar. **35**, 837–858 (1983)

Kornya, P.S. (2007). On approximating the individual risk model. N. Am. Actuar. J. **11**, 77–98

Kuon, S., Reich, A., Reimers, L. (1987). Panjer vs Kornya vs De Pril: a comparison. ASTIN Bull. **17**, 183–191

Kuon, S., Radtke, M., Reich, A. (1993). An appropriate way to switch from the individual risk model to the collective one. ASTIN Bull. **23**, 23–54

Kupper, J. (1962). Wahrscheinlichkeitstheoretische Modelle in der Schadenversicherung. Blätter Dtsch. Ges. Versicher.math. **5**, 451–501

Kurth, A., Tasche, D. (2003). Contributions to credit risk. Risk **16**, 84–88

Lemaire, J. (2004). Bonus-malus systems. In: Teugels, J.L., Sundt, B. (eds.) Encyclopedia of Actuarial Science, vol. 1, pp. 184–191. Wiley, Chichester

Lukacs, E. (1970). Characteristic Functions. Griffin, London

Luong, A., Garrido, J. (1993). Minimum quadratic distance estimation for a parametric family of discrete distributions defined recursively. Aust. J. Stat. **35**, 59–67

Maceda, E.C. (1948). On the compound and generalized Poisson distributions. Ann. Math. Stat. **19**, 414–416

Mata, A.J. (2000). Pricing excess of loss reinsurance with reinstatements. ASTIN Bull. **30**, 349–368

Mata, A.J. (2003). Asymptotic dependence of reinsurance aggregate claim amounts. ASTIN Bull. **33**, 239–263

McNeil, A.J., Frey, R., Embrechts, P. (2005). Quantitative Risk Management. Princeton University Press, Princeton

Michel, R. (1987). An improved error bound for the compound Poisson approximation. ASTIN Bull. **17**, 165–169

Milidiu, R.L. (1985). The computation of compound distributions with two-sided severities. Doctoral Dissertation, University of California, Berkeley

Murat, M., Szynal, D. (1998). On moments of counting distributions satisfying the *k*th order recursion and their compound distributions. J. Math. Sci. **92**, 4038–4043

Murat, M., Szynal, D. (2000a). On computational formulas for densities and moments of compound distributions. J. Math. Sci. **99**, 1286–1299

Murat, M., Szynal, D. (2000b). Moments of counting distributions satisfying recurrence relations. Mat. Stosow. **1**(42), 92–129

Neuhaus, W. (2004). Experience-rating. In: Teugels, J.L., Sundt, B. (eds.) Encyclopedia of Actuarial Science, vol. 2, pp. 639–646. Wiley, Chichester

Neyman, J. (1939). On a new class of 'contagious' distributions applicable in entomology and bacteriology. Ann. Math. Stat. **10**, 35–57

Norberg, R. (2004). Credibility theory. In: Teugels, J.L., Sundt, B. (eds.) Encyclopedia of Actuarial Science, vol. 1, pp. 398–406. Wiley, Chichester

Ospina, A.V., Gerber, H.U. (1987). A simple proof of Feller's characterization of the compound Poisson distributions. Insur. Math. Econ. **6**, 63–64

Ottestad, P. (1939). On the use of factorial moments in the study of discontinuous distributions. Skand. Aktuarietidskr. **22**, 22–31

Panjer, H.H. (1980). The aggregate claims distribution and stop-loss reinsurance. Trans. Soc. Actuar. **32**, 523–535. Discussion: Trans. Soc. Actuar. **32**, 537–545 (1980)

Panjer, H.H. (1981). Recursive evaluation of a family of compound distributions. ASTIN Bull. **12**, 22–26

Panjer, H.H., Lutek, B.W. (1983). Practical aspects of stop-loss calculations. Insur. Math. Econ. **2**, 159–177

Panjer, H.H., Wang, S. (1993). On the stability of recursive formulas. ASTIN Bull. **23**, 227–258

Panjer, H.H., Wang, S. (1995). Computational aspects of Sundt's generalized class. ASTIN Bull. **25**, 5–17

Panjer, H.H., Willmot, G.E. (1982). Recursions for compound distributions. ASTIN Bull. **13**, 1–11

Panjer, H.H., Willmot, G.E. (1984). Computational techniques in reinsurance models. Trans. Int. Congr. Actuar. **22.4**, 111–120

Panjer, H.H., Willmot, G.E. (1986). Computational aspects of recursive evaluation of compound distributions. Insur. Math. Econ. **5**, 113–116. Discussion: Insur. Math. Econ. **5**, 117–118 (1986)

Panjer, H.H., Willmot, G.E. (1992). Insurance Risk Models. Society of Actuaries, Schaumburg

Papageorgiou, H. (1984). On recursive formulas for aggregate claims. Scand. Actuar. J., 102–104

