Wiley Series in Probability and Statistics

LOSS MODELS

FURTHER TOPICS

STUART A. KLUGMAN · HARRY H. PANJER · GORDON E. WILLMOT

SOCIETY OF ACTUARIES

WILEY

LOSS MODELS

WILEY SERIES IN PROBABILITY AND STATISTICS

Established by WALTER A. SHEWHART and SAMUEL S. WILKS

A complete list of the titles in this series appears at the end of this volume.

LOSS MODELS

Further Topics

Stuart A. Klugman
Society of Actuaries
Schaumburg, IL

Harry H. Panjer
Department of Statistics and Actuarial Science
University of Waterloo
Ontario, Canada

Gordon E. Willmot
Department of Statistics and Actuarial Science
University of Waterloo
Ontario, Canada

SOCIETY OF ACTUARIES

WILEY

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Klugman, Stuart A., 1949–
Loss models : further topics / Stuart A. Klugman, Society of Actuaries, Schaumburg, IL, Harry H. Panjer, Department of Statistics and Actuarial Science, University of Waterloo, Ontario, Canada, Gordon E. Willmot, Department of Statistics and Actuarial Science, University of Waterloo, Ontario, Canada.
pages cm
Includes bibliographical references and index.
ISBN 978-1-118-34356-2 (cloth)
1. Insurance—Statistical methods. 2. Insurance—Mathematical models. I. Panjer, Harry H. II. Willmot, Gordon E.,– III. Title.
HG8781.K584 2013
368'.01—dc23 2013009457

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

CONTENTS

PREFACE

Readers who have been with us since the first edition may have noticed that each edition added several new topics while few were dropped. By the third edition, a limit had been reached and it was time to rethink how the material was presented. With the encouragement of our publisher, we decided to produce two books. The first, published in 2012, is the fourth edition, continuing to be called *Loss Models: From Data to Decisions* [58]. In that book we included all the topics currently covered on the examinations of the Casualty Actuarial Society and the Society of Actuaries (with some updates to specific topics). We also included a few topics we think may be worth adding in the future (and that we like to teach). When designing this companion book, we wanted to do two things. The first was to cover the topics from the third edition that had been excluded from the fourth edition. These are:

- Extreme value distributions (with expanded material on tail calculations)
- Computational methods for aggregate models [including an (a, b, m) recursion]
- Counting processes
- Copula models
- Continuous-time ruin models
- Interpolation and smoothing

The second was to add new material, particularly with regard to expanding the number of models presented and demonstrating how they apply to actuarial problems. The new topics are (though some include material that was in the third edition):

- Coxian and related distributions
- Mixed Erlang distributions
- Analytic methods for aggregate claim models
- More discrete claim count models
- Compound distributions with time dependent claim amounts

We have viewed this companion book as more of a practitioner's and researcher's resource than a textbook and thus have only created exercises where additional concepts are introduced. However, for material brought over from the third edition, those exercises have been retained. Solutions to all exercises are in an Appendix. Together with the fourth edition, we believe the two books present a comprehensive look at the current state of this aspect of actuarial work. We are thankful for the continued support and encouragement from John Wiley & Sons and the Society of Actuaries. We also thank Joan Hatton for her expert typing and Mirabelle Huynh who did a thorough job of proofreading our writing.

S. A. KLUGMAN, H. H. PANJER, G. E. WILLMOT

Schaumburg, IL and Waterloo, Ontario

CHAPTER 1

INTRODUCTION

As noted in the preface, the purpose of this book is to provide information on topics not covered in the fourth edition of *Loss Models: From Data to Decisions* [59]. In general, the emphasis here is less on data and decisions and more on what is in between, namely the vast array of models available for actuarial work. In this introduction we give a brief overview of the models covered. The material can be broken up into six sets of topics.

Univariate models for loss amounts

Three chapters are devoted to classes of univariate models. The first is the class of Coxian distributions (Chapter 2). These distributions have the desirable property that their Laplace transform (or, equivalently, their moment generating function) is a ratio of polynomials. Thus, when used as a claim size distribution, convenient explicit expressions for the associated aggregate or compound distribution may sometimes be derived. The second is the class of mixed Erlang distributions (Chapter 3). These distributions are notable because they can approximate any positive continuous distribution to an arbitrary degree of accuracy. Moreover, the mixed Erlang class contains a large number of distributions, including some whose mixed Erlang structure is not obvious. Also, calculations of most quantities of interest in an insurance loss context are computationally straightforward. The third chapter (Chapter 4) covers the two classes of extreme value distributions. This material is largely reproduced from the third edition [58] with some additional material on tail calculations.

As the name implies, these models are especially useful for management of risks that may produce large losses.

Calculation of aggregate losses

The basic methods for these calculations are covered in the fourth edition. This book contains two enhancements. Some of the univariate models introduced in the early chapters allow for exact calculation of aggregate loss probabilities. The formulas are developed in Chapter 5 along with asymptotic formulas for the right tail. Computational methods left out of the fourth edition are provided in Chapter 6. These include inversion methods, calculating with approximate distributions, and calculating from the individual risk model (which was in the second edition, but not the third). A new item is a presentation of the recursive formula when the frequency distribution is a member of the (a, b, m) class of distributions.

Loss model applications of discrete counting models

The next three chapters focus on various issues that are of interest in the loss modeling context. The first chapter (Chapter 7) introduces counting processes and, as in the third edition, deals with nonhomogeneous birth processes and mixed Poisson processes, which are useful for modeling the development of claim counts over time. Chapter 8 is new and considers properties of discrete counting distributions that are of interest in connection with loss model concepts such as deductibles and limits, recursions for compound distributions, evaluation of stop-loss moments, and computation of the risk measures VaR and TVaR in a discrete setting. The third chapter (Chapter 9) deals with models where the claim amounts depend on the time of incurral of the claim. Examples include inflation and claim payment delays.

Multivariate models

Chapter 10 covers the analysis of multivariate models based on copula functions. The material is taken from the third edition. Methods for simulation that were in a later chapter of the third edition were moved to this chapter.

