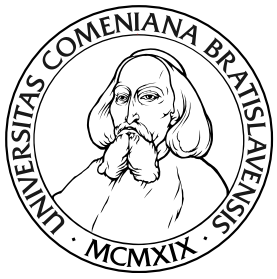
COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS, AND INFORMATICS
DEPARTMENT OF APPLIED MATHEMATICS AND STATISTICS



RECURRENT FAMILIES OF DISCRETE PROBABILITY DISTRIBUTIONS IN COLLECTIVE RISK MODELS

Ph.D Dissertation Thesis

Study program:	Ph.D Applied Mathematics
Field of study:	Statistics
Department:	Department of Applied Mathematics and Statistics
Supervisor:	doc. Mgr. Ján Mačutek, Ph.D.

BRATISLAVA 2025 **FRIDAY IKECHUKWU AGU, Mgr.**

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: M.Sc. Friday Ikechukwu Agu
Študijný program: aplikovaná matematika (Jednoodborové štúdium, doktorandské III. st., denná forma)
Študijný odbor: matematika
Typ záverečnej práce: dizertačná
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Recurrent families of discrete probability distributions in collective risk models
Rekurentné triedy diskrétnych rozdelení v kolektívnom modeli rizika

Anotácia: Práca je venovaná modelovaniu kolektívneho rizika v poisťovníctve pomocou rekurentných tried diskrétnych rozdelení pravdepodobnosti. Ťažiskom je Schröterova trieda rozdelení, kde odvádzame výpočtovo jednoduché odhady parametrov. Zaoberáme sa aj useknutými rozdeleniami zo Schröterovej triedy a vyjadrením poistných nárokov pomocou nekonečných súm.

Školiteľ: doc. Mgr. Ján Mačutek, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: doc. Mgr. Igor Melicherčík, PhD.

Dátum zadania: 03.08.2021

Dátum schválenia: 03.08.2021 prof. RNDr. Daniel Ševčovič, DrSc.
garant študijného programu

.....
študent

.....
školiteľ

Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname:	M.Sc. Friday Ikechukwu Agu
Study programme:	Applied Mathematics (Single degree study, Ph.D. III. deg., full time form)
Field of Study:	Mathematics
Type of Thesis:	Dissertation thesis
Language of Thesis:	English
Secondary language:	Slovak

Title: Recurrent families of discrete probability distributions in collective risk models

Annotation: The thesis deals with modeling collective risk in insurance using recurrent families of discrete probability distributions. The focus is on the Schröter family of distributions, where we derive computationally simple parameter estimates. We also deal with truncated distributions from the Schröter class and the expression of insurance claims using infinite sums.

Tutor:	doc. Mgr. Ján Mačutek, PhD.
Department:	FMFI.KAMŠ - Department of Applied Mathematics and Statistics
Head of department:	doc. Mgr. Igor Melicherčík, PhD.
Assigned:	03.08.2021
Approved:	03.08.2021 prof. RNDr. Daniel Ševčovič, DrSc. Guarantor of Study Programme

.....
Student

.....
Tutor

Abstract

One of the central challenges in actuarial mathematics is the accurate modeling of the number of claims and the severity of claim amounts. This thesis addresses this challenge by advancing the recursive approach to the computation of aggregate claim amounts within the collective risk model, with a particular focus on the Schröter recursive family of discrete probability distributions.

We investigate the theoretical structure, feasible parameter regions, and estimation techniques of the Schröter recursive family, and propose extensions that enhance its applicability to modern actuarial problems. Specifically, we introduce a simple and computationally efficient parameter estimation method for the Schröter family, suitable for use as initial values in optimization algorithms required for maximum likelihood and other complex estimations. Furthermore, we identified new members of the Schröter family, extended the recursion framework to truncated settings, and proposed a generalized Schröter $R_t(a, b, c)$ family along with its associated class of distributions.

The study also presents the truncated Schröter recursive algorithm, designed to model truncated claim amounts more accurately, and validates its performance using simulated and real-world data. To address the modeling of aggregate approved claim amounts, we derive an infinite-sum representation combining the zero-truncated Poisson and shifted exponential distributions and employ saddlepoint approximation methods following Daniels [27] and Wang and Sobrero [124] to obtain analytical approximations of the aggregate approved claim distribution.

Additionally, we explore the convolution of truncated and non-truncated Poisson distributions from the Schröter class and derive the corresponding probability mass function for $Z = K + K_{\text{tr}}$, along with its statistical properties and a moment-based estimation procedure. Simulation results confirm the accuracy and efficiency of the proposed methods.

In conclusion, this thesis extends the theoretical foundations of the Schröter recursive framework, develops novel computational algorithms for aggregate claim modeling, and provides practical estimation and approximation tools applicable to actuarial risk and insurance data analyses.

Abstrakt

Jednou z hlavných výziev v aktuárskej matematike je presné modelovanie počtu poistných udalostí a závažnosti poistných plnení. Táto dizertačná práca sa touto výzvou zaoberá prostredníctvom rozšírenia rekurzívneho prístupu k výpočtu súhrnných poistných plnení v rámci kolektívneho modelu rizika, so zvláštnym zameraním na Schröterovu rekurzívnu triedu diskrétnych rozdelení pravdepodobnosti.

Skúmame teoretickú štruktúru, možné hodnoty parametrov a metódy odhadu parametrov v Schröterovej rekurzívnej triede rozdelení a navrhujeme rozšírenia, ktoré zvyšujú jej uplatniteľnosť pri riešení moderných aktuárskych problémov. Konkrétne predstavujeme jednoduchú a výpočtovo efektívnu metódu odhadu parametrov v Schröterovej triede, vhodnú na použitie ako počiatočné hodnoty v optimalizačných algoritmoch potrebných pre metódu maximálnej vierohodnosti a iné zložité odhady. Ďalej identifikujeme nových členov Schröterovej triedy, rozširujeme rámec rekurzie na useknuté rozdelenia a navrhujeme zovšeobecnenú Schröterovu triedu $R_t(a, b, c)$ rozdelení.

Štúdia taktiež predstavuje Schröterov rekurzívny algoritmus pre useknuté rozdelenia, navrhnutý na presnejšie modelovanie useknutých poistných plnení, a overuje možnosť jeho aplikovania pomocou simulovaných aj reálnych dát. Na riešenie modelovania agregovaných schválených poistných plnení odvádzame nekonečnú sumárnu reprezentáciu kombinujúcu Poissonovo rozdelenie useknuté v nule a posunuté exponenciálne rozdelenie, a používame metódy aproximácie pomocou sedlového bodu podľa Daniels [27] a Wang and Sobrero [124] na získanie analytických aproximácií rozdelenia agregovaných schválených poistných plnení.

Okrem toho skúmame konvolúciu useknutých a neuseknutých Poissonových rozdelení zo Schröterovej triedy a odvodzujeme príslušnú pravdepodobnostnú funkciu pre $Z = K + K_{\text{tr}}$, spolu s jej štatistickými vlastnosťami a odhadmi založenými na momentoch. Výsledky simulácií potvrdzujú presnosť a efektívnosť navrhovaných metód.

Na záver táto dizertačná práca rozširuje teoretické základy Schröterovho rekurzívneho rámca, vyvíja nové výpočtové algoritmy pre modelovanie agregovaných poistných plnení a poskytuje praktické nástroje na odhad a aproximáciu, použiteľné pri analýze aktuárskeho rizika a poisťovacích dát.

Acknowledgements

I am deeply grateful to all the individuals and institutions that have provided guidance, support, and encouragement, making this work possible. This thesis is as much a product of their generosity and commitment as it is of my own effort.

First and foremost, I owe my deepest gratitude to my esteemed supervisor, doc. Mgr. Ján Mačutek, Ph.D., for his unwavering guidance, mentorship, and encouragement throughout my doctoral journey. His remarkable patience, invaluable insights, and dedicated attention to detail have profoundly shaped the direction of this research. The countless hours he devoted to our discussions, reviewing this thesis, and refining our manuscripts reflect his extraordinary commitment not only to my academic growth but also to my personal development as a researcher. His mentorship has been both a source of strength and an inspiration, and I will remain forever indebted to him.

I am equally grateful to the Institute for Applied Statistics at Johannes Kepler University in Linz, Austria, particularly Professor Helmut Waldl, for his collaboration and insightful guidance during my six-month research visit. Our joint work, which forms the basis of Chapter 4 of this thesis, substantially enriched my research and broadened my academic perspective. I am truly grateful for the warm hospitality and intellectual environment I experienced during my stay.

My sincere appreciation also extends to my colleagues and the staff of the Mathematical Institute, Slovak Academy of Sciences, especially Professor RNDr. Gejza Wimmer, DrSc, Professor RNDr. Anatoly Dvurečenskij, DrSc, and doc. RNDr. Karol Nemoga, CSc, whose advice and support have been invaluable. I also warmly acknowledge the Faculty of Mathematics, Physics, and Informatics at Comenius University, in particular the Dean, Professor RNDr. Daniel Ševčovič, Professor Robert Jajcay, Professor Radoslav Harman, and doc. Gábor Szűcs, among many others, for their excellent leadership and for creating an enabling environment in which my studies could flourish.

I am especially grateful to the donors of my Ph.D. grant, whose financial support enabled me to fully dedicate myself to my research and academic pursuits. Their generosity and belief in the value of scientific advancement have been instrumental in the realization of this work.

On a more personal note, I would like to express my heartfelt gratitude to my beloved wife, Mrs. Paulina, and our son, Dexter, for their unwavering love, patience, and encouragement. Their constant support and sacrifices sustained me through the

most challenging moments of this journey. To my dear mother, my siblings, and the cherished memory of my late father, I acknowledge with deep appreciation your sacrifices, unending prayers, and steadfast support. Your faith in me has been the foundation upon which this achievement rests.

Finally, I would like to thank my friend, Petrik Mačutek, for his prayers, kindness, and constant goodwill. Meeting such a genuine and supportive guy in Slovakia has been a blessing, and I am truly grateful for his presence in my life.

To all who contributed to this journey in both great and small ways, please accept my sincere appreciation. This accomplishment belongs not only to me but to all of you who stood with me along the way.

Dedication

I dedicate this work to God Almighty, the source of all wisdom, knowledge, guidance, and inspiration.

Contents

CHAPTER 1

INTRODUCTION AND THEORETICAL FOUNDATIONS OF THE COLLECTIVE RISK MODEL

In this chapter, we introduce the concept of the collective risk model and its theoretical foundation. We begin with an overview of the study and the definition of the collective risk model. Furthermore, we examined the convolution approach and discussed its significance and limitations in computing aggregate claim amounts.

1.1 Background of the Study

The computation of aggregate claim amounts of a collective risk model plays significant roles in insurance companies. Recently, the introduction of a new capital requirement under the Solvency II framework for European Union member states (Baranowska-Zając [12] and Gisler [47]) has substantially impacted computational methods for aggregate claim amounts in insurance companies. Determining premiums and profits is a crucial and interesting aspect of the insurance sector. Accurate premium setting and profit estimation depend on the reliable computation of aggregate claim amounts and the implementation of effective risk classification procedures.
The profit made by an insurance company depends on the balance between the aggregate claim amounts paid and the premium payments received. The distribution of aggregate claim amounts plays a crucial role in pricing insurance portfolios, as it enables insurers to assess the likelihood that the total claim amount will exceed or remain below a certain threshold. This information is essential for determining appropriate premium levels and estimating potential claims or risk uncertainty.
Consequently, insurers seek accurate estimates of aggregate claims of a collective risk model for informed decision-making about competitive pricing, setting risk margins, profit maximization, and ensuring financial reserves for future uncertainties and investments.

In the collective risk model, the distribution of aggregate claim amounts is derived by convolving the distributions of the number of claims and the claim severities. However, one of the key challenges is accurately modeling the aggregate claim amounts distribution when claim severities are concentrated on nonnegative integers and the

claim frequency and severities are discrete random variables.
A common method for modeling aggregate claim distributions involves fitting appropriate distributions separately to the number of claims and the claim severities. For instance, Hogg and Klugman [57] discussed the preliminary statistical problem of fitting probability distributions to insurance data on individual losses and examining the tail behavior, particularly for small and large numbers of losses. Similarly, Packová and Brebera [84] investigated approaches for selecting probability models that accurately capture insurance losses, as well as their appropriate use in risk management, and demonstrated the application of these methods to claim amount data in motor third-party liability insurance. Furthermore, Several studies, such as Gray and Pitts [49]; Pacáková and Gogola [83]; Jindrová and Pacáková [60]; Dzidzornu and Minkah [35]; Yousof et al. [132]; Ieosanurak and Moumeesri [59], and Warren [127] have explored various approaches for fitting distributions to claim amounts data.

Nevertheless, this approach may be limited in certain cases, as it does not always fully account for the convolution structure that characterizes the distribution of aggregate claim amounts. An alternative is the collective risk model, which models the aggregate claims as a compound distribution by jointly considering both the number of claims and the severity of each claim. The compound Poisson and compound negative binomial distributions are widely used within this framework to model aggregate claim amounts via the convolution method. Theoretical and practical aspects of these models have been discussed and explored in several studies, including Johnson et al. [62]; Bening and Korolev [17]; Goffard et al. [48], and Meraou et al. [77]. Moreover, Wimmer and Altmann [130] provides numerous examples of these distributions in insurance applications.

Although the convolution approach is unambiguous from a theoretical perspective, its practical implementation for computing aggregate claim amounts for a large number of claims is complex, as it requires the computation of several convolutions of the conditional distribution of the number of claims and claim severities. This problem has led to several alternative methods for computing aggregate claim amounts of a collective risk model. One of these approaches is the recursive method, which assumes that the distribution of claim amounts is discrete and can compute aggregate claims recursively as the number of claims increases. It does not involve computing several convolutions of the conditional distribution of the number of claim events and requires far less computer time (Sundt and Vernic [115]).

The study of the recursive approach for the computation of aggregate claim amounts of a collective risk model has a rich foundational history that dates back to the early work of Stroh [112] and Tilley [121], where the Laplace transform of the aggregate loss distribution was inverted in a similar approach to the characteristic function inversion method studied in Mong [79] to derive the recursive model.

Since the 1980s, actuarial research has increasingly focused on recursive methods for evaluating aggregate claim amount distributions. Building on the probability generating function (PGF) and recurrence relation originally introduced by Adelson [1], Panjer [87] advanced the field by employing these tools, together with the inverted Laplace transform, to develop a recursive model for aggregate claim amounts and to provide a comprehensive discussion of its applications. Despite its contributions, the 1980 formulation was limited in scope: it applied only to specific count distributions, most notably the Poisson, and lacked a unified recursive framework across related distributional families. Moreover, its reliance on Laplace transform inversion introduced

analytical and numerical challenges, thereby reducing computational efficiency and hindering practical implementation for more complex severity models. These limitations highlighted the need for a more general and tractable recursive approach. In response, Panjer [88] addressed this gap by introducing a general recursion for the $(a, b, 0)$ family of claim number distributions, which encompasses the Poisson, binomial, and negative binomial cases, while accommodating arbitrary claim amount distributions within a single, elegant formulation. This breakthrough, now widely known as the Panjer recursion formula or exact formula, represents a cornerstone result in collective risk theory.