Philipson, C. (1960). The theory of confluent hypergeometric functions and its application to compound Poisson processes. Skand. Aktuarietidskr. **43**, 136–162

Plackett, R.L. (1969). Stochastic models of capital investment. J. R. Stat. Soc. B **31**, 1–20. Discussion: J. R. Stat. Soc. B **31**, 20–28 (1969)

Quenouille, M.H. (1949). A relation between the logarithmic, Poisson, and negative binomial series. Biometrics **5**, 162–164

Reinhard, J.M., Snoussi, M. (2004). A monotonically converging algorithm for the severity of ruin in a discrete semi-Markov risk model. Scand. Actuar. J., 336–354

Ribas, C., Marín-Solano, J., Alegre, A. (2003). On the computation of the aggregate claims distribution in the individual life model with bivariate dependencies. Insur. Math. Econ. **32**, 201–215

Roos, B. (2005). On Hipp's compound Poisson approximations via concentration functions. Bernoulli **11**, 533–557

Ross, S.M. (1996). Stochastic Processes, 2nd edn. Wiley, New York

Ruohonen, M. (1983). Fitting a convolution of Poisson and negative binomial distributions om data. Report 118, Institute of Applied Mathematics, University of Turku

Ruohonen, M. (1988). On a model for the claim number process. ASTIN Bull. **18**, 57–68

Rytgaard, M. (2004). Stop-loss reinsurance. In: Teugels, J.L., Sundt, B. (eds.) Encyclopedia of Actuarial Science, vol. 3, pp. 1620–1625. Wiley, Chichester

Schröter, K.J. (1990). On a class of counting distributions and recursions for related compound distributions. Scand. Actuar. J., 161–175

Schröter, K.J. (1995). Verfahren zur Approximation der Gesamtschadenverteilung. Systematisierung, Techniken und Vergleiche. Verlag Versicherungswirtschaft e.V., Karlsruhe

Seal, H.L. (1971). Numerical calculation of the Bohman–Esscher family of convolution-mixed negative binomial distribution functions. Mitt. Ver. Schweiz. Versicher.math. **71**, 71–98

Sharif, A.H. (1996). Generalized recursions for total claims distributions. Dissertation, University of Waterloo

Sharif, A.H., Panjer, H.H. (1995). An improved recursion for the compound generalised Poisson distribution. Bull. Swiss Assoc. Actuar., 93–98

Sharif, A.H., Panjer, H.H. (1998). Stepwise recursions for a class of Lagrangian distributions. Actuar. Res. Clearing House, 333–364

Shimizu, R. (1962). Characterization of the normal distribution. Part II. Ann. Inst. Stat. Math. **14**, 173–178

Shiu, E. (1977). Moments for two distributions in collective risk theory. Scand. Actuar. J., 185–187

Shumway, R., Gurland, J. (1960). A fitting procedure for some generalized Poisson distributions. Skand. Aktuarietidskr. **43**, 87–108

Steutel, F.W., van Harn, K. (2004). Infinite Divisibility of Probability Distributions on the Real Line. Marcel Dekker, New York

Stroh, E. (1978). Actuarial note: the distribution function of collective risk theory as linear compounds. Actuar. Res. Clearing House

Ströter, B. (1985). The numerical evaluation of the aggregate claim density function via integral equations. Blätter Dtsch. Ges. Versicher.math. **17**, 1–14

Sundt, B. (1982). Asymptotic behaviour of compound distributions and stop-loss premiums. ASTIN Bull. **13**, 89–98. Corrigendum: ASTIN Bull. **15**, 44 (1985)

Sundt, B. (1985). On approximations for the distribution of a heterogeneous risk portfolio. Bull. Assoc. Swiss Actuar., 189–203

Sundt, B. (1986a). On stop-loss premiums and negative claim amounts. Bull. Assoc. Swiss Actuar., 89–94

Sundt, B. (1986b). Discussion on W. Hürlimann: Error bounds for stop-loss premiums calculated with the Fast Fourier Transform. Scand. Actuar. J., 114–116

Sundt, B. (1988). A recursive algorithm for convolutions of discrete uniform distributions. Insur. Math. Econ. **7**, 283–285

Sundt, B. (1991a). On excess of loss reinsurance with reinstatements. Bull. Swiss Assoc. Actuar., 51–66

Sundt, B. (1991b). On asymptotic rates on line in excess of loss reinsurance. Insur. Math. Econ. **10**, 61–67

Sundt, B. (1991c). On approximating aggregate claims distributions and stop-loss premiums by truncation. Insur. Math. Econ. **10**, 133–136

Sundt, B. (1992). On some extensions of Panjer's class of counting distributions. ASTIN Bull. **22**, 61–80