Continuous-time ruin models

The material in Chapter 11 is taken directly from the third edition. It contains the classic analysis of the infinite-time ruin problem.

Interpolation and smoothing

While this material was covered in the third edition, two changes have been made for Chapter 12. First, some of the earlier material has been eliminated or streamlined. The goal is to efficiently arrive at the smoothing spline, the method most suitable for actuarial problems. More emphasis is placed on the most common application, the smoothing of data from experience studies. A traditional actuarial method, Whittaker–Henderson, has been added along with discussion of its similarity to smoothing splines.

CHAPTER 2

COXIAN AND RELATED DISTRIBUTIONS

2.1 Introduction

For the analysis of aggregate claims, the typical models involve compound distributions, which result in analytical complexities. A useful feature of compound distributions is the simplicity of the probability generating function (for discrete cases) and the Laplace transform (for continuous cases). This characteristic can be exploited to obtain useful results from either a mathematical or computational viewpoint. Because the class of Coxian distributions is defined through its Laplace transform, members of the class are well suited for use as claim amount distributions in aggregate claims models. In this chapter we briefly discuss two fairly broad classes of models that have been used in applications involving loss models. Both are subclasses of the class of Coxian distributions, which we now define.

Definition 2.1 *A distribution is from the* **Coxian-n class** *if its Laplace transform* $\tilde{f}(s) = \int_0^\infty e^{-sx} f(x)dx$ *may be expressed as*

$$\tilde{f}(s) = \frac{a(s)}{\prod_{i=1}^{m} (\lambda_i + s)^{n_i}}, \tag{2.1}$$

where $\lambda_i > 0$ *for* $i = 1, 2, \ldots, m$ *and (without loss of generality) we assume that* $\lambda_i \neq \lambda_j$ *for* $i \neq j$. *We further assume that* n_i *is a nonnegative integer for* $i = 1, 2, \ldots, m$ *and that* $n = \sum_{i=1}^{m} n_i > 0$. *Also,* $a(s)$ *is a polynomial of degree* $n - 1$ *or less.*

As $\tilde{f}(0) = 1$, it follows that $a(0) = \prod_{i=1}^{m} \lambda_i^{n_i}$. Furthermore, a partial fraction expansion of (2.1) yields

$$\tilde{f}(s) = \sum_{i=1}^{m} \sum_{j=1}^{n_i} p_{ij} \left(\frac{\lambda_i}{\lambda_i + s} \right)^j, \tag{2.2}$$

where

$$p_{ij} = \frac{\lambda_i^{-j}}{(n_i - j)!} \frac{d^{n_i - j}}{ds^{n_i - j}} \left\{ \prod_{\substack{k=1 \\ k \neq i}}^{m} \frac{a(s)}{(\lambda_k + s)^{n_k}} \right\} \Bigg|_{s=-\lambda_i} . \tag{2.3}$$

Note that

$$\left(\frac{\lambda_i}{\lambda_i + s} \right)^j = \int_0^\infty e^{-sx} e_{\lambda_i, j}(x) dx, \tag{2.4}$$

where

$$e_{\lambda_i, j}(x) = \frac{\lambda_i^j x^{j-1} e^{-\lambda_i x}}{(j-1)!}, \quad x \geq 0, \tag{2.5}$$

is the probability density function (pdf) of an Erlang-j random variable with scale parameter λ_i. Then, from (2.2) and also (2.4), the Coxian-n class has pdf of the form

$$f(x) = \sum_{i=1}^{m} \sum_{j=1}^{n_i} p_{ij} e_{\lambda_i, j}(x)$$

with $e_{\lambda_i, j}(x)$ given by (2.5), which is a finite combination of Erlang pdfs. We now discuss the special case when $n_i = 1$ for $i = 1, 2, \ldots, m$.

2.2 Combinations of exponentials

Suppose that X has pdf of the form

$$f(x) = \sum_{i=1}^{n} q_i \lambda_i e^{-\lambda_i x}, \quad x \geq 0, \tag{2.6}$$

where $\lambda_i \geq 0$ for $i = 1, 2, \ldots, n$. The condition that $\int_0^\infty f(x) dx = 1$ implies that $\sum_{i=1}^{n} q_i = 1$, and if $0 \leq q_i \leq 1$ for $i = 1, 2, \ldots, n$ then (2.6) may be interpreted as a mixture of exponential pdfs.

But (2.6) is a pdf in many cases even if some of the q_i are negative, in which case (2.6) may be referred to as a combination (or a generalized mixture) of exponentials. Two examples where this occurs are now given.

EXAMPLE 2.1 Generalized Erlang distribution

Suppose that X_i has the exponential distribution with mean $1/\lambda_i$ for $i = 1, 2, \ldots, n$, where $\lambda_i \neq \lambda_j$. Let $S_n = X_1 + \cdots + X_n$. Then S_n is said to have a **generalized Erlang** pdf where it is further assumed that $X_1, X_2, \ldots, X_n$ are independent. Clearly,

$$\mathrm{E}\left(e^{-sS_n}\right) = \prod_{i=1}^{n} \left(\frac{\lambda_i}{\lambda_i + s} \right),$$

which is of the form (2.1) with $n_i = 1$ for $i = 1, 2, \ldots, m$. Then a partial fraction expansion yields immediately that

$$\prod_{j=1}^{n} \left(\frac{\lambda_j}{\lambda_j + s} \right) = \sum_{j=1}^{n} q_j \frac{\lambda_j}{\lambda_j + s}. \tag{2.7}$$

Thus, for $i = 1, 2, \ldots, n$,

$$q_i \lambda_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} q_j \lambda_j \frac{\lambda_i + s}{\lambda_j + s} = \lambda_i \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{\lambda_j}{\lambda_j + s},$$

and substitution of $s = -\lambda_i$ yields

$$q_i = \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{\lambda_j}{\lambda_j - \lambda_i}, \quad i = 1, 2, \ldots, n. \tag{2.8}$$