Although Panjer [88] provided a major breakthrough by deriving a unified recursion for the $(a, b, 0)$ family, its scope was limited to three count distributions, which include the Poisson, binomial, and negative binomial, whose probabilities satisfy a simple linear recursion. While this result offered substantial computational efficiency and elegance, it excluded a wide range of claim frequency models that do not fall within the $(a, b, 0)$ structure, such as mixed, overdispersed, or other non-standard distributions that often arise in practice. Recognizing this limitation, Panjer and Willmot [86] expanded the recursive approach to a broader class of frequency distributions in which successive probabilities can be expressed as ratios of polynomials in descending factorial form, which finds useful applications to the problem of computing the probability of ruin in a particular mixed Poisson process in Hesselager [54]. This generalization remarkably expanded the applicability of recursive methods for compound distributions, rendering them suitable for a more realistic and flexible set of actuarial models.

Furthermore, while Panjer and Willmot [86] work was a big step forward, the resulting formulas were without limitation, as it is often cumbersome and lacked the unifying simplicity of the $(a, b, 0)$ class. Moreover, this framework still excluded many distributions of practical and theoretical interest and could suffer from numerical inefficiencies in computation. To address these limitations, Schröter [103] introduced a systematic family of recursive distributions that generalized the earlier approaches, providing a more flexible and unified foundation for the application of recursive methods in aggregate claims modeling. Despite its usefulness, it is not free from several challenges. This study makes a substantial contribution to the literature on the Schröter recursive method and addresses associated challenges.

In this thesis, we investigate the Schröter family of discrete distributions introduced by Schröter [103], together with their truncated variants. The study begins with an examination of the parametric space of this family of distributions. Specifically, the parametric space problem concerns the identification of restricted regions for the parameters in the recursive formula proposed by Schröter [103] for finite and infinite cases of n, the asymptotic properties of the distributions, and the estimation of their parameters. Within this context, we analyze the Schröter recursive algorithm and propose an alternative version based on truncation, designed to facilitate the computation of aggregate claim amounts when individual claims are subject to truncation.

Beyond this, the thesis develops and analyzes a novel compound distribution based on the Schröter family, along with its saddlepoint-based formulation, for modeling aggregate claim approvals in the insurance industry. This involves deriving the distribution of approved claims modeled as a zero-truncated Poisson-Binomial distribution with shifted exponential severity, establishing closed-form expressions for the probability mass and density functions, and addressing the computational challenges of accurate and efficient estimation. To this end, saddlepoint approximation techniques

are employed, supplemented by simulation-based validation. Additionally, the work explores the convolution of truncated distributions from the Schröter family and their potential applications. The thesis is organized into five chapters, each addressing a distinct component of the study.

Chapter **1** introduces the fundamental concepts underpinning the thesis. It provides an overview of the collective risk model, the convolution approach, and truncated distributions, while situating the study within the broader state of the art and outlining key challenges.

Chapter **2** focuses on recursive approaches, particularly the Panjer and Schröter recursions, and the associated families of distributions. It also examines recent developments within this framework, as well as the challenges, with particular emphasis on the parametric space problem for finite and infinite cases of n and parameter estimation problem for the Schröter family. The contribution of this chapter is based on our joint work in Agu et al. [4].

Chapter 3 is devoted to the truncated Schröter family and the corresponding recursive algorithm. It develops both theoretical approaches for deriving the truncated Schröter recursion and numerical methods for evaluating the efficiency of the proposed algorithm. This chapter is based on the work of [2].

Chapter **4** investigates the compound distribution infinite-sum representation defined by the zero-truncated Poisson-Binomial with shifted exponential severity for modeling aggregate claim approvals. The analysis incorporates the truncated Poisson distribution from the Schröter family, offering new insights into the mathematical modeling of aggregate approved claims. This chapter presents results based on the joint work of Agu and Wald [3].

Chapter **5** examines the convolution of truncated distributions from the Schröter family, with particular attention to parameter estimation and numerical application issues, thereby opening new directions for future research. The results of this chapter is based on the joint work of Agu et al. [5].

1.2 The Concept of the Collective Risk Model and Progress

In non-life insurance theory, the analysis of claims for damages is conducted using mathematical models designed to determine the optimal premium that an insurance company should charge to avoid bankruptcy.

The fundamental theory of the collective risk model originates with Filip Lundberg (1909–1934), whose work framed the insurer's surplus as a stochastic process and set the agenda for ruin-related analysis. Building on this foundation, subsequent syntheses by Cramér [26] and the later survey of Cramér [25] consolidated the theory from the viewpoint of random processes, while the broader review of Arfwedson [10] mapped early strands of development. Within this formative phase, complementary directions emerged: the supplementary investigations of Lundberg [71], the early numerical implementations of Lundberg [72], and the linkage to homogeneous random processes in Segerdahl [105]. Concerns with insolvency risk were already visible in Saxon [102], foreshadowing the more systematic ruin-theoretic treatments to follow.

The 1950s marked a period of structured consolidation and applied extensions. Programmatic studies, such as Arfwedson [9] and Cramér [26], expanded the stochastic-

process framing of collective risk. A comprehensive survey in the Skandia Jubilee volume Cramér [26] synthesized these advances from the perspective of stochastic processes. At the same time, practical relevance widened as the collective risk methods were tentatively applied to sectoral problems, most notably crop insurance by Philipson [92]. The decade also laid the groundwork for later technical advances. For example, Prabhu [96] addressed the ruin problem when the time to ruin lacks a closed form, deriving explicit results for negative and positive drift and extending the analysis to general additive claim processes.

Progress in the 1960s reflected a parallel deepening of both theory and practice in collective risk modeling. On the applied side, Kahn [63] offered an accessible treatment that explicitly linked theory to the practice of stop-loss reinsurance, thereby clarifying the operational implications of collective risk models. On the theoretical side, Beekman [14] advanced the understanding of the collective-risk stochastic process itself, refining aggregation dynamics and deriving auxiliary analytical identities that enriched the technical literature. These developments built upon the earlier surplus process framework of Lundberg and the mid-century consolidation achieved through the surveys and contributions of Cramér [26] and Arfwedson [10] and later extended in Cramér [25]. The combined efforts of Prabhu [96], Kahn [63], and Beekman [14] thus mark a transition toward more targeted results on ruin probabilities, reinsurance structures, and computational techniques, establishing the modern foundations of collective risk theory. A comprehensive synthesis of these advancements and the broader literature is provided in Philipson [94].

Building on these historical developments, contemporary collective risk models formalize claims arriving at an insurance company during a given period as random variables (Ramasubramanian [98]). These random variables take values on the real line, representing the monetary amounts of claims, and their distribution captures both the occurrence and severity of claims. Each random variable is associated with a probability distribution function that characterizes the aggregate claim amounts (Djurić [33]). In actuarial science, this modeling framework is fundamental for analyzing the total claim liability of an insurer, since it encapsulates the aggregate exposure arising from all insured policies over a specified horizon.

Within this probabilistic framework, two classical approaches to insurance risk are distinguished: the individual risk model and the collective risk model (Cramér [23]). In the individual risk model, the gain or loss on each policy is treated as a random variable, and the insurer's total gain or loss is obtained by summing across policies. According to the central limit theorem, if the number of policies is sufficiently large, this sum is approximately normally distributed, permitting approximate inference about the company's total financial outcome under specified conditions.

In contrast, the collective risk model, pioneered by Lundberg, directs attention away from individual contracts to the portfolio as a whole. In this setting, both the total number of claims and their severities are modeled stochastically, with the aggregate claims over a given period constituting the central object of analysis. This perspective enables a rigorous treatment of the probability distribution of the total claim amount, the timing of claims, and the insurer's probability of insolvency. Mathematically, the collective risk framework is firmly situated within the broader theory of stochastic processes, which provides powerful tools for characterizing insurance risk dynamics. The objective is to achieve a unified framework that not only describes the distribution of aggregate claims but also delivers practically useful results for premium calculation,

solvency assessment, and reinsurance design (Cramér [25]).

Definition 1. *Let $\{X_i\}_{i=1}^{\infty}$ denote the claim severities, where the random variables X_i are independent and identically distributed (i.i.d.) over the non-negative reals with probability density function (pdf) $f(x)$, $x \geq 0$. Let f^{*n} denote the n-fold convolution of f, where $n = 0, 1, 2, \ldots$.*

Similarly, let N be a discrete random variable representing the number of claims, with probability mass function (PMF) $p_n = \Pr(N = n)$, $n = 0, 1, 2, \ldots$. Assume that $\{X_i\}$ and N are stochastically independent. This unique assumption provides significant mathematical convenience and elegance in the use of the collective risk model. In the collective risk model, the total claim amount generated by a portfolio of N claims over a given period is defined as:

$$S = \sum_{i=1}^{N} X_i. \tag{1.1}$$

The collective claim amount, denoted by S, arises as a compound distribution of X and N. In particular, if there are no claims (that is, $N = 0$), then the aggregate claims amount equals zero. Consequently, the event $(S \leq x)$ is only meaningful when there are $N = n$ claims. The moments and analytical properties of S, including its compound distribution and the models for the distribution of N, have been extensively discussed and explored in Dickson [31] and Hofmann [56].

When N is taken as a fixed positive integer rather than a random variable, the collective risk model reduces to the individual risk model, which is particularly relevant in portfolio analysis. Although several studies, such as Liu and Wang [70] and Blier-Wong et al. [20], have considered cases where N and X_i are dependent, modeling such dependence remains challenging because of the high dimensionality of the joint distribution and the frequent scarcity of reliable data. For this reason, throughout the present study, we adopt the simplifying assumption that N and X_i are stochastically independent.

1.2.1 Model Assumptions for the Collective Risk Model

The collective risk model defines the aggregate claim amount as the sum of individual claim amounts over a random number of claims. The formal assumptions are summarized as follows.

Definition 2 (Model Assumptions)**.** *The aggregate claim amount*

$$S = X_1 + X_2 + ... + X_n,$$

is subject to the following assumptions:

1. *N is a discrete random variable taking values in $\mathbb{N}_0$. In particular, $N = 0$ implies $S = 0$.*

2. *$X_1, X_2, \ldots$ are i.i.d. non-negative random variables with common distribution F_X.*

3. N and $(X_i)_{i\geq 1}$ are independent.

Remark 1.
- *Assumption (2) implies that the claim sizes X_i do not influence one another. For example, observing a large X_1 reveals no information about X_j for $j \geq 2$.*
- *$X_i > 0$ holds almost surely for all $i \in \mathbb{N}$, ensuring strictly positive claim amounts.*

The mathematical foundation of the collective risk model was laid by the pioneering work of Filip Lundberg in the early 20th century. Lundberg's framework, later extended by Cramér [24], introduced the classical ruin theory and provided probabilistic tools to assess an insurer's solvency under the independence assumption between N and $\{X_i\}$. Since then, the collective risk model has received extensive attention in actuarial research, stimulating a wide range of methodological extensions and practical applications. For instance, Déniz and Ojeda [29] proposed a model in which claim severities follow an Erlang distribution and claim frequencies are governed by a discrete generalized Lindley distribution. They derived explicit expressions for the aggregate claim size distribution and validated their approach using two automobile insurance datasets, demonstrating its competitiveness relative to the compound Poisson and negative binomial models.

Furthermore, Dickson et al. [32] employed the Bayesian approach to derive predictive aggregate claims distributions within the framework of the collective risk model. They compared predictive distributions with fitted distributions that ignore parameter uncertainty and demonstrated that actuarial quantities such as premiums can be substantially understated when this uncertainty is neglected. Their study also highlighted situations where the moments of the predictive individual claim distribution do not exist, and they explored suitable approaches for applying such predictive distributions to address various insurance problems. The main conclusion of their work indicates that parameter uncertainty has a considerable impact on the moments and percentiles of aggregate claims distributions in a collective risk model, with uncertainty in the claim number distribution being more influential than in the claim severity distribution whenever the moments of the latter exist. While the objective Bayesian approach may sometimes yield predictive distributions without finite moments, it has been shown that suitable modifications can render such distributions practically applicable for insurance purposes. Similarly, an extensive review and applications of the collective model have been explored in the work of Kahn [63] and **?**].

1.3 The Concept of the Convolution Approach

The key challenge in the collective risk model lies in deriving the distribution of the aggregate claim amount S. One of the earliest and most natural approaches to this problem is the convolution method, which serves as the mathematical tool for combining individual claim amounts into the aggregate S. When this convolution is further mixed over the random claim count N, it yields the standard formulation of the collective risk model.

Parallel to the development of recursive methods, the convolution approach emerged as a fundamental technique for evaluating aggregate claim distributions. This method builds on the principle that the distribution of aggregate claims can be obtained by successively convolving the claim severity distribution with itself, weighted by the claim

frequency distribution. While conceptually straightforward, direct convolution can become computationally intensive as the number of claims increases; thus, early research concentrated on efficient formulations.

An early milestone was set by Hovinen [58], who introduced a systematic method for computing convolutions within the framework of collective risk theory. The work of Hovinen [58] offered one of the first structured attempts to reduce the computational burden of repeated convolutions, thereby laying the groundwork for later algorithmic refinements. This marked an important step from purely theoretical formulations toward methods with practical computational relevance. Building on this, Philipson [93] established convolution theorems for compound Poisson processes, clarifying the mathematical structure underlying convolution-based aggregate models. By supplying explicit convolution identities, the work extended the range of severity distributions that could be handled analytically, thereby strengthening the applicability of convolution methods in actuarial modeling.

Taken together, these contributions represent the initial phase of progress in the convolution approach: from Hovinen's early algorithmic formulation Hovinen [58] to the establishment of general convolution theorems Philipson [93]. This progression highlights the computational challenges and the theoretical richness of convolution methods, which continue to serve as a baseline against which newer approaches, such as recursive algorithms and transform-based techniques, are evaluated in modern actuarial science.

Definition 3 (The Convolution in the Collective Risk Model)**.** *Let N denote the number of claims and $X_1, X_2, \ldots, X_N$ be i.i.d. claim amounts with the distribution function F_X.*

If $N = n$ is fixed, then the aggregate claim amount is:

$$S = X_1 + X_2 + \cdots + X_n,$$

and its distribution is given by the n-fold convolution of F_X:

$$f_S(s \mid N = n) = f_X^{*n}(s),$$

where convolution is defined as:

$$(f_X * f_Y)(s) = \int_0^s f_X(x) f_Y(s - x)\, dx.$$

If N is random, then the distribution of the aggregate claim S is a mixture over all possible N:

$$h_S(s) = \sum_{n=0}^{\infty} P(N = n)\, f_X^{*n}(s),$$

with $f_X^{*0}(s) = \delta(s)$ (the Dirac mass at zero; no claims).