Sundt, B. (1995). On some properties of De Pril transforms of counting distributions. ASTIN Bull. **25**, 19–31

Sundt, B. (1998). A generalisation of the De Pril transform. Scand. Actuar. J., 41–48

Sundt, B. (1999a). Recursions for convolutions of discrete uniform distributions revisited. Insur. Math. Econ. **24**, 15–21

Sundt, B. (1999b). An Introduction to Non-Life Insurance Mathematics, 4th edn. Verlag Versicherungswirtschaft GmbH., Karlsruhe

Sundt, B. (1999c). On multivariate Panjer recursions. ASTIN Bull. **29**, 29–45

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

Sundt, B. (2000b). The multivariate De Pril transform. Insur. Math. Econ. **27**, 123–136

Sundt, B. (2000c). On error bounds for approximations to multivariate distributions. Insur. Math. Econ. **27**, 137–144

Sundt, B. (2000d). On multivariate Vernic recursions. ASTIN Bull. **30**, 111–122

Sundt, B. (2002). Recursive evaluation of aggregate claims distributions. Insur. Math. Econ. **30**, 297–322

Sundt, B. (2003a). Some recursions for moments of n-fold convolutions. Insur. Math. Econ. **33**, 479–486

Sundt, B. (2003b). Some recursions for moments of compound distributions. Insur. Math. Econ. **33**, 487–496

Sundt, B., Dickson, D.C.M. (2000). Comparison of methods for evaluation of the n-fold convolution of an arithmetic distribution. Bull. Swiss Assoc. Actuar., 129–140

Sundt, B., Ekuma, O. (1999). The De Pril transform of a compound $\mathcal{R}_k$ distribution. Bull. Swiss Assoc. Actuar., 179–189

Sundt, B., Jewell, W.S. (1981). Further results on recursive evaluation of compound distributions. ASTIN Bull. **12**, 27–39

Sundt, B., Vernic, R. (2002). On error bounds for approximations to multivariate distributions II. ASTIN Bull. **32**, 57–69

Sundt, B., Vernic, R. (2004). Recursions for compound mixed multivariate Poisson distributions. Blätter Dtsch. Ges. Versicher.- Finanzmath. **26**, 665–691

Sundt, B., Vernic, R. (2006). Two binomial methods for evaluating the aggregate claims distribution in De Pril's individual risk model. Belg. Actuar. Bull. **6**, 5–13

Sundt, B., Dhaene, J., De Pril, N. (1998). Some results on moments and cumulants. Scand. Actuar. J., 24–40

Szynal, D., Teugels, J.L. (1995). On moments of a class of counting distributions. Ann. Univ. Mariae Curie-Skłodowska Lublin Polonia A **49**, 199–211

Teicher, H. (1954). On the multivariate Poisson distribution. Skand. Aktuarietidskr. **37**, 1–9

Vandebroek, M., De Pril, N. (1988). Recursions for the distribution of a life portfolio: a numerical comparison. Bull. R. Assoc. Belg. Actuar. **82**, 21–35

Vernic, R. (1997). On the bivariate generalized Poisson distribution. ASTIN Bull. **27**, 23–31

Vernic, R. (1999). Recursive evaluation of some bivariate compound distributions. ASTIN Bull. **29**, 315–325

Vernic, R. (2000). A multivariate generalization of the generalized Poisson distribution. ASTIN Bull. **30**, 57–67

Vernic, R. (2002). On the evaluation of some bivariate compound Poisson distributions. Math. Rep. **4**, 311–320

Vernic, R. (2004). On a bivariate generalisation of Sundt's class of counting distributions. Math. Rep. **6**, 185–201

Vernic, R., Dhaene, J., Sundt, B. (2009). Inequalities for the De Pril approximation to the distribution of the number of policies with claims. Research Paper AFI_0928, Katholieke Universiteit Leuven.

Waldmann, K.H. (1994). On the exact calculation of the aggregate claims distribution in the individual life model. ASTIN Bull. **24**, 89–96

Waldmann, K.H. (1995). Exact calculation of the aggregate claims distribution in the individual life model by use of an n-layer model. Blätter Dtsch. Ges. Versicher.math. **22**, 279–287

Waldmann, K.H. (1996). Modified recursions for a class of compound distributions. ASTIN Bull. **26**, 213–224

Walhin, J.F. (2001). Some comments on two approximations used for the pricing of reinstatements. Bull. Swiss Assoc. Actuar., 29–47