We remark that (2.7) and (2.8) also follow directly from (2.2) and (2.3), respectively. Thus, from (2.7), S_n has pdf

$$f_{S_n}(x) = \sum_{i=1}^{n} q_i \lambda_i e^{-\lambda_i x}, \tag{2.9}$$

where q_i is given by (2.8) and (2.9) is of the form (2.6). □

The use of the partial fraction expansion in the previous example is essentially equivalent to Lagrange's polynomial representation. That is, if $x_1, x_2, \ldots, x_n$ are distinct numbers and $g(x)$ is a polynomial of degree $n-1$ or less, then $g(x)$ may be expressed in terms of the functional values $g(x_i)$ for $i = 1, 2, \ldots, n$ as

$$g(x) = \sum_{i=1}^{n} g\left(x_i\right) \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j}. \tag{2.10}$$

An important special case of (2.10) is when $g(x) = 1$, yielding the identity

$$\sum_{i=1}^{n} \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j} = 1. \tag{2.11}$$

If $x = 0$ in (2.11) and $x_j = \lambda_j$, it follows immediately that $\sum_{i=1}^{n} q_i = 1$, where q_i is given by (2.8), a condition necessary for (2.9) to be a pdf. Also, (2.11) may be viewed as a polynomial identity in x. The left-hand side is a polynomial of degree $n-1$ in x, with coefficient of x^{n-1} satisfying

$$\sum_{i=1}^{n} \left\{ \prod_{\substack{j=1 \\ j \neq i}}^{n} \left(x_i - x_j\right) \right\}^{-1} = \sum_{i=1}^{n} \left\{ \prod_{\substack{j=1 \\ j \neq i}}^{n} \left(x_j - x_i\right) \right\}^{-1} = 0, \tag{2.12}$$

because the coefficient of x^{n-1} on the right-hand side of (2.11) is 0. This proves the outer equality in (2.12), and the left-hand equality in (2.12) follows by multiplication by $(-1)^{n-1}$. It is clear from (2.8) and (2.9) that

$$f_{S_n}(0) = \sum_{i=1}^{n} q_i \lambda_i = \left\{ \prod_{j=1}^{n} \lambda_j \right\} \sum_{i=1}^{n} \left\{ \prod_{\substack{j=1 \\ j \neq i}}^{n} (\lambda_j - \lambda_i) \right\}^{-1},$$

and thus (2.12) with $x_j = \lambda_j$ implies that $f_{S_n}(0) = 0$ if $n = 2, 3, \ldots$.

We now consider a second example of a combination of exponentials.

EXAMPLE 2.2 A logbeta distribution

Suppose that Y has the beta pdf

$$f_Y(y) = \frac{\Gamma(n+\alpha+1)}{\Gamma(\alpha+1)(n-1)!} y^{\alpha}(1-y)^{n-1}, \quad 0 < y < 1,$$

where $\alpha > -1$ and n is a positive integer. Now consider the random variable X defined by $Y = e^{-\lambda X}$, so that $X = -\frac{1}{\lambda}\ln Y$. Thus the cdf of X is

$$F_X(x) = \Pr\left(-\frac{1}{\lambda}\ln Y \leq x\right) = \Pr(-\lambda x \leq \ln Y) = \Pr\left(Y > e^{-\lambda x}\right),$$

and differentiation yields the pdf

$$f_X(x) = \lambda e^{-\lambda x} f_Y\left(e^{-\lambda x}\right) = \lambda e^{-\lambda x} \frac{\Gamma(n+\alpha+1)}{\Gamma(\alpha+1)(n-1)!} \left(e^{-\lambda x}\right)^{\alpha} \left(1 - e^{-\lambda x}\right)^{n-1}.$$

Noting that $\Gamma(n+\alpha+1)/\Gamma(\alpha+1) = \prod_{j=1}^{n}(\alpha+j)$ and applying a binomial expansion to $(1-e^{-\lambda x})^{n-1}$ yield

$$f_X(x) = \left\{ \prod_{j=1}^{n}(\alpha+j) \right\} \lambda e^{-\lambda(\alpha+1)x} \sum_{k=0}^{n-1} \frac{(-1)^k}{k!(n-1-k)!} e^{-\lambda k x}.$$

A change in the index of summation from k to $i = k+1$ yields

$$\begin{aligned} f_X(x) &= \left\{ \prod_{j=1}^{n}(\alpha+j) \right\} \lambda e^{-\lambda \alpha x} \sum_{i=1}^{n} \frac{(-1)^{i-1}}{(i-1)!(n-i)!} e^{-\lambda i x} \\ &= \sum_{i=1}^{n} \frac{(-1)^{i-1}}{(i-1)!(n-i)!} \left\{ \prod_{\substack{j=1 \\ j \neq i}}^{n} (\alpha+j) \right\} \lambda(\alpha+i) e^{-\lambda(\alpha+i)x}, \end{aligned}$$

that is,

$$f_X(x) = \sum_{i=1}^{n} q_i \lambda_i e^{-\lambda_i x}, \tag{2.13}$$

where $\lambda_i = \lambda(\alpha + i)$ and

$$q_i = \frac{(-1)^{i-1}}{(i-1)!(n-i)!} \prod_{\substack{j=1 \\ j \neq i}}^{n} (\alpha + j), \quad i = 1, 2, \ldots, n. \tag{2.14}$$

It is useful to note that if $i = 1, 2, \ldots, n$,

$$\begin{aligned}\prod_{\substack{j=1 \\ j \neq i}}^{n} (j - i) &= \{(1-i)(2-i)\ldots(i-1-i)\}\{(i+1-i)(i+2-i)\cdots(n-i)\} \\ &= \left\{(-1)^{i-1}(i-1)!\right\}\{(n-i)!\} \\ &= (i-1)!(n-i)!/(-1)^{i-1},\end{aligned}$$

and thus (2.14) may also be expressed as

$$q_i = \prod_{\substack{j=1 \\ j \neq i}}^{n} \left(\frac{j+\alpha}{j-i}\right). \tag{2.15}$$