The general formulation above was further refined by Feller [41], who expressed the aggregate distribution as:

$$h_S(s) = \sum_{n=0}^{\infty} P(S \leq s, N = n),$$

$$h_S(s) = \sum_{n=0}^{\infty} P_n\, f_X^{*n}(s),$$

where $f_X^{*n}(s)$ denotes the n-fold distribution of claim severities, $f_X^{*0}(s)$ is defined as 1 for $s \geq 0$ and 0 otherwise, and P_n is the PMF of the claim frequency N. This theoretical formulation provides a general and rigorous basis for calculating the aggregate claim distribution of the collective risk model. To make this computation feasible, De Pril [28] proposed an explicit method for evaluating $f^{*n}(s)$ for $n = 1, 2, \ldots$, thereby making the convolution approach more accessible for numerical applications. The utilization and conceptual development of convolution methods in the collective risk framework have also been examined in detail by Kahn [63] and Philipson [94], who discussed their practical relevance in insurance settings.

The convolution approach has also been extended beyond classical formulations to more specialized families of claim distributions. For example, Beekman [15] examined the distribution of aggregate claims for cases in which claim values are equi-spaced and equi-probable, and further developed convolution-type series for the infinite-time ruin function, including corresponding approximations and an analysis of their associated errors. This work demonstrated how convolution methods can be adapted to structured severity distributions and provided insights into the accuracy of approximation techniques in ruin-theoretic contexts.

More recently, Pak [85] employed convolution ideas in the context of loss modeling by expressing the total loss distribution as the convolution of expected and unexpected components. Specifically, combining normal and exponential distributions to construct a convolution-based model and evaluating its performance using Danish data on large fire insurance losses. The findings of Pak [85] indicated that the proposed approach yields useful information on the distribution of losses and, in some cases, outperforms existing benchmark models. Together, these contributions highlight the adaptability of the convolution approach, both for extending classical aggregate claim models and for incorporating new perspectives in modern loss distribution analysis.

Despite these advances, limitations remain. As Dickson [31] observed, the explicit form of the n-fold convolution f_X^{*n} is not always obtainable for common claim severity distributions. Even when an explicit form exists, the aggregate distribution requires computation of an infinite sum, making the convolution method computationally demanding as the claim count grows. Thus, although the convolution approach is mathematically transparent and theoretically rigorous, its practical implementation can be computationally complex and time-consuming. This recognition motivated the development of alternative approaches, such as recursive algorithms and transform-based techniques, which are examined in Chapter **2** of this thesis.

1.4 Truncated Distributions

Having established the convolution approach for deriving the distribution of aggregate claims in the collective risk model, we now turn to truncated distributions. Such distributions arise naturally in the sciences and play a central role in various applications, particularly in insurance modeling, where claim amounts and claim counts are often modified by contractual conditions. For instance, deductibles eliminate claims below a minimum threshold, policy limits cap the size of large claims, and reporting requirements may exclude smaller or unreported claims altogether. Similarly, claim frequency distributions may be truncated when only positive numbers of claims are admissible, as in zero-truncated Poisson or related models. These truncated versions of claim severity and frequency distributions capture more realistic insurance settings and are crucial

for accurately modeling aggregate risk. Their properties and applications will therefore be employed extensively in the subsequent Chapters of this thesis.

1.4.1 Review on Truncated Distributions

The concept of truncated distributions has evolved substantially since its early introduction, with milestones marked by contributions in statistics and actuarial science. One of the earliest formal treatments can be traced to Stevens [111], who introduced the truncated normal distribution in a biological context and demonstrated how empirical data often deviate from classical assumptions when observations are restricted to a finite range. Thereafter, Thompson [120] emphasized the applied importance of truncation by modeling insect populations with truncated lognormal distributions, while also referencing the earlier works by other researchers on truncated models for word frequencies, accidents, ecology, and bacteriology. These studies further highlighted the natural occurrence of truncation across different applied domains.

The 1950s brought theoretical refinements. Smith [110] investigated truncated and sufficient statistics, clarifying the implications of truncation for estimation. Shapiro [106] analyzed the behavior of sums of independent truncated random variables, thereby extending the understanding of how truncation shapes aggregation properties. These theoretical insights were complemented by comprehensive expositions, most notably the treatment of truncated distributions in Johnson et al. [61], which formalized the theory, consolidated earlier contributions, and provided extensive applications, establishing a central reference in probability and statistics. An encyclopedia-style synthesis was later provided by Lawless [69], who offered an accessible overview with emphasis on actuarial applications. Additional listings of developments can be found in El-Din et al. [36], which reflects the breadth of research on truncated distributions across disciplines.

Within the actuarial science domain, truncated distributions garnered substantial interest due to their natural connection with deductibles, policy limits, and reporting thresholds. These contractual modifications directly impact claim sizes and, by extension, the aggregate distribution of claims. Early recognition of this link appears in Seal [104], who provided a comprehensive discussion of the stochastic theory of risk and explicitly linked truncated and modified claim size distributions to insurance practice. Building on this, Gerber [45] explored collective risk theory and demonstrated how truncation approximates real-world claim amounts subject to policy restrictions. Later, Klugman et al. [65] incorporated truncated distributions into modern actuarial models of insurance losses, showing how truncation can be seamlessly embedded into credibility theory, aggregate claims analysis, and computational examples. Collectively, these contributions demonstrate that truncation is not merely a mathematical device but an essential feature reflecting the very structure of insurance contracts. The historical trajectory is clear: initial recognition in applied sciences [111; 120] was followed by theoretical advances on sufficiency and aggregation [110; 106]. These developments laid the groundwork for more comprehensive statistical treatments, such as those in Johnson et al. [61] and Lawless [69]. Building on this foundation, actuarial applications incorporated truncation as a natural component of contract design, most notably in the works of Seal [104], Gerber [45], and Klugman et al. [65].

Recent actuarial work has sharpened methodology for modeling loss data affected by left truncation (deductibles) and right censoring (policy limits). Frees [43] consolidated practical treatment of truncation/censoring and coverage modifications, provid-

ing accessible derivations and implementation guidance. Complementing this applied focus, Nadarajah [81] addressed the limitations of long-tailed distributions, which often lack finite moments, by introducing truncated versions of heavy-tailed families such as the Student's t, inverted beta, Fréchet, Lévy, and F distributions. Their truncated forms possess finite moments of all orders, with explicit expressions derived and real-data applications illustrating their tractability. In a similar approach, Ahmed et al. [6] explored predictive actuarial models incorporating the truncated Birnbaum–Saunders distribution, analyzing the effects of inflation, the behavior of risk rates and risk measures, and reciprocal transformation properties. Ahmed et al. [6] also developed parameter estimation procedures and demonstrated the empirical usefulness of these models on real loss data. More recently, Kouadria and Zeghdoudi [66] introduced the truncated new-XLindley distribution, highlighting monotonic behavior, order statistics, and maximum likelihood estimators (MLEs) for the upper, lower, and doubly truncated cases, and demonstrated its flexibility through medical data applications.

Methodological innovation in the 2020s has increasingly emphasized robustness, risk measurement, and model selection under truncation/censoring. Poudyal and Brazauskas [95] developed robust estimators for truncated and censored severities (e.g., Pareto) using trimmed and winsorized moments, supported by asymptotic theory and applications in insurance data. Extending risk measurement under contractual modifications, Biswas and Sen [19] proposed nonparametric estimation of spectral risk measures for left-truncated/right-censored data, deriving asymptotic normality, Edgeworth expansions, and bootstrap accuracy. In addressing model uncertainty, Zhao et al. [135] compared common severity families (Fisk, Lognormal, Lomax, Paralogistic, Weibull) across varying truncation/censoring levels, providing guidance on model selection under deductible/limit structures. From a modeling-innovation perspective, Gatti and Wüthrich [44] extended the MBBEFD family to include lower-truncated and right-censored claims, demonstrating how deductible and limit adjustments affect fitted severities and linking contract structures with flexible severity families. Complementary progress in collective risk modeling has addressed dependence-aware frequency–severity models (e.g., copula-linked compounds), which are directly compatible with truncated/censored severities; see, for instance, Shi and Zhao [107].[1]

Taken together, this body of work demonstrates a clear evolution: from the early recognition of truncation in the applied sciences, through theoretical formalization, to actuarial adoption and modern methodological refinement. Since 2019, the literature has shifted focus from fitting truncated/censored severities toward robust inference, risk measurement, model selection, and flexible severity families under deductible/limit structures. This ongoing development integrates naturally with collective risk modeling via convolution and recursive methods, where the severity distribution is modeled in its truncated or censored form to reflect realistic insurance contracts.

1.4.2 The Zero Truncated and Zero-Modified Distributions

In insurance applications, the company is often interested in the number of events that have resulted in claims being filed. Once a claim has been reported, the minimum observed claim count is 1, since the probability of observing zero claims is excluded Dickson [31]. This motivates the use of zero-truncated distributions, where the prob-

[1] Although these dependence models do not require truncation, but are routinely paired with truncated/censored severities induced by contract terms.

ability of zero occurrences is removed. Formally, the zero-truncated PMF is defined as:

$$Q_n = \frac{P_n}{1 - P_0}, \qquad n = 1, 2, \ldots, \tag{1.2}$$

where $P_n = \Pr(N = n)$ denotes the original PMF of the claim frequency distribution and $P_0 = \Pr(N = 0)$ is the probability of zero claims.

Similarly, in practice, insurance data often exhibit situations where insured individuals fail to report claims. For example, a policyholder may not submit a small claim to avoid higher premiums in the next contract period, or minor damages may simply be ignored. These behaviors inflate the number of observed zero counts beyond what the assumed frequency distribution predicts. To address this, the zero-modified distribution is introduced, in which the probability at zero is adjusted by an additional parameter $\alpha \in (0, 1)$ that controls the proportion of excess (or deficit) zeros. The corresponding PMF is given by:

$$Q_n = \frac{(1 - \alpha)P_n}{1 - P_0}, \qquad n = 1, 2, \ldots, \tag{1.3}$$

where α represents the probability mass shifted to account for the excess zeros in the data.

Equations (1.2) and (1.3) provide fundamental tools for constructing models that simultaneously incorporate zero truncation and zero modification. These formulations are particularly useful in deriving recursive formulas for aggregate claim distributions under realistic insurance settings, where under-reporting and structural truncation may occur.

Several zero-truncated and zero-modified frequency distributions have been proposed and analyzed in the actuarial and statistical literature. Examples include the zero-truncated and zero-modified Poisson, negative binomial, and Weibull distributions, among others. Their mathematical properties and practical applications have been extensively discussed in the work of Zuur et al. [136], Zhang et al. [134], Sitho et al. [109], Raqab et al. [99], and Monisha et al. [80].

1.5 Final Remarks

In conclusion, this chapter has introduced the fundamental aspects of the collective risk model, including its underlying assumptions, the convolution approach, and the role of truncated distributions. We have highlighted the significance of these frameworks within the collective risk setting and their wide-ranging applications in the insurance domain. Furthermore, the discussion has emphasized both the strengths and challenges of the collective risk model and the convolution approach, particularly regarding their computational complexity and time-consuming nature, and underscored the importance of recursive methods as a practical alternative.

With this foundation in place, the stage is now set to examine the recursive approach in greater detail, focusing on the Panjer and Schröter classes of discrete distributions. Particular attention will be devoted to challenges related to the parametric space in the infinite-n case and the associated issues of parameter estimation for the Schröter family of distributions.

CHAPTER 2

THE RECURSIVE APPROACH IN THE COLLECTIVE RISK THEORY: THE PANJER AND SCHRÖTER FAMILIES AND THE PARAMETRIC PROBLEM

In this chapter, we focus on recursive approaches for computing aggregate claim amounts, with particular attention to the Panjer and Schröter recursions and their associated families of distributions. We review recent developments and achievements within this framework, highlighting key challenges, including the parametric space problem for the infinite-n case in the Schröter family. A new approach to parameter estimation for the Schröter family is also introduced, based on a moment-matching method. The chapter concludes with final remarks summarizing the contributions, which draw on results from our joint work in Agu et al. [4].

2.1 The Recursive Approach

In the collective risk model, fitting a distribution to the claim count presents several challenges. Typically, claim counts are modeled using the binomial, Poisson, or negative binomial distributions, with parameters often chosen based on judgment or prior assumptions. In practice, however, the situation is more complex: for a given insured, the exposure changes over time, and observations are dominated by claims that have been incurred but not yet reported. These factors make it challenging to accurately estimate the distributions of claim counts. Additionally, computing the aggregate loss distribution can be demanding and time-consuming.

Several approximation and numerical methods have been proposed to address these difficulties. One well-known technique is the normal power approximation introduced by Beard et al. [13], which computes the moments of the aggregate claim distribution from those of the claim count and severity distributions, and then matches them to a convenient reference distribution. While simple, this method can be inaccurate in practice. Improved results were reported by Venter [122] using transformed beta and gamma distributions, though these approximations do not always capture the true distribution of aggregate claims and can still yield outcomes that may not be reliable. Other studies, such as Cooley and Tukey [22] and Heckman and Meyers [52], applied

fast Fourier transform methods to compute aggregate distributions; however, these approaches remain computationally intensive. Similarly, Meyers [78] discussed Monte Carlo simulation as a flexible approach; however, this method also suffers from high computational costs despite its simplicity.

An alternative approach was proposed by Mong [79], who developed a more computationally efficient method based on inverting the characteristic function of the claim severity distribution. While promising, this approach requires that the characteristic function exist in a tractable form, which is often not the case. Mong [79] further employed a shifted Gamma distribution to approximate claim severities, but the method's success depends heavily on the suitability of this approximation and the explicit availability of the characteristic function. Since the collective risk model is fundamentally a convolution of the discrete distributions of claim frequency and claim severity, such methods are inherently constrained by convolution-related difficulties and the derivation of the characteristic function.

The recursive approach offers a powerful alternative and is widely recognized as an exact method for computing aggregate claims. Assuming the claim severity distribution is discrete, recursive formulas enable the computation of aggregate claims with high precision across a broad range of severities iteratively. Unlike the convolution approach, recursion avoids repeated convolutions of frequency and severity distributions, thereby requiring far less computational time. As a result, recursive methods overcome many of the challenges associated with convolution and other numerical or approximation-based approaches, establishing themselves as a cornerstone in modern collective risk modeling.

2.2 The Panjer Recursive Formula

In this section, we examine the Panjer recursive formula, its subsequent extensions, and the associated families of distributions.

The first step toward recursive computation of aggregate claim distributions was taken by Panjer [87], who built on the PGF and the recurrence relation introduced by Adelson [1]. By combining these tools with the inverted Laplace transform, Panjer derived the recurrence relation

$$P_n = \frac{1}{n} \sum_{i=1}^{\min(n,k)} E_i P_{n-i}, \tag{2.1}$$

where P_n denotes the probability of exactly n aggregate claims, $E_i \neq 0$ for $i = 1, \ldots, n$, $P_0 = \exp\left(-\sum_{i=1}^{k} \theta_i\right)$, and $P_{-n} = 0$ by definition. While this result provided an important advancement, its dependence on Laplace inversion limited its computational practicality. To overcome this difficulty, Panjer [88] introduced the celebrated Panjer recursion[1], which bypasses Laplace inversion and applies directly to discrete severity distributions. This recursion specifies that the counting distribution $\{P_n\}$ satisfies the recursive formula:

$$P_n = \left(a + \frac{b}{n}\right) P_{n-1}, \qquad n = 1, 2, \ldots, \tag{2.2}$$

[1]The main focus of this section

with the parameters a and b chosen such that $P_n \geq 0$ and $\sum_{n=0}^{\infty} P_n = 1$, and $P\_n = 0$ for $n \leq 0$. Under this assumption, the aggregate claim distribution $h(s)$ can be expressed recursively as:

$$h(s) = \frac{1}{1 - af_0} \sum_{i=1}^{s} \left(a + \frac{bi}{s} \right) f_i h(s - i), \qquad s = 1, 2, \ldots, \tag{2.3}$$

where $h(0) = P_0$, $f_i = P(X = i)$, $f_0 = 0$, and $P_0 = P(N = 0)$. The family of counting distributions satisfying this recursion is called the Panjer $R_0(a, b)$ class. Panjer [88] demonstrated that the Poisson, binomial, negative binomial, and geometric distributions all belong to the $R_0(a, b)$ class. A key feature of this class is the freedom to specify P_0, implying that these distributions are not uniquely determined by the parameters (a, b). Subsequent research has extensively examined the structure and parameter space of the $R_0(a, b)$ class, as reflected in the works of Rolski et al. [101], Dickson [31], and Yartey [131].