Walhin, J.F. (2002a). On the practical multiline excess of loss pricing. In: CAS Forum, Winter 2002, pp. 121–151
Walhin, J.F. (2002b). On the optimality of multiline excess of loss covers. Belg. Actuar. Bull. **2**, 92–96
Walhin, J.F. (2002c). Some comments on the pricing of an exotic excess of loss treaty. J. Actuar. Pract. **10**, 175–191
Walhin, J.F., Paris, J. (1998). On the use of equispaced discrete distributions. ASTIN Bull. **28**, 241–255
Walhin, J.F., Paris, J. (1999). Using mixed Poisson processes in connection with bonus-malus systems. ASTIN Bull. **29**, 81–99
Walhin, J.F., Paris, J. (2000a). The effect of excess of loss reinsurance with reinstatements on the cedent's portfolio. Blätter Dtsch. Ges. Versicher.math. **24**, 615–627
Walhin, J.F., Paris, J. (2000b). Recursive formulae for some bivariate counting distributions obtained by the trivariate reduction method. ASTIN Bull. **30**, 141–155
Walhin, J.F., Paris, J. (2001a). Excess of loss reinsurance with reinstatements: premium calculation and ruin probability of the cedent. Blätter Dtsch. Ges. Versicher.math. **25**, 1–12
Walhin, J.F., Paris, J. (2001b). The mixed bivariate Hofmann distribution. ASTIN Bull. **31**, 123–138
Walhin, J.F., Paris, J. (2001c). Some comments on the individual risk model and multivariate extension. Blätter Dtsch. Ges. Versicher.math. **25**, 257–270
Walhin, J.F., Herfurth, L., De Longueville, P. (2001). The practical pricing of excess of loss treaties: actuarial, financial, economic and commercial aspects. Belg. Actuar. Bull. **1**, 40–57
Wang, S. (1994). Recurrence equations and insurance risk computing. Dissertation, University of Waterloo
Wang, S. (1995). On two-sided compound binomial distributions. Insur. Math. Econ. **17**, 35–41
Wang, S., Panjer, H.H. (1993). Critical starting points for stable evaluation of mixed Poisson probabilities. Insur. Math. Econ. **13**, 287–297
Wang, S., Panjer, H.H. (1994). Proportional convergence and tail-cutting techniques in evaluating aggregate claim distributions. Insur. Math. Econ. **14**, 129–138
Wang, S., Sobrero, M. (1994). Further results on Hesselager's recursive procedure for calculation of some compound distributions. ASTIN Bull. **24**, 161–166
White, R.P., Greville, T.N.E. (1959). On computing the probability that exactly k of n independent events will occur. Trans. Soc. Actuar. **11**, 88–95. Discussion: Trans. Soc. Actuar. **11**, 96–99 (1959)
Williams, R.E. (1980). Computing the probability function for aggregate claims. Proc. Can. Inst. Actuar. **XI**, 39–47
Willmot, G.E. (1986a). A class of counting distributions with insurance applications. Doctoral Dissertation, University of Waterloo
Willmot, G.E. (1986b). Mixed compound Poisson distributions. ASTIN Bull. **16**, S59–S79
Willmot, G.E. (1987). The Poisson-inverse Gaussian distribution as an alternative to the negative binomial. Scand. Actuar. J., 113–127
Willmot, G.E. (1988). Sundt and Jewell's family of discrete distributions. ASTIN Bull. **18**, 17–29
Willmot, G.E. (1989). Limiting tail behaviour of some discrete compound distributions. Insur. Math. Econ. **8**, 175–185
Willmot, G.E. (1990). Asymptotic tail behaviour of Poisson mixtures with applications. Adv. Appl. Probab. **22**, 147–159
Willmot, G.E. (1993). On recursive evaluation of mixed Poisson probabilities and related quantities. Scand. Actuar. J., 114–133
Willmot, G.E. (2000). On evaluation of the conditional distribution of the deficit at the time of ruin. Scand. Actuar. J., 63–79
Willmot, G.E. (2004). Thinned distributions. In: Teugels, J.L., Sundt, B. (eds.) Encyclopedia of Actuarial Science, vol. 3, pp. 1672–1673. Wiley, Chichester
Willmot, G.E., Panjer, H.H. (1987). Difference equation approaches in evaluation of compound distributions. Insur. Math. Econ. **6**, 43–56

Willmot, G.E., Sundt, B. (1989). On evaluation of the Delaporte distribution and related distributions. Scand. Actuar. J., 101–113

Willmot, G.E., Woo, J.K. (2007). On the class of Erlang mixtures with risk theoretic applications. N. Am. Actuar. J. **11**, 99–115. Discussion: N. Am. Actuar. J. **11**, 115–118 (2007)

Witdouck, S., Walhin, J.F. (2004). Some remarks on the cedent's retention risk in presence of an annual aggregate deductible or reinstatements. Blätter Dtsch. Ges. Versicher.- Finanzmath. **26**, 461–481

List of Notation

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0, © Springer-Verlag Berlin Heidelberg 2009

Author Index

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0,

Subject Index

B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications,* EAA Lecture Notes,
DOI 10.1007/978-3-540-92900-0,