Now, with $x_i = -i$, (2.10) becomes

$$g(x) = \sum_{i=1}^{n} g(-i) \prod_{\substack{j=1 \\ j \neq i}}^{n} \left(\frac{x+j}{j-i}\right),$$

implying from (2.15) that $g(\alpha) = \sum_{i=1}^{n} q_i g(-i)$ for any polynomial $g(x)$ of degree $n-1$ or less. Thus, with $g(x) = 1$, it follows that $\sum_{i=1}^{n} q_i = 1$, a condition that again must hold for (2.13) to be a pdf. □

The class of combinations of exponentials is an important class of distributions as it is dense in the set of probability distributions on $[0, \infty)$, implying that any such probability distribution may be approximated by a combination of exponentials. Dufresne [19] considers this approximation problem and uses logbeta pdfs of the type considered in Example 2.2 in this context. Interestingly, the terminology "logbeta" is also due to Dufresne [19] and is more appropriate than the use of the term "lognormal" in that the log (not the exponential) of a lognormal random variable is normally distributed.

Assuming without loss of generality that $\lambda_1 < \lambda_2 < \cdots < \lambda_n$, necessary conditions for (2.6) to be a valid pdf are that $q_1 > 0$ and $\sum_{i=1}^{n} q_i \lambda_i \geq 0$, and these conditions are also sufficient if there is not more than one sign change in the sequence $\{q_1, q_2, \ldots, q_n\}$, obviously the case if $n = 2$. See Steutel and van Harn [89, pp. 338–339] for further details. Again assuming that $\lambda_1 < \lambda_2 < \cdots < \lambda_n$, Bartholomew [7] shows that alternative sufficient conditions for (2.6) to be a valid pdf are that $\sum_{i=1}^{k} q_i \lambda_i \geq 0$ for $k = 1, 2, \ldots, n$.

2.3 Coxian-2 distributions

In the Coxian-n case with $n = 2$, $a(s)$ is a linear function of s, and thus (2.1) may be expressed as

$$\tilde{f}(s) = \frac{\lambda_1(1-p)s + \lambda_1\lambda_2}{(\lambda_1 + s)(\lambda_2 + s)}, \tag{2.16}$$

where $\lambda_1 > 0$, $\lambda_2 > 0$, and $\lambda_1 = \lambda_2$ is possible. We wish to consider values of p for which (2.16) is the Laplace transform of a pdf. First, note that if $p = 0$ then $\tilde{f}(s) = \lambda_1/(\lambda_1 + s)$, which is the Laplace transform of an exponential (with mean $1/\lambda_1$) pdf. Similarly, if $p = 1 - \lambda_2/\lambda_1$, that is, $\lambda_2 = \lambda_1(1-p)$, then $\tilde{f}(s) = \lambda_2/(\lambda_2 + s)$, again of exponential form. Thus we exclude the cases with $p = 0$ and $p = 1 - \lambda_2/\lambda_1$ in what follows.

It is clear from (2.16) that

$$\tilde{f}(s) = (1-p)\frac{\lambda_1}{s+\lambda_1} + p\frac{\lambda_1\lambda_2}{(s+\lambda_1)(s+\lambda_2)}, \tag{2.17}$$

which implies that

$$f(x) = \lambda_1(1-p)e^{-\lambda_1 x} + \lambda_1\lambda_2 p e^{-\lambda_1 x}h(x), \tag{2.18}$$

where

$$h(x) = \int_0^x e^{(\lambda_1-\lambda_2)y}dy. \tag{2.19}$$

Clearly, $h(x)$ is easy to evaluate, but its form depends on whether $\lambda_1 = \lambda_2$ or not. In any event, $h(0) = 0$ from (2.19), implying from (2.18) that $f(0) = \lambda_1(1-p)$, and so the condition $p \le 1$ is required for $f(x)$ to be a valid pdf. The Laplace transform of the tail $\overline{F}(x) = \int_x^\infty f(y)dy$ is, from (2.17),

$$\begin{aligned}\frac{1-\tilde{f}(s)}{s} &= \frac{(1-p)}{s}\left\{1 - \frac{\lambda_1}{\lambda_1+s}\right\} + \frac{p}{s}\left\{1 - \frac{\lambda_1\lambda_2}{(\lambda_1+s)(\lambda_2+s)}\right\} \\ &= (1-p)\left\{\frac{1-\frac{\lambda_1}{\lambda_1+s}}{s}\right\} + p\left\{\frac{1-\frac{\lambda_1}{\lambda_1+s}}{s} + \frac{\lambda_1}{\lambda_1+s}\,\frac{1-\frac{\lambda_2}{\lambda_2+s}}{s}\right\},\end{aligned}$$

from which it follows that

$$\overline{F}(x) = (1-p)e^{-\lambda_1 x} + p\left\{e^{-\lambda_1 x} + \lambda_1\int_0^x e^{-\lambda_1(x-y)-\lambda_2 y}dy\right\}.$$

Thus, again using (2.19),

$$\overline{F}(x) = e^{-\lambda_1 x}\left\{1 + p\lambda_1 h(x)\right\}, \quad x \ge 0. \tag{2.20}$$

If $\lambda_1 \ge \lambda_2$ then from (2.19) $\lim_{x\to\infty} h(x) = \infty$, and thus from (2.20) it is clear that $p \ge 0$ because if $p < 0$ then $e^{\lambda_1 x}\overline{F}(x)$ would become negative for large x. But it was assumed that $p \neq 0$, and thus if $\lambda_1 \ge \lambda_2$ it follows that $0 < p \le 1$. Thus if $\lambda_1 = \lambda_2 = \lambda$, (2.18) and (2.19) yield

$$f(x) = (1-p)\lambda e^{-\lambda x} + p\lambda^2 x e^{-\lambda x}, \tag{2.21}$$

which is the pdf of the mixture of two Erlang pdfs, both with the same scale parameter λ. We remark that pdfs of the form (2.21) will be discussed in much detail later.