The performance of the Panjer recursion relative to alternatives such as the fast Fourier transform has been analyzed by Embrechts and Frei [37]. Building on this line, Fackler [38] provided a unified parameterization linking the Poisson, binomial, and negative binomial distributions through the recursive formula:

$$P_n = \left(1 + \frac{\lambda}{\alpha} \right)^{-\alpha} \frac{\lambda^n}{n!} \prod_{i=0}^{n-1} \frac{\alpha + i}{\alpha + \lambda}, \qquad n = 0, 1, 2, \ldots, \tag{2.4}$$

with parameters $\alpha \in \mathbb{R}$ and $\lambda > 0$. The Panjer coefficients are then $a = \left(1 + \frac{\lambda}{\alpha}\right)$ and $b = \frac{(\alpha-1)\lambda}{\alpha+\lambda}$. This characterization shows that the negative binomial arises when $\alpha > 0$, the binomial when $-a \in \mathbb{N}$ with $-a > \lambda$, and the Poisson as $\alpha \to \pm\infty$. Farooq [40] further demonstrated that the $R_0(a, b)$ class can be embedded in the exponential family, while Fackler [39] extended the analysis to the $R_k(a, b)$ class using a binomial series representation.

Although the $R_0(a, b)$ class has proven central, its assumption of a fixed positive mass at zero may not always reflect empirical data. To address this, Sundt and Jewell [114] extended the recursion to distributions satisfying

$$P_n = \left(a + \frac{b}{n} \right) P_{n-1}, \qquad n = 2, 3, \ldots, \tag{2.5}$$

with P_1 as the initial value. This class, called $R_1(a, b)$, includes the Poisson, binomial, negative binomial, and logarithmic distributions, with parameter constraints ensuring feasibility. Sundt and Jewell [114] further generalized equation (2.5) to

$$P_n = \left(a + \frac{b}{n} \right) P_{n-1}, \qquad n = t + 1, t + 2, \ldots, \tag{2.6}$$

for $t \geq 1$, though the recursion requires higher-order convolutions and is computationally demanding.

A broader generalization was introduced by Panjer and Willmot [86], who studied the recursions of the form:

$$P_n = \frac{\sum_{i=0}^{k} a_i n^i}{\sum_{i=0}^{k} b_i n^i} P_{n-1}, \qquad n = 1, 2, \ldots, \tag{2.7}$$

for polynomial order k. For $k = 1$ and $k = 2$, Panjer and Willmot [86] derived algorithms for compound distributions, showing that distributions such as Poisson, binomial, negative binomial, hypergeometric, hyper-Poisson, logarithmic, Waring, and Pólya–Eggenberger satisfy this form. Subsequent extensions by Willmot and Panjer [128], Hesselager [54], and Wang and Sobrero [124] expanded applicability to additional severity distributions and provided more general recursive structures.

Subsequent contributions further enriched the framework. For instance, Hess et al. [53] introduced the Panjer distribution of order k, denoted $R_k(a, b)$, with recursion

$$P_{n+1} = \left(a + \frac{b}{n+1}\right) P_n, \qquad n \geq k, \tag{2.8}$$

and the corresponding aggregate recursion is expressed as:

$$h(s) = \frac{1}{1 - af_0} \left[\sum_{i=1}^{s} \left(a + \frac{bi}{s}\right) f_i h(s-i) + p_k f_s^{k*}\right], \qquad s \geq 1. \tag{2.9}$$

This class has been shown to yield finite moments of all orders, and its application to real-life datasets has been investigated in Yartey [131]. Similarly, Sundt [113] developed the generalized $R_k(a, b)$ class, with nesting properties $R_{k-1}(a, b) \subseteq R_k(a, b)$ and theoretical convolution results, though numerical aspects were limited. Further refinements by Willmot and Sundt [129] and Vernic [123], among others, explored computational strategies for these broader classes.

More recent contributions have proposed additional generalizations. Beknazaryan and Adamic [16] introduced the recursion:

$$P_n = \frac{P_k(n)}{n^m} P_{n-1}, \qquad n \geq 1, \tag{2.10}$$

where $P_k(n) = \sum_{i=0}^{k} a_i n^i$ is a polynomial of degree k, encompassing Panjer's original recursion as the special case $k = m = 1$. The authors showed that the pmf factorizes into products of Poisson, binomial, negative binomial, or geometric distributions. Applications of Panjer recursion have since extended well beyond insurance, including to ruin theory Dickson [30], truncated distributions Szűcs [116], and even linguistics Mačutek and Altmann [75]. In practice, Panjer-type recursion has also been combined with fast Fourier transform methods, as demonstrated by Nugrahainy and Azizah [82], who applied it to zero-truncated negative binomial frequency with Burr severities and used the aggregate distribution to estimate gross premiums. In summary, the Panjer recursion has evolved from its original $R_0(a, b)$ formulation into a broad family of recursive structures encompassing $R_k(a, b)$ and beyond. These extensions demonstrate both its flexibility and its enduring role as a cornerstone of recursive methods in collective risk modeling.

2.2.1 The Panjer Class Characterization

First, we consider the $R_0(a, b)$ Panjer class, the method, and conditions that result in the distributions defined in Panjer [88]. From equation (2.2), if $n = 1$, we have $P_1 = (a + b) P_0$, which implies that $a + b \geq 0$ (positive probability weight), and this

condition can be made to be strict. In the case where $a + b = 0$, we have:

$$\begin{aligned} P_1 &= (a+b)P_0 = 0 \\ P_2 &= (a+b)P_1 = 0 \\ &\vdots \\ P_n &= \left(a + \tfrac{b}{n}\right) P_{n-1} = 0 \end{aligned}$$

This implies that $P_0 = 1$ since $\sum_{i=0}^{N} P_i = 1$; thus, $P(N = 0)$ holds.

Theorem 1 (Panjer Class)**.** *The following statements, as studied and presented in Panjer [88] and further discussed in Hofmann [56], are equivalent for $a, b \in \mathbb{R}$ with $a + b > 0$:*

1. *For all $n \in \mathbb{N}$:*

$$P_n = \left(a + \frac{b}{n}\right) P_{n-1}. \tag{2.11}$$

2. *For all $s \in [0, 1)$:*

$$(1 - as)G'_N(s) = (a + b)G_N(s), \tag{2.12}$$

   where $G_N(s) = \sum_{n=0}^{\infty} P_n s^n$ is the PGF of N.

3. *For all $n \in \mathbb{N}, s \in [0, 1)$:*

$$(1 - as)G_N^{(n)}(s) = (na + b)G_N^{(n-1)}(s). \tag{2.13}$$

Proof. Assume equation (2.11) holds. Then $P_0 > 0$ and $P_1 = (a+b)P_0 > 0$. We derive:

$$\begin{aligned} P_n &= \left(a + \frac{b}{n}\right) P_{n-1} \\ &= \frac{(n-1)a + (a+b)}{n} P_{n-1} \\ &\geq \frac{n-1}{n} a P_{n-1}. \end{aligned}$$

Thus,

$$P_n \geq \frac{1}{n} P_1, \qquad \forall n \in \mathbb{N}.$$

If $a \geq 1$, then

$$P_0 + P_1 \sum_{n=1}^{\infty} \frac{1}{n} \leq \sum_{n=0}^{\infty} P_n = 1,$$

which is a contradiction since the harmonic series diverges. Hence, $a < 1$.

(1) $\implies$ (2): Define

$$G_N(s) = \sum_{n=0}^{\infty} P_n s^n.$$

Differentiating gives

$$\begin{aligned} G'_N(s) &= \sum_{n=1}^{\infty} nP_n s^{n-1} \\ &= \sum_{n=1}^{\infty} \left(a + \frac{b}{n}\right) nP_{n-1} s^{n-1} \\ &= \sum_{n=1}^{\infty} \left((n-1)a + (a+b)\right) P_{n-1} s^{n-1}. \end{aligned}$$

Re-indexing yields

$$G'_N(s) = as \sum_{i=1}^{\infty} iP_i s^{i-1} + (a+b) \sum_{i=0}^{\infty} P_i s^i.$$

Hence,

$$(1 - as)G'_N(s) = (a+b)G_N(s).$$

(2) $\implies$ (3): Differentiate $(1-as)G'_N(s) = (a+b)G_N(s)$ repeatedly and use induction.

(3) $\implies$ (1): Using the fact that $P_n = \frac{1}{n!}G_N^{(n)}(0)$:

$$\begin{aligned} P_n &= \frac{1}{n!}(na+b)G_N^{(n-1)}(0) \\ &= \left(a + \frac{b}{n}\right) P_{n-1}. \end{aligned}$$

□

Theorem 1 provides the Panjer characterization of all distributions satisfying the Panjer first-order probability recursion and differential equation for the PGF. Equivalently, the PGF $G_N(s)$ satisfies:

$$(1 - as)G'_N(s) = (a+b)G_N(s), \quad G_N(1) = 1.$$

Thus:

1. There exist $a, b \in \mathbb{R}$ with $a + b > 0$ such that $P_n = \left(a + \frac{b}{n}\right) P_{n-1}$ for all $n \in \mathbb{N}$.
2. The claim count N must follow either a Poisson, binomial, or negative binomial distribution.

In simple terms, within this class (with $a+b > 0$), the admissible counting distributions reduce precisely to the Poisson, binomial, and negative binomial families, as established by Panjer [88].

Theorem 2 (Panjer recursion for the total claim amount)**.** *Let $a, b \in \mathbb{R}$ with $a+b > 0$. Suppose the claim count N belongs to the Panjer $R_0(a, b)$ class and let $X_1, X_2, \ldots$ be i.i.d. severities supported on $\{nh : n \in \mathbb{N}_0\}$ with step size $h > 0$. Write $f_i = P(X_1 = ih)$ (so $f_0 = P(X_1 = 0)$) and let*

$$P(S = nh) = h(n), \qquad n = 0, 1, 2, \ldots.$$

Then,

$$P(S=0) = \begin{cases} P_0, & \text{if } f_0 = 0, \\ G_N(f_0), & \text{if } f_0 > 0, \end{cases} \tag{2.14}$$

and, for $n \geq 1$,

$$P(S=nh) = \frac{1}{1-af_0} \sum_{i=1}^{n} \left(a + \frac{i\,b}{n}\right) f_i\, P\big(S=(n-i)h\big). \tag{2.15}$$

Proof. Let $G_N(s) = \sum_{n\geq 0} P_n s^n$ be the PGF of N and $G_X(s) = \sum_{i\geq 0} f_i s^i$ the PGF of X_1/h as defined in [56]. For the aggregate claims S, we have

$$G_S(s) = G_N\big(G_X(s)\big).$$

At $s = 0$:

$$P(S=0) = G_S(0) = G_N\big(G_X(0)\big) = G_N(f_0),$$

which equals P_0 if $f_0 = 0$.

Now, using

$$(1-as)\, G_N'(s) = (a+b)\, G_N(s),$$

substitute $s = G_X(u)$:

$$\big(1 - aG_X(u)\big)\, G_N'\big(G_X(u)\big) = (a+b)\, G_N\big(G_X(u)\big).$$

Since $G_S(u) = G_N(G_X(u))$, differentiation gives

$$G_S'(u) = G_N'\big(G_X(u)\big)\, G_X'(u).$$

Substituting yields

$$\big(1 - aG_X(u)\big)\, G_S'(u) = (a+b)\, G_S(u).$$

Extracting coefficients at u^n, and writing $G_X(u) = \sum_{i\geq 0} f_i u^i$, $G_S(u) = \sum_{n\geq 0} h(n) u^n$, gives:

$$(1-af_0)\, h(n) = \sum_{i=1}^{n} \left(a + \frac{i\,b}{n}\right) f_i\, h(n-i).$$

[2]. □

2.2.2 Members of the Panjer $R_0(a,b)$ Class

Here, we defined the $R_0(a,b)$ class of discrete distributions along with the recursive relationship that forms the foundation of this class and the conditions under which they satisfy the Panjer class. The $R_0(a,b)$ class has only three members of counting distributions (Panjer [88]).

Also, it is worth noting that the Panjer class is the same as the Katz family introduced in Katz [64], which is a family of discrete distributions that satisfy the first-order recurrence relation defined as:

$$\frac{P_{n+1}}{P_n} = a + \frac{b}{1+n}, \qquad n = k, k+1, .., \tag{2.16}$$

[2]This result has also been explored in Hofmann [56]

and for Panjer class, $k = 0$. This class has been extended and well discussed in Gurland and Tripathi [50] and Pestana and Velosa [91].

Let N be a discrete random variable with the PMF P_n. Then, P_n belongs to the $R_0(a, b)$ class if it can be expressed as:

$$\frac{P_n}{P_{n-1}} = a + \frac{b}{n}, \qquad n = 1, 2, 3, ... \tag{2.17}$$

which generates the probabilities for all integers n starting from 1 with the parameters a and b and the initial probability P_0. Note that each P_n is expressed in terms of P_0. Therefore, when a and b are fixed, P_0 is typically determined to ensure that the probabilities sum up to 1. However, it is essential to acknowledge that not all combinations of the parameters a and b can form a probability distribution under the aforementioned recursive relation, but knowing the nth consecutive probabilities of this class determines the entire distribution. Equation (2.17) can be written in the following pattern:

$$\frac{nP_n}{P_{n-1}} = an + b. \tag{2.18}$$

We show how the parameters a, b, and P_0 of the members of the Panjer class $R_0(a, b)$ are derived in the next section.

2.2.3 Recursive Representation of the Panjer Family

2.2.4 The Poisson Distribution

Let P_n be the PMF of the Poisson random variable N such that:

$$P_n = \frac{\lambda^n e^{-\lambda}}{n!}, \qquad n = 0, 1, 2, ... \tag{2.19}$$

Then, the probability function at $n = 0$ is $P_0 = e^{-\lambda}$ and

$$\frac{P_n}{P_{n-1}} = \frac{\frac{\lambda^n e^{-\lambda}}{n!}}{\frac{\lambda^{n-1} e^{-\lambda}}{(n-1)!}} = \frac{\lambda}{n}.$$

Comparing this with equation (2.2), we have that $a = 0$ and $b = \lambda$.