If $\lambda_1 < \lambda_2$ then from (2.19)

$$h(x) = \frac{1 - e^{-(\lambda_2-\lambda_1)x}}{\lambda_2 - \lambda_1},$$

and thus from (2.20)

$$\lim_{x\to\infty} e^{\lambda_1 x}\overline{F}(x) = 1 + p\lambda_1 \lim_{x\to\infty} h(x) = 1 + p\frac{\lambda_1}{\lambda_2-\lambda_1}.$$

This limit obviously cannot be negative, and it follows that $\lambda_2 - \lambda_1 + p\lambda_1 \geq 0$, i.e., $p \geq 1 - \lambda_2/\lambda_1$, which is equivalent to $\lambda_2 \geq \lambda_1(1-p)$. But again it is assumed that $p \neq 1 - \lambda_2/\lambda_1$, and therefore if $\lambda_1 < \lambda_2$ then $1 - \lambda_2/\lambda_1 < p \leq 1$ but $p \neq 0$.

If $\lambda_1 \neq \lambda_2$ then

$$\frac{\lambda_1\lambda_2}{(s+\lambda_1)(s+\lambda_2)} = \frac{\lambda_2}{\lambda_2-\lambda_1}\frac{\lambda_1}{s+\lambda_1} + \frac{\lambda_1}{\lambda_1-\lambda_2}\frac{\lambda_2}{s+\lambda_2}, \tag{2.22}$$

which follows directly or from (2.7) and (2.8). Substitution of (2.22) into (2.17) yields

$$\begin{aligned}\tilde{f}(s) &= \left(1 - p + p\frac{\lambda_2}{\lambda_2-\lambda_1}\right)\frac{\lambda_1}{s+\lambda_1} + \left(p\frac{\lambda_1}{\lambda_1-\lambda_2}\right)\frac{\lambda_2}{s+\lambda_2}\\ &= \left(1 - p\frac{\lambda_1}{\lambda_1-\lambda_2}\right)\frac{\lambda_1}{s+\lambda_1} + \left(p\frac{\lambda_1}{\lambda_1-\lambda_2}\right)\frac{\lambda_2}{s+\lambda_2}.\end{aligned}$$

That is, if $\lambda_1 \neq \lambda_2$,

$$f(x) = (1-\alpha)\lambda_1 e^{-\lambda_1 x} + \alpha\lambda_2 e^{-\lambda_2 x}, \tag{2.23}$$

where

$$\alpha = p\frac{\lambda_1}{\lambda_1-\lambda_2}. \tag{2.24}$$

If $0 < p \leq 1$ then (2.23) is either a mixture or a combination of two exponential pdfs. However, if $p < 0$ then one must have $\lambda_1 < \lambda_2$, and $\alpha > 0$ from (2.24). But if p is negative, one must have $1 - \lambda_2/\lambda_1 < p$, or equivalently $\lambda_1 - \lambda_2 < \lambda_1 p$, and because $\lambda_1 - \lambda_2$ must also be negative, $1 > \lambda_1 p/(\lambda_1 - \lambda_2)$, i.e., $\alpha < 1$. Thus, if $p < 0$ then $0 < \alpha < 1$ and (2.23) is a mixture.

To summarize, when $\lambda_1 \neq \lambda_2$, the pdf $f(x)$ is given by (2.23) with α given by (2.24). If $p > 0$ then (2.23) is either a mixture or a combination of two exponential pdfs, whereas if $p < 0$ then (2.23) is a mixture.

Again from (2.17)

$$\begin{aligned}\frac{1-\tilde{f}(s)}{s} &= \frac{1}{s}\left\{1 - (1-p)\frac{\lambda_1}{s+\lambda_1} - p\frac{\lambda_1\lambda_2}{(s+\lambda_1)(s+\lambda_2)}\right\}\\ &= \frac{1}{s}\left\{(1-p)\left(1 - \frac{\lambda_1}{s+\lambda_1}\right) + p\left(1 - \frac{\lambda_1\lambda_2}{(s+\lambda_1)(s+\lambda_2)}\right)\right\}\\ &= \frac{1-p}{s+\lambda_1} + p\frac{s+\lambda_1+\lambda_2}{(s+\lambda_1)(s+\lambda_2)}\\ &= \frac{(1-p)(s+\lambda_2) + p(s+\lambda_2+\lambda_1)}{(s+\lambda_1)(s+\lambda_2)},\end{aligned}$$

that is,

$$\int_0^\infty e^{-sx}\overline{F}(x)dx = \frac{1-\tilde{f}(s)}{s} = \frac{s+\lambda_2+\lambda_1 p}{(s+\lambda_1)(s+\lambda_2)}. \tag{2.25}$$

With $s = 0$, (2.25) gives the mean, namely,

$$\int_0^\infty \overline{F}(x)dx = \frac{\lambda_2+\lambda_1 p}{\lambda_1\lambda_2}. \tag{2.26}$$

The equilibrium pdf $f_e(x) = \frac{\overline{F}(x)}{\int_0^\infty \overline{F}(y)dy}$ thus has Laplace transform, from (2.25) and (2.26), given by

$$\begin{aligned}\tilde{f}_e(s) &= \frac{s + \lambda_2 + \lambda_1 p}{(s+\lambda_1)(s+\lambda_2)} \frac{\lambda_1 \lambda_2}{\lambda_2 + \lambda_1 p} \\ &= \frac{\lambda_1 \left(\frac{\lambda_2}{\lambda_2 + \lambda_1 p}\right) s + \lambda_1 \lambda_2}{(s+\lambda_1)(s+\lambda_2)}.\end{aligned}$$

That is,

$$\tilde{f}_e(s) = \frac{\lambda_1(1-p_e)s + \lambda_1\lambda_2}{(s+\lambda_1)(s+\lambda_2)}, \tag{2.27}$$

where

$$p_e = \frac{\lambda_1 p}{\lambda_2 + \lambda_1 p}. \tag{2.28}$$

Comparison of (2.27) with (2.16) implies immediately (by the uniqueness of the Laplace transform) that $f_e(x)$ is again a Coxian-2 pdf, but with p replaced by p_e from (2.28).

CHAPTER 3

MIXED ERLANG DISTRIBUTIONS

3.1 Introduction

It is clear from the previous chapter that it is very difficult to get explicit closed form expressions for quantities associated with the aggregate claims distribution on a portfolio of insurance business. While there are various numerical methods available, it is nevertheless convenient to utilize analytic techniques when possible. Of course, there is always a tradeoff between mathematical simplicity on the one hand and realistic modeling on the other.

An alternative approach that addresses many of these issues may be referred to as a partially parametric or semiparametric approach. Clearly, the exponential distribution has many attractive mathematical properties in the present context but is not terribly realistic as a model for claims in many situations. It is possible, however, to capitalize on these properties in more complex models.

In the present chapter, we consider the class of mixed Erlang distributions. There are various reasons why the use of mixed Erlang distributions is of interest. First, mixed Erlang distributions are extremely flexible in terms of distributional shape. In fact, the mixed Erlang class is dense in the set of positive continuous distributions [96, pp. 163–164] so that any such distribution may be approximated arbitrarily accurately by a member of the mixed Erlang class. Second, the mixed Erlang class is extremely large and includes many distributions as members, despite the fact that their membership in the mixed Erlang class is not at all obvious at first glance. Third, many quantities of interest in connection with aggregate claims and stop-loss analysis are easily computable under the mixed Erlang assumption. Lee and Lin [62] discuss fitting of mixed Erlang distributions to data using maximum likelihood estimation and the expectation–maximization (EM) algorithm.

Two other classes of distributions may also be viewed as semiparametric, namely the class of combinations of exponentials discussed in Section 2.2 and the class of phase-type distributions introduced in the queueing theoretic community. Both of these classes are also dense in the class of positive continuous distributions, and, in fact, the class of phase-type distributions is a subset of the mixed Erlang class [85]. Perhaps not surprisingly, these two classes are also generalizations of the exponential distribution. While both of these are also useful in various situations, the infinite series expansion methodology in the mixed Erlang case has the advantage of avoiding the location of roots needed for the partial fraction expansions typically used with combinations of exponentials and also avoids the determination of eigenvalues needed for evaluation of matrix-exponentials in the phase-type case.

We will understand X to have a mixed Erlang distribution (we will not give a formal definition as it is possible to have Erlang mixtures over both Erlang parameters) if it has a pdf for $x > 0$ of the form

$$f(x) = \sum_{n=1}^{\infty} q_n \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} = \sum_{n=1}^{\infty} q_n e_{\lambda,n}(x), \tag{3.1}$$

where $e_{\lambda,n}(x)$ is the Erlang-n pdf (2.5). For mathematical convenience, we assume that X may also have a discrete mass point q_0 at 0 and that the "mixing weights" form a discrete counting distribution with probability generating function (pgf)

$$Q(z) = \sum_{n=0}^{\infty} q_n z^n. \tag{3.2}$$

Thus, using (2.4) and (3.1), X has a mixed Erlang distribution if its Laplace transform is

$$\mathrm{E}\left(e^{-sX}\right) = \sum_{n=0}^{\infty} q_n \left(\frac{\lambda}{\lambda+s}\right)^n = Q\left(\frac{\lambda}{\lambda+s}\right), \tag{3.3}$$

where $Q(z)$ is given by (3.2).

In many modeling applications, the distribution $\{q_0, q_1, \ldots\}$ is finite, so that $q_{r+k} = 0$ for some r and $k = 0, 1, 2, \ldots$. Also, if $q_0 = 0$ then X has a continuous distribution.

It is also useful to note that (3.3) reveals that a mixed Erlang distribution may also be viewed as a compound distribution where the primary distribution has pgf $Q(z)$ and the secondary distribution is exponential with mean $1/\lambda$. For this reason, a mixed Erlang distribution may serve as a model for aggregate claims but is also well suited for use as a claim size distribution. In this context, it will be seen that it is often not at all restrictive to assume that there is no parametric structure to the distribution $\{q_0, q_1, \ldots\}$. The term "semiparametric" stems from the possibly nonparametric nature of $\{q_0, q_1, \ldots\}$ coupled with the parametric exponential assumption, combined via (3.3).

3.2 Members of the mixed Erlang class

It is clear from (3.1) with $q_1 = 1$ that the exponential distribution is a member of the mixed Erlang class. More generally, the Erlang-r distribution is the special case $q_r = 1$.

Tijms [96, pp. 358–359] discusses two special cases. For both, $q_k \in (0,1)$ has an arbitrary value for some k. One case then has $q_1 = 1 - q_k$ and the other has $q_{k-1} = 1 - q_k$. As mentioned previously, the phase-type distributions are members of the mixed Erlang class. The following well-known distribution is also a member of the class.

EXAMPLE 3.1 Noncentral chi-squared with even degrees of freedom

It is well-known [52, p. 437] that the Laplace transform of the noncentral chi-squared distribution with noncentrality parameter μ and $2m$ degrees of freedom is given by

$$\begin{aligned}\tilde{f}(s) &= (1+2s)^{-m} e^{-\frac{\mu s}{1+2s}} \\ &= (1+2s)^{-m} e^{\frac{\mu}{2}\left(\frac{1}{1+2s}-1\right)} \\ &= \sum_{j=0}^{\infty} \frac{\left(\frac{\mu}{2}\right)^j e^{-\frac{\mu}{2}}}{j!} (1+2s)^{-(m+j)}.\end{aligned}$$