2.2.5 The Binomial Distribution

If the counting distribution P_n is the binomial distribution, then we have that:

$$P_n = \binom{k}{n} q^n (1 - q)^{k-n}, \qquad k = 0, 1, 2, ...; n = 0, 1, 2, ..., k. \tag{2.20}$$

The probability at $n = 0$ is $P_0 = (1-q)^k$. Similarly,

$$\begin{aligned}
\frac{P_n}{P_{n-1}} &= \frac{\binom{k}{n} q^n (1-q)^{k-n}}{\binom{k}{n-1} q^{n-1}(1-q)^{k-n-1}}, \\
&= \frac{\frac{k! q^n (1-q)^{k-n}}{n!(k-n)!}}{\frac{k! q^{n-1}(1-q)^{(k-(n-1))!}}{(n-1)!(k-(n-1))!}}, \\
&= \frac{k! q^n (1-q)^{k-n}}{n(n-1)!(k-n)!} \times \frac{(n-1)!(k-n+1)(k-n)!}{k! q^{n-1}(1-q)^{k-n+1}}, \\
&= \frac{q}{n(1-q)}[k+1-n] \\
&= \frac{q(k+1)}{n(1-q)} - \frac{qn}{n(1-q)}.
\end{aligned}$$

By comparing this with equation (2.2), we have that $a = -\frac{q}{(1-q)}$ and $b = \frac{q(k+1)}{(1-q)}$ for $q \neq 1$.

2.2.6 The Negative Binomial Distribution

Similarly, if the counting distribution P_n is the negative binomial distribution, then we have that:

$$P_n = \binom{n+k-1}{n} (1-q)^n q^k, \qquad k > 0, n = 0, 1, 2, ... \tag{2.21}$$

The probability at $n = 0$ is $P_0 = q^k$. Similarly,

$$\begin{aligned}
\frac{P_n}{P_{n-1}} &= \frac{\binom{n+k-1}{n}(1-q)^n q^k}{\binom{n+k-2}{n-1}(1-q)^{n-1} q^k}, \\
&= \frac{\binom{n+k-1}{n}(1-q)}{\binom{n+k-2}{n-1}}, \\
&= \frac{(n+k-1)(n+k-2)!(1-q)}{n(n-1)!(k-1)!} \times \frac{(n-1)!(k-1)!}{(n+k-2)!}, \\
&= \frac{(1-q)}{1} + \frac{(1-q)(k-1)}{n}.
\end{aligned}$$

If we compare this with equation (2.2), we have that $a = (1-q)$ and $b = (1-q)(k-1)$ for $q < 1$ and for $k \neq 1$.

2.2.7 The Geometric Distribution: A Special Case

Suppose the counting distribution P_n is the geometric distribution, then we have that:

$$P_n = (1-q)^n q, \qquad , n = 0, 1, 2, ... \tag{2.22}$$

The probability at $n = 0$ is $p_0 = q$. Additionally, we have that

$$\begin{aligned}
\frac{P_n}{P_{n-1}} &= \frac{q(1-q)^n}{q(1-q)^{n-1}} \\
&= (1-q).
\end{aligned}$$

Also, comparing this with equation (2.2), we have that $a = (1 - q)$ and $b = 0$.

The $R_0(a, b)$ class offers the capability to virtualize (using equation (2.17)) and effectively select its members for modeling, especially when a certain number of claims is present. For instance, when dealing with claim data involving multiple cases, denoted as $i = 0, 1, 2, ...$ under various insurance policies, say M_i, we can plot the graph of i against the ratio $\frac{iM_i}{M_i-1}$ and analyze its pattern in alignment with equation (2.17). A negative slope in the graph ($a < 0$) corresponds to the binomial distribution, while a positive slope ($a > 0$) leads to a negative binomial distribution (with geometric distribution as a special case). On the other hand, a level line in the graph ($a = 0$) results in a Poisson distribution. This approach enables us to select a suitable distribution for modeling based on the observed pattern in the graph.

2.2.8 Members of the Panjer $R_1(a, b)$ Class

In the $R_0(a, b)$ class, all members have a fixed positive probability at zero. However, in reality, these members may not appropriately describe the characteristics of some datasets. For example, the sample data might suggest that the probability at zero exceeds what is indicated by the distributions in the $R_0(a, b)$ class. Hence, an alternative approach involves assigning a higher initial probability and subsequently generating the probabilities P_k for $k = 2, 3$, *ldots* recursively. This recursive relation forms the defining characteristic of the $R_1(a, b)$ class (Yartey [131]).

Definition 4 (Panjer $R_1(a, b)$ class)**.** *A counting distribution $\{P_n\}_{n\geq 0}$ belongs to the Panjer $R_1(a, b)$ class if, for some constants $a, b \in \mathbb{R}$,*

$$\frac{P_n}{P_{n-1}} = a + \frac{b}{n}, \qquad n = 2, 3, 4, \ldots, \tag{2.23}$$

with $P_1 > 0$ taken as the initial value, and with parameters and initial probabilities chosen so that $P_n \geq 0$ and $\sum_{n=0}^{\infty} P_n = 1$.

Remarks. 1) Equation (2.23) "degenerates" at $n = 1$, so the recursion starts at P_1 (contrast with the $R_0(a, b)$ class, where the recursion starts at P_0). 2) The value of P_0 is not determined by equation (2.23) and can be set subject to normalization:
If $P_0 = 0$, the resulting model is a zero-truncated member of $R_1(a, b)$.
If $P_0 \in (0, 1)$ is specified (possibly differing from a base model's zero probability), the model is a zero-modified member of $R_1(a, b)$.

This formulation is consistent with the extensions in which $R_1(a, b)$ starts the Panjer-type recursion at $n = 2$ with P_1 as the initial probability, thereby offering flexibility for frequency models that exclude zero counts (zero-truncation) or adjust the mass at zero (zero-modification).

2.3 The Schröter Recursive Formula

This section is devoted to the study of the Schröter recursive formula and its associated family of distributions. We begin by revisiting the parameter characterization originally introduced in Schröter [103], which provides the foundational framework for the Schröter family. Building on this, we extend the analysis to explore the parametric space and feasible region of the family, offering new insights into its structural

properties. Finally, we address the problem of parameter estimation by employing a moment-matching approach, and we present numerical results that illustrate the performance and applicability of the proposed method.

2.3.1 The Schröter Formula

While the Panjer recursive formula effectively solves the computational challenges of the convolution method, it is limited to a small number of distributions. This limitation led Schröter [103] to introduce Schröter's second-order recursive formula, which extends Panjer's formula and provides a more accurate and broader representation of the behavior of claim severity and the number of claim distributions. The Schröter's second-order recursive formula is:

$$P_n = \left(a + \frac{b}{n}\right) P_{n-1} + \frac{c}{n} P_{n-2}, \qquad n = 1, 2, ..., \tag{2.24}$$

where a, b, and c are parameters of the Schröter's recursive formula, and by definition $P_{-1} = 0$. This class degenerates at 0 and can be referred to as the $R_0(a, b, c)$ class.

The corresponding recursion formula for equation (2.24) is expressed as:

$$h(s) = \frac{1}{1 - af_0} \sum_{i=1}^{s} [(a + \frac{b}{s} i) f_i + \frac{c}{2s} i f_i^{2*}] h(s - i), \tag{2.25}$$

such that f^{2*} is obtained by the convolution formula given as

$$f_i^{2*} = \sum_{k=0}^{i} f_k f_{i-k}$$

The Panjer recursion and recursive formulas emerge as special cases of equations (2.24) and (2.25), respectively, when $c = 0$. The class of distributions defined by equation (2.24) encompasses, among others, the convolution of the Poisson distribution with distributions belonging to the Panjer class (see Schröter [103], pp. 172–173).

Equation (2.24) can be expressed as:

$$P_n = \left(a + \frac{b}{n}\right) P_{n-1} + \frac{c}{n} P_{n-2}, \qquad n = 2, 3, \ldots, \tag{2.26}$$

which we denote as the $R_1(a, b, c)$ class of the Schröter family of discrete distributions. Here, the truncation level begins at 1, and since $n > 1$, it is unnecessary to consider the case P_{-1}.

Although Schröter [103] identified a portion of the parametric space for the Schröter recursive formula, the complete characterization of this space remains unresolved.[3] Parameter estimation for this family has also posed challenges.[4] In a conference paper, Szűcs [117] applied numerical estimation by minimizing the Kolmogorov–Smirnov test statistic, while Szűcs [118] discussed applications of the Schröter class in modeling claim counts under the collective risk model using various calibration methodologies.

[3] In this section, we extend this analysis and provide new results on the feasible region for the parameters of the Schröter family.

[4] We propose a new, straightforward estimation approach for the Schröter family, the results are presented in the subsequent section.

More recently, Agu et al. [4] proposed a simple method for estimating the parameters of the Schröter family. Their numerical illustrations showed that the resulting parameter values can serve as effective initial values for iterative computational algorithms. However, the non-convergence of the associated standard deviations raises concerns about the unbiasedness of the estimates, suggesting that the method is best regarded as a source of starting values for optimization procedures.

Despite these developments, both in parameter characterization and estimation, the full implementation of the Schröter recursion formula continues to rely on the convolution approach[5].

2.3.2 The Schröter Characterization

This section presents the characterization of the Schröter recursive formula as studied in Schröter [103].

Theorem 3 (Characterization of the Schröter Recursive Formula)**.** *Let $\{P_n\}_{n\geq 0}$ be a counting pmf satisfying the second–order Schröter recursive formula:*

$$P_n \;=\; \left(a+\frac{b}{n}\right) P_{n-1} \;+\; \frac{c}{n}\,P_{n-2}, \qquad n\geq 1, \quad P_{-1}:=0, \tag{2.27}$$

for constants $a<1$, $b\in\mathbb{R}$, $c\in\mathbb{R}$ with $P_0>0$. Then the PGF $G(s)=\sum_{n\geq 0}P_n s^n$ satisfies

$$(1-as)\,G_N'(s) \;=\; (a+b+c\,s)\,G_N(s), \qquad s\in[0,1]. \tag{2.28}$$

The solution is

$$G_N(s)=\exp\Big\{-\frac{c}{a}(s-1)\;+\;\delta\,\ln\Big(\frac{1-a}{1-as}\Big)\Big\}, \quad a\neq 0, \quad \delta=\frac{a(a+b)+c}{a^2}, \tag{2.29}$$

$$G_N(s)=\exp\Big\{\frac{c}{2}\,(s^2-1)+b\,(s-1)\Big\}, \qquad a=0. \tag{2.30}$$

Moreover,

$$\mathbb{E}[N] \;=\; \frac{a+b+c}{1-a}, \qquad \mathrm{Var}(N) \;=\; \frac{a+b+(2-a)c}{(1-a)^2}. \tag{2.31}$$

Admissible parameters satisfy

$$a+b\;\geq\;0, \qquad a+b+c\;\geq\;0, \qquad b>-1, \qquad \text{and if } c<0 \text{ then } a>0. \tag{2.32}$$

Finally, N is infinitely divisible (hence compound Poisson) if and only if

$$\text{(i) } a<0 \text{ and } \delta=0; \qquad \text{(ii) } a=0 \text{ and } c\geq 0; \qquad \text{(iii) } a>0 \text{ and } \delta\geq 0. \tag{2.33}$$

6

[5] As an alternative, we introduced an efficient Schröter recursive algorithm in Chapter 3 of this thesis.

[6] Here, Schröter [103] only explicitly specified the conditions and partial characterization and stated that this specification does not yield a complete characterization of the parameter space for the Schröter family, and certain regions of this parametric space do not correspond to discernible distributions. This implies that the exact parametric space problem for the Schröter family remains unsolved. Although the complete characterization problem has not been addressed, in Section 2.4.2, we consider the feasible region, which extends the parametric characterization analysis in Schröter [103].

Lemma 1 (Weighted convolution identity)**.** *Let $X_1, \ldots, X_n$ be i.i.d. on $\mathbb{N}_0$ with pmf $f_i = \Pr(X_1 = i)$, and denote by f_i^{*n} the n–fold convolution. Then for each $n \geq 1$ and $i \geq 1$,*

$$f_i^{*n} = \frac{1}{n} \sum_{r=1}^{i} r\, f_r\, f_{i-r}^{*(n-1)}. \tag{2.34}$$

Suppose that the counting distribution P_n of N satisfies equation (2.27), Schröter [103] derived some fundamental properties of such distributions and showed that these distributions have the PGF defined as

$$G_N(s) = e^{\frac{-c}{a}(s-1)} \left(\frac{1-a}{1-as} \right)^{\delta}, \tag{2.35}$$

where $\delta = \left(\frac{a(a+b)+c}{a^2} \right)$, the corresponding PMF as:

$$P_n = e^{\frac{c}{a}} (1-a)^{\delta} \sum_{i=0}^{n} \binom{\delta + i - 1}{i} \frac{\left(-\frac{c}{a}\right)^{n-i} a^i}{(n-i)!}, \tag{2.36}$$

Equations (2.31), (2.35), and (2.36) are viable under the condition $a \neq 0$, and they assume a more intricate form when $a = 0$. Specific parametric domains corresponding to a few distinct distributions within this classification have been specified in (Schröter [103]; page 174).

2.3.3 The Convolution of the Poisson and $R_0(a, b)$ Class

Proposition 1. *The convolution of Poisson and the members of the $R_0(a, b)$ Panjer class generates certain members of the Schröter $R_0(a, b, c)$ class (Schröter [103]; page 172).*

If N_1 has a Poisson distribution and N_2 belongs to the $R_0(a, b)$ Panjer class, then $N_3 = N_1 + N_2$ belongs to the Schröter class. This result is essential and useful for deriving even members of the higher-order truncated Schröter class.

2.3.4 Members of the Schröter $R_0(a, b, c)$ Class

In this section, we show that all members of the Panjer $R_0(a, b)$ class are also included within the Schröter $R_0(a, b, c)$ class.

Let P_n denote the PMF of N. The distribution P_n belongs to the Schröter $R_0(a, b, c)$ class if it can be expressed in the form of equation (2.28).

2.3.5 Examples

Poisson distribution. Let P_n denote the PMF of a Poisson distribution with PGF:

$$G(s) = e^{\lambda(s-1)}. \tag{2.37}$$

Thus,

$$\frac{G'(s)}{G(s)} = \lambda. \tag{2.38}$$

Comparing with equation (2.28), we obtain $a = 0$, $b = \lambda$, and $c = 0$.

Binomial distribution. The PGF of the binomial distribution is

$$G(s) = (q + ps)^n. \tag{2.39}$$

We have:

$$\frac{G'(s)}{G(s)} = \frac{np(q+ps)^{n-1}}{(q+ps)^n} = \frac{np}{q+ps}. \tag{2.40}$$

Thus, $a = -p$, $a + b = np$, $c = 0$. Hence, $b = p(n-1)$.

Negative binomial distribution. The PGF of the negative binomial distribution is

$$G(s) = \frac{p^r}{(1-(1-p)s)^r}, \qquad |s| < \tfrac{1}{p}. \tag{2.41}$$

Then,

$$\frac{G'(s)}{G(s)} = \frac{r(1-p)}{1-(1-p)s} = \frac{r-rp}{1-(1-p)s}. \tag{2.42}$$

We deduce that $a = (1-p)$, $a + b = r(1-p)$, $b = (1-p)(r-1)$, and $c = 0$.