That is,

$$\tilde{f}(s) = \sum_{n=m}^{\infty} \frac{\left(\frac{\mu}{2}\right)^{n-m} e^{-\frac{\mu}{2}}}{(n-m)!} \left(\frac{\frac{1}{2}}{\frac{1}{2}+s}\right)^n,$$

which is of the form (3.3) with $\lambda = 1/2$ and $q_n = (\frac{\mu}{2})^{n-m} e^{-\frac{\mu}{2}}/(n-m)!$ for $n = m, m+1, \ldots$ and $q_n = 0$ otherwise. Thus the mixing weights are of Poisson form (with mean $\mu/2$) shifted to the right by m. □

The mixed Erlang examples discussed to this point are "obvious" in the sense that the mixing weights $q_0, q_1, \ldots$ are of a simple form. But many other distributions are of mixed Erlang form with more complicated mixing weights. A key observation in this regard is (5.21) from Example 5.2. Let λ in (5.21) be replaced by λ_i and $\lambda/(1-q)$ by λ. A key observation in this regard is the algebraic identity

$$\frac{\lambda_i}{\lambda_i + s} = \frac{\lambda}{\lambda + s} \frac{\frac{\lambda_i}{\lambda}}{1 - \left(1 - \frac{\lambda_i}{\lambda}\right) \frac{\lambda}{\lambda+s}}, \tag{3.4}$$

a relationship which is of interest when $0 < \lambda_i \leq \lambda < 0$. We note that the right-hand side of (3.4) is of the form (3.3) with

$$Q(z) = z \left\{ \frac{\lambda_i}{\lambda} \Big/ \left[1 - \left(1 - \frac{\lambda_i}{\lambda}\right) z\right] \right\},$$

which for $\lambda_i < \lambda$ is the pgf of a zero-truncated geometric distribution. Thus (3.4) expresses (in Laplace transform form) the exponential distribution as a compound zero-truncated geometric distribution with a different exponential secondary distribution. For distributions whose Laplace transform is a function of $\lambda_i/(\lambda_i + s)$ for different values of λ_i, it is often possible to "change" to a common value of λ using (3.4) and hence express the Laplace transform in the mixed Erlang form (3.3). The following example illustrates this idea.

EXAMPLE 3.2 Exponential mixtures with a finite range

Consider the exponential mixture with pdf

$$f(x) = \int_0^\lambda \mu e^{-\mu x} dB(\mu), \quad x > 0, \tag{3.5}$$

where $B(\mu)$ is a cumulative distribution function (cdf) (discrete or continuous) satisfying $B(0) = 0$ and $B(\lambda) = 1$. The Laplace transform of (3.5) is

$$\tilde{f}(s) = \int_0^\lambda \frac{\mu}{\mu + s} dB(\mu),$$

which may be expressed using (3.4) with $\lambda_i = \mu$ as

$$\tilde{f}(s) = \int_0^\lambda \frac{\lambda}{\lambda + s} \left[\frac{\frac{\mu}{\lambda}}{1 - \left(1 - \frac{\mu}{\lambda}\right) \frac{\lambda}{\lambda + s}} \right] dB(\mu).$$

That is, $\tilde{f}(s) = Q(\frac{\lambda}{\lambda+s})$ where

$$Q(z) = \int_0^\lambda \frac{\frac{\mu}{\lambda} z}{1 - \left(1 - \frac{\mu}{\lambda}\right) z} dB(\mu), \tag{3.6}$$

a relation of the form (3.3). It is clear that (3.6) is the pgf of a mixture of zero-truncated geometric pgfs. As (3.6) may be expressed as

$$Q(z) = \int_0^\lambda \left[\sum_{n=1}^\infty \frac{\mu}{\lambda} \left(1 - \frac{\mu}{\lambda}\right)^{n-1} z^n \right] dB(\mu),$$

it follows by comparison with (3.2) that $q_0 = 0$ and

$$q_n = \int_0^\lambda \frac{\mu}{\lambda} \left(1 - \frac{\mu}{\lambda}\right)^{n-1} dB(\mu), \quad n = 1, 2, \ldots . \tag{3.7}$$

Thus, by the uniqueness of the Laplace transform, the pdf $f(x)$ in (3.5) may be re-expressed as (3.1) with q_n given by (3.7). We note that (3.7) is particularly simple in the mixed exponential–beta model. That is, if

$$B'(\mu) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{1}{\lambda} \left(\frac{\mu}{\lambda}\right)^{\alpha-1} \left(1 - \frac{\mu}{\lambda}\right)^{\beta-1}, \quad 0 < \mu < \lambda,$$

then

$$\begin{aligned} q_n &= \alpha \frac{\Gamma(\alpha + \beta)\Gamma(\beta + n - 1)}{\Gamma(\alpha + \beta + n)\Gamma(\beta)} \\ &= \frac{\alpha}{\beta + n - 1} \prod_{j=0}^{n-1} \left(\frac{\beta + j}{\alpha + \beta + j} \right). \end{aligned}$$

□

The following example illustrates why it is usually not necessary to generalize (3.1) by allowing for countable (or even unbounded) mixtures over the Erlang scale parameter.

EXAMPLE 3.3 A double Erlang mixture

Consider the "generalization" of (3.1) given by

$$f(x) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} q_{ij} \frac{\lambda_i^j x^{j-1} e^{-\lambda_i x}}{(j-1)!}, \quad x > 0, \tag{3.8}$$

where $q_{ij} \geq 0$ for all i and j and $\sup_i \lambda_i < \infty$. Then let $\lambda \geq \sup_i \lambda_i$, and using (3.4), the Laplace transform of (3.8) may be expressed as

$$\begin{aligned} \tilde{f}(s) &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} q_{ij} \left(\frac{\lambda_i}{\lambda_i + s}\right)^j \\ &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} q_{ij} \left(\frac{\lambda}{\lambda + s}\right)^j \left[\frac{\frac{\lambda_i}{\lambda}}{1 - \left(1 - \frac{\lambda_i}{\lambda}\right)\frac{\lambda}{\lambda+s}}\right]^j. \end{aligned}$$

That is, $\tilde{f}(s) = Q(\frac{\lambda}{\lambda+s})$, where

$$Q(z) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} q_{ij} \left[\frac{\frac{\lambda_i}{\lambda} z}{1 - \left(1 - \frac{\lambda_i}{\lambda}\right) z}\right]^j. \tag{3.9}$$

It is not hard to see that (3.9) has the form of a mixture of shifted Pascal pgfs. Thus the double-mixture pdf (3.8) is actually a "single" mixture of Erlangs of the form (3.1), as long as $\sup_i \lambda_i < \infty$. This must be the case if there are a finite number of λ_is in (3.8) . To identify the mixing weights, the pgf (3.9) may be expressed as

$$Q(z) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} q_{ij} \sum_{k=0}^{\infty} \binom{j+k-1}{j-1} \left(\frac{\lambda_i}{\lambda}\right)^j \left(1 - \frac{\lambda_i}{\lambda}\right)^k z^{j+k}.$$

A change in the third index of summation from k to $n = j + k$ yields

$$\begin{aligned} Q(z) &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} q_{ij} \sum_{n=j}^{\infty} \binom{n-1}{j-1} \left(\frac{\lambda_i}{\lambda}\right)^j \left(1 - \frac{\lambda_i}{\lambda}\right)^{n-j} z^n \\ &= \sum_{i=1}^{\infty}\sum_{n=1}^{\infty} z^n \sum_{j=1}^{n} q_{ij} \binom{n-1}{j-1} \left(\frac{\lambda_i}{\lambda}\right)^j \left(1 - \frac{\lambda_i}{\lambda}\right)^{n-j}, \end{aligned}$$

that is, $Q(z) = \sum_{n=1}^{\infty} q_n z^n$ where

$$q_n = \sum_{i=1}^{\infty}\sum_{j=1}^{n} q_{ij} \binom{n-1}{j-1} \left(\frac{\lambda_i}{\lambda}\right)^j \left(1 - \frac{\lambda_i}{\lambda}\right)^{n-j}, \quad n = 1, 2, 3, \ldots,$$

which is a finite sum if there are a finite number of λ_is. To summarize, the pdf (3.8) is actually of the form (3.1) with q_n as given. □

We next present a general example of a distribution that includes various special cases of interest.

EXAMPLE 3.4 A generalized mixed Erlang convolution

Suppose that $X_1, X_2, \ldots, X_k$ are independent random variables such that X_i has Laplace transform

$$\tilde{f}_i(s) = \left(\frac{\lambda_i}{\lambda_i + s}\right)^{\alpha_i} Q_i\left(\frac{\lambda_i}{\lambda_i + s}\right),$$

where $\alpha_i \geq 0$, $Q_i(z)$ is a pgf, and $\lambda_i > 0$. Clearly, if $\alpha_i > 0$ and $Q_i(z) = 1$ for all i, then $\tilde{f}_i(s)$ is the Laplace transform of a gamma distribution. Conversely, if $\alpha_i = 0$ and $Q_i(z) \neq 1$ for all i, then $\tilde{f}_i(s)$ is a mixed Erlang Laplace transform. Then let

$$X = a_1 X_1 + a_2 X_2 + \cdots + a_k X_k,$$

where $a_i > 0$ for all $i = 1, 2, \ldots, k$. The Laplace transform of X is

$$\begin{aligned} \mathrm{E}\left(e^{-sX}\right) &= \mathrm{E}\left[\exp\left(-s\sum_{i=1}^{k} a_i X_i\right)\right] = \prod_{i=1}^{k} E\left(e^{-a_i s X_i}\right) = \prod_{i=1}^{k} \tilde{f}_i(a_i s) \\ &= \prod_{i=1}^{k}\left[\left(\frac{\frac{\lambda_i}{a_i}}{\frac{\lambda_i}{a_i} + s}\right)^{\alpha_i} Q_i\left(\frac{\frac{\lambda_i}{a_i}}{\frac{\lambda_i}{a_i} + s}\right)\right]. \end{aligned} \tag{3.10}$$

The Laplace transform in (3.10) is very general. If $\alpha_i = Q_i(z) = 1$ for all i, the generalized Erlang distribution of Example 2.1 results. If $Q_i(z) = 1$ for all i, then (3.10) is the Laplace transform of the convolution of gamma distributions. If $\alpha_i = 0$ for all i, then (3.10) is the convolution of mixed Erlang distributions. Of course, if $a_i = 1$ for all i, then $X = X_1 + X_2 + \cdots + X_k$, and if $a_1 + a_2 + \cdots + a_k = 1$ then X is a weighted average of the X_i.

If $m = \alpha_1 + \alpha_2 + \cdots + \alpha_k$ is a nonnegative integer, then (3.10) is a mixed Erlang Laplace transform, as we now demonstrate. Let $\lambda \geq \sup_i \lambda_i / a_i$, and using (3.4), we may write

$$\frac{\frac{\lambda_i}{a_i}}{\frac{\lambda_i}{a_i} + s} = \left(\frac{\lambda}{\lambda + s}\right)\left[\frac{\frac{\lambda_i}{a_i \lambda}}{1 - \left(1 - \frac{\lambda_i}{a_i \lambda}\right)\left(\frac{\lambda}{\lambda + s}\right)}\right] = \frac{z}{1 + \beta_i - \beta_i z},$$

where $z = \lambda/(\lambda + s)$ and $\beta_i = (a_i \lambda - \lambda_i)/\lambda_i$. Hence, (3.10) may be expressed as

$$\mathrm{E}\left(e^{-sX}\right) = Q\left(\frac{\lambda}{\lambda + s}\right)$$

where

$$Q(z) = \prod_{i=1}^{k}\left[\left(\frac{z}{1 + \beta_i - \beta_i z}\right)^{\alpha_i} Q_i\left(\frac{z}{1 + \beta_i - \beta_i z}\right)\right],$$

or equivalently

$$Q(z) = z^m \prod_{i=1}^{k}\left[(1 + \beta_i - \beta_i z)^{-\alpha_i} Q_i\left(\frac{z}{1 + \beta_i - \beta_i z}\right)\right]. \tag{3.11}$$