Geometric distribution. The PGF of the geometric distribution is

$$G(s) = \frac{p}{1-(1-p)s}. \tag{2.43}$$

We compute

$$\frac{G'(s)}{G(s)} = \frac{1-p}{1-(1-p)s}. \tag{2.44}$$

Thus, $a = (1-p)$, $a + b = 1-p$, $b = 0$, and $c = 0$.

Other examples of members of the Schröter $R_0(a, b, c)$ class can be found in Schröter [103, page 174].

2.3.6 Other Distributions

Beyond the classical members, several other discrete distributions can also be shown to belong to the Schröter class. Using results from Wimmer and Altmann [130], we present two such examples:

(i) The Bartlett distribution. The PMF is

$$P_n = \sum_{i=0}^{n} \frac{e^{-\alpha}\alpha^{n-i}}{(n-i)!} pq^i, \qquad n = 1, 2, \ldots,\ \alpha \geq 0,\ 0 < p \leq 1,\ q = 1-p. \tag{2.45}$$

The PGF is

$$G(s) = \frac{pe^{\alpha(s-1)}}{1-qs}. \tag{2.46}$$

Then,

$$\frac{G'(s)}{G(s)} = \frac{\alpha + q - \alpha qs}{1-qs}. \tag{2.47}$$

Hence, $a = 1-p$, $a + b = \alpha + 1 - p$, $b = \alpha$, and $c = -\alpha(1-p)$.

(ii) The Binomial-Geometric distribution. The PMF is

$$P_n = (1-\theta)q^n \sum_{i=0}^{x} \binom{n}{x-i} \left(\frac{p}{q}\right)^x \left(\frac{q\theta}{p}\right)^i, \qquad x = 0, 1, \ldots,\ p \neq 0,\ n \in \mathbb{N}_0,\ 0 < p \leq 1,\ 0 \leq \theta \leq 1. \tag{2.48}$$

The PGF is

$$G(s) = \frac{(q+ps)^n(1-\theta)}{1-\theta s}. \tag{2.49}$$

We have

$$\frac{G'(s)}{G(s)} = \frac{(1-\theta s)np + \theta(q+ps)}{(q+ps)(1-\theta s)}. \tag{2.50}$$

If $q + ps = 1$, this simplifies to

$$\frac{G'(s)}{G(s)} = \frac{np - \theta nps + \theta}{1-\theta s}. \tag{2.51}$$

Thus, $a = \theta$, $a + b = np + \theta$, $b = np$, and $c = -\theta np$.

2.3.7 Members of the $R_1(a, b, c)$ Class

If equation (2.27) is truncated at 1, we can rewrite equation (2.28) as

$$\frac{G'(s)}{G(s)} = \frac{a_1 + b_1 + c_1 s}{1 - a_1 s}. \tag{2.52}$$

Also, Panjer [87] showed for equation (2.2) that

$$\frac{G'(s)}{G(s)} = \frac{a+b}{1-as}. \tag{2.53}$$

We derive members of the $R_1(a, b, c)$ class by extending Proposition 1, namely, by convolving the truncated Poisson distribution with members of the $R_1(a, b)$ class. The truncated distributions can be obtained using

$$q_n = \frac{p_n}{1-p_0}, \tag{2.54}$$

where P_n is the pmf of the counting distribution. Using equation (2.54), we can therefore derive the pmfs of the $R_1(a, b)$ members (Yartey [131]).

Let $G_i(s) = E(s^{n_i})$ denote the PGF of n_i, and let $P(s) = E(s^{n_N})$ denote the PGF of n_N for $i = 1, \ldots, N$. Then

$$P(s) = E(S^{n_1} \cdots S^{n_N}) = \prod_{i=1}^{N} G_i(s). \tag{2.55}$$

If the n_i are independent and identically distributed, then

$$P(s) = [G(s)]^N. \tag{2.56}$$

Now let $N_1 \sim \text{Poisson}(\lambda)$ with PGF $G_1(s)$ and N_2 be from the $R_1(a, b)$ class with PGF $G_2(s)$. Then

$$N_3 = N_1 + N_2, \qquad G_{N_3}(s) = G_1(s)G_2(s). \tag{2.57}$$

Taking logarithms and differentiating both sides of equation (2.57), it follows that N_3 has a counting distribution in the $R_1(a, b, c)$ class.

2.4 The Convolution of Truncated Poisson and Other Truncated Distributions

Using equation (2.54), the PMFs and PGFs of truncated Poisson, binomial, negative binomial, and geometric distributions are:

Truncated Poisson distribution.

$$q_n = \frac{\lambda^n e^{-\lambda}}{n!(1-e^{-\lambda})}, \qquad n = 0, 1, 2, \ldots, \tag{2.58}$$

$$G(s) = \frac{e^{\lambda(s-1)}}{1-e^{-\lambda}}. \tag{2.59}$$

Truncated binomial distribution.

$$q_n = \binom{N}{n} p^n q^{N-n}, \qquad q = 1-p, \tag{2.60}$$

$$G(s) = \frac{(p+qs)^N}{1-q^N}. \tag{2.61}$$

Truncated negative binomial distribution.

$$q_n = \frac{\binom{N+r-1}{n} p^r q^n}{1-p^r}, \tag{2.62}$$

$$G(s) = \frac{p^r}{(1-qs)^r(1-p^r)}. \tag{2.63}$$

Truncated geometric distribution.

$$q_n = pq^{n-1}, \qquad n = 1, 2, \ldots, \tag{2.64}$$

$$G(s) = \frac{p}{(1-qs)(1-p^r)}. \tag{2.65}$$

2.4.1 Examples

Convolution of Truncated Poisson and Truncated Binomial. Here N_1 is truncated Poisson and N_2 truncated binomial (illustration of 1). Then,

$$\frac{G'_{N_3}(s)}{G_{N_3}(s)} = \lambda + \frac{Nq}{p+qs} = \frac{Nq + \lambda p + \lambda qs}{p+qs}. \tag{2.66}$$

Thus, $a = -q$, $a+b = Nq+\lambda p$, $b = q(N+1)+\lambda p$, and $c = \lambda q$. Hence, the convolution of N_1 and N_2 is a member of the Schröter $R_1(a, b, c)$ class.

Convolution of truncated Poisson and truncated negative binomial.

$$\frac{G'_{N_3}(s)}{G_{N_3}(s)} = \lambda + \frac{rq}{1-qs} = \frac{rq+\lambda+\lambda qs}{1-qs}. \tag{2.67}$$

Hence, $a = q$, $a + b = rq + \lambda$, $b = q(r-1) + \lambda$, and $c = -\lambda q$. Additionally, the convolution is a member of the Schröter $R_1(a, b, c)$ class.

Convolution of truncated Poisson and truncated geometric.

$$\frac{G'_{N_3}(s)}{G_{N_3}(s)} = \lambda + \frac{q}{1-qs} = \frac{\lambda(1-qs)+q}{1-qs}. \tag{2.68}$$

Thus, $a = q$, $a + b = \lambda + q$, $b = \lambda$, and $c = -\lambda q$.

Hence, the convolution of truncated Poisson with truncated binomial, truncated negative binomial, and truncated geometric distributions all belong to the Schröter $R_1(a, b, c)$ class of counting distributions.

2.4.2 The Feasible Region for the Parameters of the Schröter Recursive Formula for Finite Case

In this section, we extend the parametric characterization in Schröter [103] by investigating feasible regions for the parameters of the Schröter recursion and then passing to the limit as $n \to \infty$. We first derive finite-n feasibility regions by constraining the recursion coefficients, and subsequently establish infinite-n feasibility by employing the generating-function representation. This section brings new results concerning the feasibility of the Schröter recursive formula, asymptotic behavior, parameter estimation, and probabilistic practical guidance for modeling and predicting claim counts.

We start from the Schröter's second–order recursion

$$P_n \;=\; \left(a + \frac{b}{n}\right) P_{n-1} \;+\; \frac{c}{n}\, P_{n-2}, \qquad n \geq 2, \qquad P_{-1} := 0, \quad P_0 > 0. \tag{2.69}$$

A convenient sufficient condition ensuring nonnegativity and bounded growth at a fixed n is to require

$$0 \leq a + \frac{b}{n} \leq 1 \quad \text{and} \quad 0 \leq \frac{c}{n} \leq 1.$$

This yields explicit half-space constraints in (a, b, c) for each fixed n.

Base cases. For $n = 1$, equation (2.26) gives $P_1 = (a+b)P_0$, so $a+b \geq 0$ is necessary for $P_1 \geq 0$. For $n = 2$,

$$P_2 = \left(a + \frac{b}{2}\right) P_1 + \frac{c}{2}\, P_0,$$

and the four bounding planes $a + \frac{b}{2} = 0$, $a + \frac{b}{2} = 1$, $\frac{c}{2} = 0$, $\frac{c}{2} = 1$ delineate a feasible polyhedral slice (see Figure 2.1). Intersecting these half-spaces gives:

$$a + \tfrac{b}{2} \in [0, 1] \quad \text{and} \quad \tfrac{c}{2} \in [0, 1].$$

Similarly, for $n = 3$ (Figure 2.2), we obtain:

$$a + \tfrac{b}{3} \in [0, 1] \quad \text{and} \quad \tfrac{c}{3} \in [0, 1].$$

Finite-n pattern. Continuing in this way, a sufficient pointwise condition for all finite n is

$$a + \tfrac{b}{n} \in [0,1] \quad \text{and} \quad \tfrac{c}{n} \in [0,1] \qquad \text{for each fixed } n \in \mathbb{N}.$$

The feasible region consists of all points (a, b, c) such that $P_n \geq 0$ and $\sum_{n=0}^{\infty} P_n = 1$. Points lying outside this region correspond to parameter values that are not admissible. The lower boundaries of feasibility are determined by the conditions $a + \frac{b}{n} = 0$, $\frac{c}{n} = 0$, $a + \frac{b}{n} \geq 0$, and $\frac{c}{n} \geq 0$, which ensure that the coefficients $a + \frac{b}{n}$ and $\frac{c}{n}$ remain nonnegative and that the resulting probabilities are greater than or equal to zero. Similarly, the upper boundaries are defined by $a + \frac{b}{n} = 1$, $\frac{c}{n} = 1$, $a + \frac{b}{n} \leq 1$, and $\frac{c}{n} \leq 1$, thereby constraining the recursion so that the coefficients do not exceed one. This guarantees that the recursion remains bounded and that the sum of probabilities generated by the formula does not exceed one. Consequently, points (a, b, c) within the intersection of these regions yield nonnegative probabilities at each recursive step, while points outside necessarily violate one of the bounding inequalities. It is important to note, however, that these conditions, while practically useful, are not sufficient on their own to guarantee the feasibility of the entire infinite sequence for the Schröter family of distributions.

2.4.3 Feasibility in the Limit as $n \to \infty$ for the Schröter Recursive Formula

To address the infinite case, we start with the PGF:

$$G_N(s) \;=\; \sum_{n \geq 0} P_n s^n, \qquad s \in [0,1].$$

By multiplying the recursion by ns^{n-1} and summing over $n \geq 1$ yields the first–order ordinary differential equation (ODE)

$$(1 - as)\, G'_N(s) \;=\; (a + b + c\, s)\, G_N(s), \qquad G_N(1) = 1. \tag{2.70}$$

By solving the ODE, i.e., we solve (2.70) by separation of variables:

$$\frac{G'_N(s)}{G_N(s)} \;=\; \frac{a + b + cs}{1 - as}.$$

and integrating both sides gives:

$$\int \frac{G'_N(s)}{G_N(s)}\, ds \;=\; \int \frac{a + b + cs}{1 - as}\, ds.$$

The left-hand side integrates to:

$$\ln G_N(s).$$

For the right-hand side, write

$$\frac{a + b + cs}{1 - as} = \frac{a + b}{1 - as} + \frac{cs}{1 - as}.$$

The first term is solved as:

$$\int \frac{a + b}{1 - as}\, ds = -\frac{a + b}{a} \ln(1 - as).$$

Next, the second term is solved as:

$$\int \frac{cs}{1-as}\,ds.$$

Let us set $u = 1 - as$, so $du = -ads$ and $s = \frac{1-u}{a}$. Then

$$\int \frac{cs}{1-as}\,ds = \int \frac{c\cdot\frac{1-u}{a}}{u}\cdot\Big(-\frac{du}{a}\Big) = -\frac{c}{a^2}\int\Big(\frac{1}{u}-1\Big)du.$$

Thus,

$$\int \frac{cs}{1-as}\,ds = -\frac{c}{a^2}\ln u + \frac{c}{a^2}u = -\frac{c}{a^2}\ln(1-as) + \frac{c}{a^2}(1-as).$$

Therefore,

$$\ln G_N(s) = -\frac{a+b}{a}\ln(1-as) - \frac{c}{a^2}\ln(1-as) + \frac{c}{a^2}(1-as) + C,$$

where C is the constant of integration.

Simplify:

$$\ln G_N(s) = -\frac{a(a+b)+c}{a^2}\ln(1-as) + \frac{c}{a^2}(1-as) + C.$$

Since $G_N(1) = 1$, we must have $\ln G_N(1) = 0$. At $s = 1$ (boundary condition),

$$0 = -\frac{a(a+b)+c}{a^2}\ln(1-a) + \frac{c}{a^2}(1-a) + C.$$

Thus,

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

Hence,

$$\ln G_N(s) = -\frac{a(a+b)+c}{a^2}\ln(1-as) + \frac{c}{a^2}(1-as) + \frac{a(a+b)+c}{a^2}\ln(1-a) - \frac{c}{a^2}(1-a).$$

Exponentiating,

$$G_N(s) = \exp\left\{-\frac{c}{a}(s-1)\right\}\left(\frac{1-a}{1-as}\right)^{\delta}, \qquad \delta = \frac{a(a+b)+c}{a^2}, \quad a \neq 0.$$

Special case $a = 0$. When $a = 0$, the ODE (2.70) reduces to:

$$G_N'(s) = (b + cs)\,G_N(s).$$

Integrating,

$$\ln G_N(s) = b(s-1) + \frac{c}{2}(s^2-1).$$

Thus,

$$G_N(s) = \exp\Big\{b(s-1) + \tfrac{c}{2}(s^2-1)\Big\}.$$

Therefore,

$$G_N(s) = \exp\left\{ -\frac{c}{a}(s-1) \right\} \left(\frac{1-a}{1-as} \right)^{\delta}, \qquad \delta = \frac{a(a+b)+c}{a^2}, \qquad a \neq 0, \tag{2.71}$$

$$G_N(s) = \exp\Big\{ \tfrac{c}{2}(s^2-1) + b(s-1) \Big\}, \qquad a = 0. \tag{2.72}$$

These equations coincide with (2.71)–(2.72) and are the necessary forms arising from the recursion. Next, we identify parameter regions where G_N is (the product of) known PGFs, which yields sufficient conditions for full infinite–sequence feasibility.

Theorem 4 (Limit-Feasibility of the Schröter Recursive Formula)**.** *Let G_N be the PGF given by equations* (2.73) *and* (2.74)*, i.e.*

$$G_N(s) = \exp\Big\{ -\tfrac{c}{a}(s-1) \Big\} \Big(\tfrac{1-a}{1-as} \Big)^{\delta}, \qquad \delta = \tfrac{a(a+b)+c}{a^2}, \quad a \neq 0, \tag{2.73}$$

$$G_N(s) = \exp\Big\{ \tfrac{c}{2}(s^2-1) + b(s-1) \Big\}, \qquad a = 0. \tag{2.74}$$

If one of the following disjoint parameter conditions holds, then G_N is a valid PGF on $[0,1]$; consequently the recursion (equation (2.69)) defines a proper PMF $\{P_n\}_{n\geq 0}$ (that is, $P_n \geq 0$ for all n and $\sum_{n\geq 0} P_n = 1$):

(1) $0 < a < 1$, $\delta \geq 0$, *and* $c \leq 0$.
Then

$$G_N(s) = \underbrace{\exp\{(-c/a)(s-1)\}}_{\text{Poisson with mean } -c/a} \times \underbrace{\Big(\tfrac{1-a}{1-as} \Big)^{\delta}}_{\text{Negative–binomial } (\delta, p=a)},$$

a product of PGFs (interpretable as an independent Poisson sum of a negative–binomial count), and therefore a PGF.

(2) $a = 0$, $b \geq 0$, $c \geq 0$.
Writing $\mu = b + \frac{c}{2} \geq 0$ and

$$\Psi(s) \;=\; \frac{b}{\mu}\, s + \frac{c/2}{\mu}\, s^2 \qquad (\text{when } \mu > 0),$$

we have

$$\log G_N(s) \;=\; b(s-1) + \tfrac{c}{2}(s^2-1) \;=\; \mu\big(\Psi(s) - 1\big),$$

so $G_N(s) = \exp\{\mu(\Psi(s)-1)\}$ is the PGF of a compound Poisson distribution with rate μ and severity PGF Ψ. (If $\mu = 0$ then $G_N \equiv 1$, the degenerate case.)

(3) $a < 0$, $\delta = 0$, *and* $-c/a \geq 0$.
If $\delta = 0$, the negative–binomial factor in (2.73) *is unity and*

$$G_N(s) = \exp\{(-c/a)(s-1)\},$$

which is the Poisson PGF with mean $-c/a \geq 0$, hence a PGF.

Proof. The theorem follows directly from the closed forms (2.73)–(2.74) derived by integrating the ODE $(1-as)G_N'(s) = (a+b+cs)G_N(s)$ and enforcing $G_N(1) = 1$.

(1) For $0 < a < 1$, the factor $\left(\frac{1-a}{1-as}\right)^{\delta}$ is the PGF of a negative–binomial distribution with parameters $(\delta, p = a)$ provided $\delta \geq 0$ and $p = a \in (0,1)$. The exponential factor $\exp\{(-c/a)(s-1)\}$ is the PGF of a Poisson distribution with mean $-c/a$; the condition $c \leq 0$ ensures $-c/a \geq 0$ when $a > 0$. The product of two PGFs is a PGF (corresponding to an independent convolution), hence G_N is a PGF under these parameter constraints.

(2) With $a = 0$ we have $\log G_N(s) = b(s-1)+\frac{c}{2}(s^2-1)$. Writing this as $\mu(\Psi(s)-1)$ with $\mu = b+c/2$ and Ψ as above exhibits G_N as the PGF of a compound Poisson distribution (Poisson rate μ and severity PGF Ψ). The conditions $b \geq 0$, $c \geq 0$ guarantee $\mu \geq 0$ and that Ψ is a convex combination of s and s^2, hence a valid severity PGF. If $\mu = 0$ the PGF is identically one, the trivial (degenerate) distribution.

(3) If $a < 0$ and $\delta = 0$, then the negative–binomial factor equals one and (2.73) reduces to the Poisson PGF $\exp\{(-c/a)(s-1)\}$. The requirement $-c/a \geq 0$ ensures the Poisson mean is nonnegative. Thus G_N is a PGF.

In all three cases the normalization $G_N(1) = 1$ is satisfied by construction (the integration constant in the ODE solution was fixed using that boundary condition), and coefficients of the power series expansion are nonnegative because each factor is the PGF of a nonnegative integer-valued law (or their product). Hence the recursion (2.69) yields a bona fide PMF. □

Remark 2 (Scope and implications)**.** ***Theorem 4*** *provides explicit and easily verifiable* sufficient *conditions guaranteeing that the Schröter recursion generates a proper infinite sequence* $\{P_n\}_{n\geq 0}$*. These conditions cover large parameter regions of practical interest where* G_N *factorizes into well-known PGFs (Poisson, negative-binomial, compound Poisson). They are* not *claimed to be necessary; a complete characterization of all triples* (a,b,c) *that produce a valid PGF beyond these compound–Poisson factorizations remains open.*

Furthermore, the parameter statements in part (1) are consistent: for $0 < a < 1$ *the condition* $c \leq 0$ *is equivalent to* $-c/a \geq 0$*, which is the condition used when interpreting the exponential term as a Poisson PGF.*

Figures 2.1 and 2.2 illustrate the feasible regions in the parameter space (a,b,c) for the Schröter recursion at $n = 2$ and $n = 3$, respectively. Each figure depicts the intersection of the half-spaces defined by the bounding inequalities

$$0 \leq a + \tfrac{b}{n} \leq 1 \quad \text{and} \quad 0 \leq \tfrac{c}{n} \leq 1.$$

In Figure 2.1 ($n = 2$), the four bounding planes

$$a + \tfrac{b}{2} = 0, \quad a + \tfrac{b}{2} = 1, \quad \tfrac{c}{2} = 0, \quad \tfrac{c}{2} = 1$$

form a polyhedral slice of the feasible set. This region captures all admissible parameter values for which the recursion produces nonnegative probabilities at the second step.

In Figure 2.2 ($n = 3$), the analogous bounding planes

$$a + \tfrac{b}{3} = 0, \quad a + \tfrac{b}{3} = 1, \quad \tfrac{c}{3} = 0, \quad \tfrac{c}{3} = 1$$

yield a similar but shifted polyhedral region, now reflecting feasibility constraints at the third step. Together, these figures demonstrate how the feasible region evolves with

n, tightening the admissible parameter space as the recursion progresses. While these conditions guarantee pointwise nonnegativity for fixed n, they do not by themselves ensure global feasibility for the entire infinite sequence.

2.4.4 Relation to the Finite n Regions and Practical Guidance

The finite–n constraints $0 \leq a + \frac{b}{n} \leq 1$ and $0 \leq \frac{c}{n} \leq 1$ are convenient stepwise sufficient conditions that ensure nonnegativity and boundedness locally in n. Imposing them uniformly in n forces $0 \leq a \leq 1$ and $c \geq 0$, which primarily intersects case (2). Cases (1) and (3) show that global feasibility (for the entire sequence) can still hold even if some stepwise coefficients in equation (2.69) are negative (e.g., $c < 0$ in (1)), because feasibility is ultimately determined by the global PGF structure rather than pointwise coefficient bounds.

Note: Given (a, b, c):

1. If $a = 0$ and $b, c \geq 0$, use equation (2.71): N is compound Poisson with a two–point severity.
2. If $0 < a < 1$, compute $\delta = \dfrac{a(a+b)+c}{a^2}$. If $\delta \geq 0$ and $c \leq 0$, use equation (2.69): N is Poisson$(-c/a)$ plus NB(δ, a).
3. If $a < 0$ and $a(a+b)+c = 0$ with $-c/a \geq 0$, then N is Poisson$(-c/a)$.

In all three cases, it should be noted that the full parametric space of the Schröter recursive formula, beyond the compound–Poisson admissible regions identified in **Theorem 4**, may still contain feasible but non–compound–Poisson cases. Characterizing this entire residual space remains an open problem in general, and thus **Theorem 4** settles a substantial portion of the feasibility question while leaving the complete description unresolved.

2.4.5 Asymptotic and Limiting Behavior of the Schröter Recursive Formula

An important question in the study of the Schröter recursion is the long-run behavior of the sequence $\{P_n\}$ as $n \to \infty$.

While the feasibility analysis identifies parameter regions ensuring that the recursion defines a proper PMF, it is also crucial to understand how the probabilities decay in the tail. In particular, the asymptotic form reveals whether the recursion produces light-tailed or heavy-tailed distributions, allowing for comparison with classical discrete families. The following result characterizes the limiting behavior.

Theorem 5. *Let P_n be a claim number distribution satisfying the Schröter recursive formula for $a, b, c \in \mathbb{R}$. For large n, the Schröter recursive formula defines a decreasing sequence converging to zero, with a tail asymptotically equivalent to that of a geometric distribution with base $a \in [0, 1]$ and constant $K \in [0, 1]$.*

Proof. Consider the recursion

$$P_n = \left(a + \frac{b}{n}\right) P_{n-1} + \frac{c}{n} P_{n-2}, \qquad n \geq 2.$$

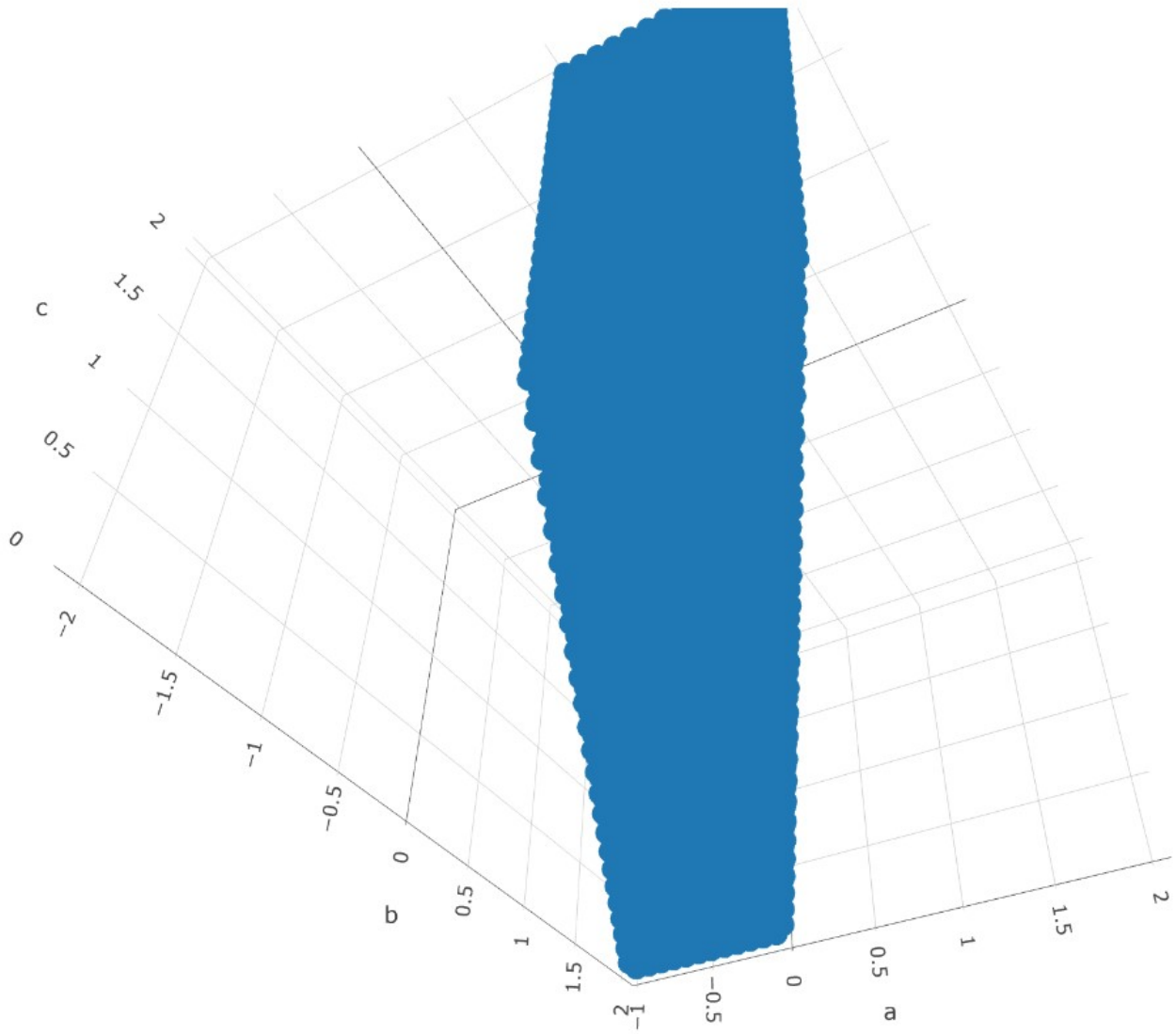


Figure 2.1: The plot of the planes for $n = 2$

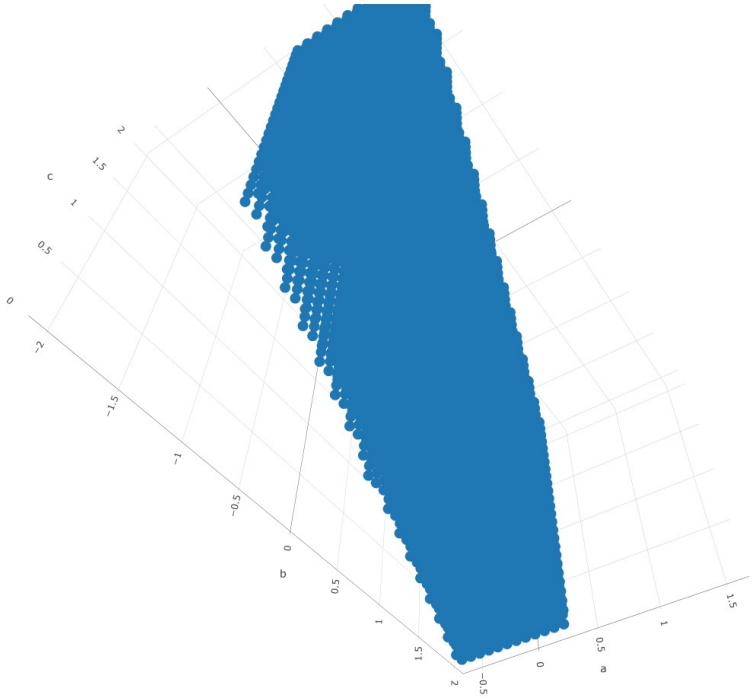


Figure 2.2: The plot of the planes for $n = 3$

As $n \to \infty$, the terms $\frac{b}{n}$ and $\frac{c}{n}$ vanish, so their contributions become negligible. Hence, the recursion reduces to the asymptotic form:

$$P_n \approx aP_{n-1}.$$

Applying this iteratively gives:

$$\begin{aligned} P_{n-1} &= aP_{n-2}, \\ P_{n-2} &= aP_{n-3}, \\ &\vdots \\ P_1 &= aP_0. \end{aligned}$$

Let P_0 be a constant K. By successive substitution we obtain

$$P_n = a^n K.$$

Thus, P_n decays geometrically with base a and scale K (see Figure 2.3). For $0 \leq a < 1$, the sequence decreases monotonically and converges to zero. This behavior aligns with the geometric distribution's tail, confirming the claimed asymptotic equivalence. □

In summary, as $n \to \infty$, the probabilities P_n decrease and converge to zero, with a tail structure analogous to that of the geometric distribution. This provides additional insight into the long-term stability of the Schröter recursion and clarifies its asymptotic contribution within the broader family of recursive distributions.

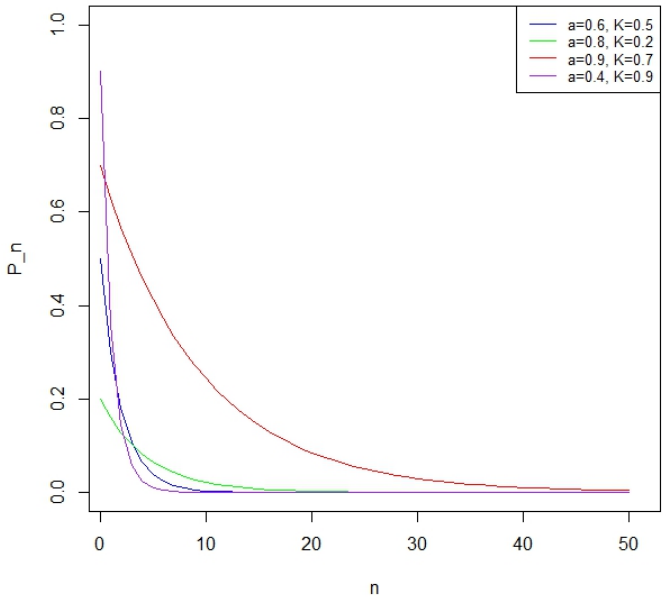


Figure 2.3: The asymptotic decay of P_n as $n \to \infty$.

2.4.6 The Parameters Estimation of the Schröter Recursive Formula

Due to the inherent complexity of the probability mass function defined in equation (2.36), classical methods for parameter estimation do not yield explicit analytical solutions. For instance, the maximum likelihood method yields a system of equations that lacks explicit solutions, necessitating numerical techniques for their resolution.
In light of this, Luong and Garrido [74] proposed an estimation method specifically tailored for recursively defined probability distributions, which is based on minimizing the quadratic distance. The recursive nature of the formula is treated akin to a linear regression model. Subsequently, Luong and Doray [73] and Doray and Haziza [34] further refined and expanded upon this concept. The estimation technique based on minimizing the quadratic distance exhibits desirable properties subject to certain conditions. Nonetheless, the computations still require the utilization of numerical methods and a good knowledge of software, involving operations such as matrix inversion.
Hence, we introduced a simple approach to obtain the parameter estimates for distributions from the Schröter family.

2.4.7 The Parameters Estimate

To obtain the parameters estimate of the Schröter formula, we replaced $E(N)$ and $Var(N)$ in equation (2.31) with their sample counterparts and the parameters with their estimates, we have:

$$\bar{x} = \frac{\hat{a} + \hat{b} + \hat{c}}{1 - \hat{a}}, \tag{2.75}$$

$$s^2 = \frac{\hat{a} + \hat{b} + (2 - \hat{a})\hat{c}}{(1 - \hat{a})^2}. \tag{2.76}$$

Let N represent the sample size, $f_0, f_1, f_2, ...$, the observed frequencies of values $0, 1, 2, ...$, and k the number for which the sum of three neighboring frequencies $f_k + f_{k-1} + f_{k-2}$ attains its maximum. We represent the empirical analog of equation (2.27) with the empirical probabilities $\hat{P}_k, \hat{P_{k-1}}$, and $\hat{P}_k$ (for $\hat{P}_k = \frac{f_k}{N}$) as:

$$\hat{P}_k = \left(\hat{a} + \frac{\hat{b}}{k}\right) \hat{P_{k-1}} + \frac{\hat{c}}{k} \hat{P_{k-2}}, \tag{2.77}$$

$$\begin{aligned} k\hat{P}_k &= \left(k\hat{a} + \hat{b}\right) \hat{P_{k-1}} + \hat{c}\hat{P_{k-2}} \\ &= \left(\hat{a}(k - 1 + 1) + \hat{b}\right) \hat{P_{k-1}} + \hat{c}\hat{P_{k-2}} \\ &= \hat{a}(k-1)\hat{P_{k-1}} + (\hat{a} + \hat{b})\hat{P_{k-1}} + \hat{c}\hat{P_{k-2}} \end{aligned}$$

The system of equations (2.73)-(2.75) is solved to obtain the estimate of the parameters a, b, and c. From equation (2.73), we have:

$$\bar{x}(1 - \hat{a}) - \hat{c} = \hat{a} + \hat{b}. \tag{2.78}$$

Using equation (2.76) in equation (2.75), we have

$$k\hat{P}_k = \hat{a}(k-1)\hat{P_{k-1}} + [\bar{x}(1 - \hat{a}) - \hat{c}]\hat{P_{k-1}} + \hat{c}\hat{P_{k-2}}, \tag{2.79}$$

$$k\hat{P}_k = \hat{a}(k - 1 - \bar{x})\hat{P_{k-1}} + \bar{x}\hat{P_{k-1}} + \hat{c}[\hat{P_{k-2}} - \hat{P_{k-1}}].$$

$$\hat{c} = \frac{k\hat{P}_k - \hat{a}(k - 1 - \bar{x})\hat{P_{k-1}} - \bar{x}\hat{P_{k-1}}}{\hat{P_{k-2}} - \hat{P_{k-1}}}. \tag{2.80}$$

Combining equations (2.74) and (2.76), we have

$$s^2(1 - \hat{a})^2 = (2 - \hat{a})\hat{c} + \bar{x}(1 - \hat{a}) - \hat{c}, \tag{2.81}$$

$$s^2(1 - \hat{a}) = \hat{c} + \bar{x}.$$

By substituting equation (2.78) into equation (2.79) and solving, we have:

$$s^2[\hat{P_{k-2}} - \hat{P_{k-1}}] - as^2[\hat{P_{k-2}} - \hat{P_{k-1}}] = \bar{x}[\hat{P_{k-2}} - \hat{P_{k-1}}] + k\hat{P}_k - \hat{a}[k - 1 - \bar{x}]\hat{P_{k-1}} - \bar{x}P_{k-1}, \tag{2.82}$$

$$(s^2 - \bar{x})[\hat{P_{k-2}} - \hat{P_{k-1}}] - k\hat{P}_k + \bar{x}\hat{P_{k-1}} = \hat{a}[s^2(\hat{P_{k-2}} - \hat{P_{k-1}}) - (k - 1 - \bar{x})\hat{P_{k-1}}].$$

Then,

$$\hat{a} = \frac{(s^2 - \bar{x})[\hat{P_{k-2}} - \hat{P_{k-1}}] - k\hat{P}_k + \bar{x}\hat{P_{k-1}}}{[s^2([\hat{P_{k-2}} - \hat{P_{k-1}}]) - (k - 1 - \bar{x}])\hat{P_{k-1}}}. \tag{2.83}$$

Using equations (2.76), (2.79), and (2.81), we have that:

$$\hat{c} = s^2(1 - \hat{a}) - \bar{x} \tag{2.84}$$

$$\hat{b} = \bar{x}(1 - \hat{a}) - \hat{a} - \hat{c}, \tag{2.85}$$

where $\hat{a}, \hat{b}$, and $\hat{c}$ are the estimates of a, b, and c of the Schröter recursive formula.

2.4.8 Simulation and Real World Application of the Parameter Estimates

The previous subsections focused on the theoretical characterization, feasibility, asymptotic properties, and parameter estimation of the Schröter recursive formula.

To complement these theoretical results, we now turn to empirical evaluation. In particular, we demonstrate the practical application of the parameter estimation method by conducting simulation experiments and applying the approach to real-world car accident claim data. This dual evaluation highlights the statistical properties of the estimators under controlled conditions and their practical relevance in actuarial applications.

2.4.9 Simulation Study

To investigate the finite-sample behavior of the parameter estimates, we conducted a simulation study. We set the initial parameter values as $a = 0.6$, $b = 2.6$, and $c = -1.1$, and generated 100,000 random numbers, repeating this procedure 10,000 times for each parameter from the distribution defined in equation (2.36). The resulting samples were then used to estimate the parameters, and the descriptive statistics of the estimated values are summarized in Table 2.1. This experiment provides insights into the bias, variance, and overall stability of the estimation procedure in large samples.

Table 2.1: Estimates of the Parameters

	$\hat{a}$	$\hat{b}$	$\hat{c}$
Mean	0.600	2.601	-1.103
Standard deviation	0.021	0.085	0.212
Minimum	0.525	2.305	-1.897
Maximum	0.678	2.918	-0.346

Nonetheless, the mean from Table 2.1 shows that the estimates of a and b could be unbiased, but their standard deviation, especially $\hat{c}$, appears not to converge to zero.

2.4.10 Car Accident Injuries Applications

We fit the Schröter recursive formula, defined in equation (2.27), to car accident data from the Olomouc region, Czech Republic, encompassing the period between January 1 and December 31, 2021, and the severity of injuries, including deadly, serious, and minor cases. This data was obtained from the webpage of the Czech public radio broadcaster, accessed on October 10, 2022, and can be found at www.irozhlas.cz/nehody. In Table 2.2, we present the number of days with the corresponding number of injuries from the data.
where x denotes the number of injuries and $f(x)$ represents number of days with x injuries. Using equations (2.81)-(2.83), the model shows a good fit (see Figure 2.4) based on the Pearson chi-square test $p_{value} = 0.1677$ and we obtained $\hat{a} = 0.451, \hat{b} = 1.127$, and $\hat{c} = 0.254$. Although, other methods may be used to obtain improved results using our parameters estimate as initial values.

Figure 2.5 illustrates the bar plot representing the fitted Schröter recursive model to the car accident data based on the parameters estimated.

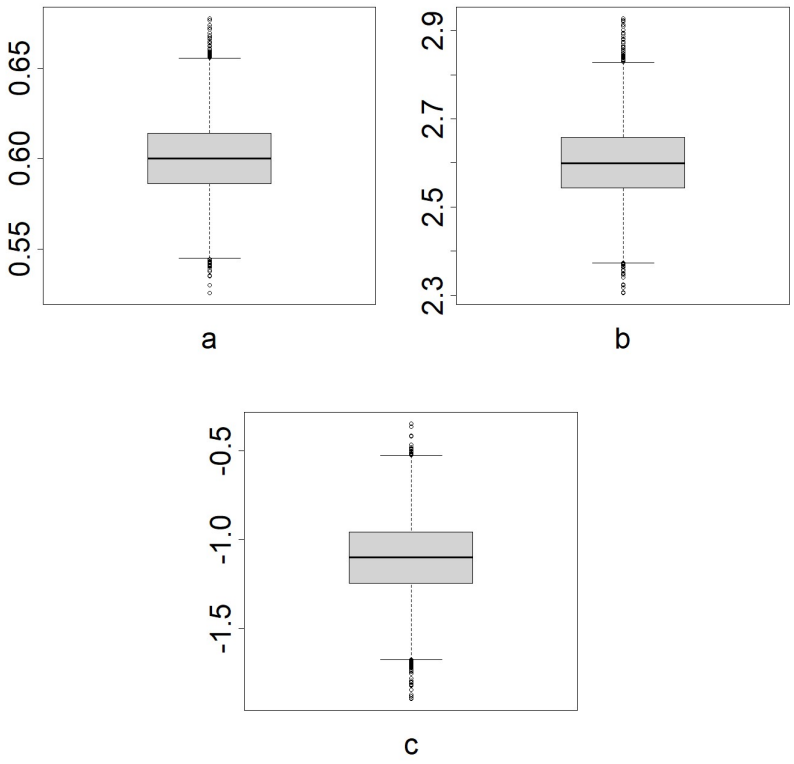


Figure 2.4: Boxplots of parameter estimates for $\hat{a}$, $\hat{b}$, and $\hat{c}$

The simulation study confirms that the proposed estimation method yields stable and reliable estimates for a and b, with sample means close to their true values and relatively small variability (Table 2.1, Figure 2.4). In contrast, the parameter $\hat{c}$ shows higher variability, indicating slower convergence in finite samples, though its mean remains close to the true value $c = -1.1$. These results validate the consistency and robustness of the approach for large samples.

In the real-data application, fitting the Schr"oter recursive model to the car accident injury dataset produced parameter estimates $\hat{a} = 0.451$, $\hat{b} = 1.127$, and $\hat{c} = 0.254$. The fitted distribution closely matched the observed frequencies (Figure 2.5), with the Pearson chi-square test yielding a non-significant result (p-value $= 0.1677$). Thus, this result indicates that the model provides an adequate description of the accident data. Hence, the achieved results indicate that the Schr"oter recursive formula performs well under simulation and provides a flexible and practically applicable model in actuarial settings, with potential for further refinement using alternative estimation techniques.

2.5 An Extended Schröter Recursive Formula

The Schröter recursive model has been widely applied in insurance mathematics to describe the distribution of claim counts within a portfolio over a given period. However, in practice, the occurrence of claims is rarely determined by the recursion alone. Claim frequencies are often influenced by external factors, such as changes in policyholder behavior, macroeconomic conditions, anticipated adjustments in insurance premiums, regulatory interventions, or even social media dynamics that amplify reporting tenden-

Table 2.2: Car Accident Injuries

x	$f(x)$
0	40
1	64
2	60
3	55
4	33
5	39
6	29
7	22
8	8
9	7
10	5
12	3

cies.

To capture these real-world complexities, we extend the Schröter recursion by incorporating systematic explanatory factors that may affect the likelihood of claims being filed. This generalization enables the model to more accurately reflect practical insurance environments and provides a framework for studying finite-period contract scenarios that explicitly consider additional covariates.

Let Y denote the number of discrete claims, and let X_i, $i = 0, 1, \ldots, r$ represent a finite collection of explanatory factors, assumed mutually independent. We define a systematic component as

$$g(\eta) = \beta_0 + \beta_1 X_1 + \cdots + \beta_r X_r, \tag{2.86}$$

where $\beta_0, \beta_1, \ldots, \beta_r$ are parameters. The Schröter recursion in equation (2.37) can then be generalized to:

$$P_n = \left(a + \frac{b}{n}\right) P_{n-1} + \frac{c}{n} P_{n-2} + g(\eta). \tag{2.87}$$

Here, n denotes the finite number of contract periods under consideration, and $g(\eta)$ introduces the contribution of covariates into the recursion.

Since Y is discrete and belongs to the exponential family of distributions, the systematic component can be linked to the mean of Y through a generalized linear model (GLM) framework. Let $g(u) = X^T\beta$ be the canonical link function relating $\mathbb{E}[Y] = u$ to the linear predictor $X^T\beta$, with inverse $u = f(X^T\beta)$. The canonical form of the exponential family density is:

$$f(y; \theta, \Psi) = e^{\left(\frac{y\theta - b(\theta)}{a(\Psi)} - c(y;\Psi)\right)}, \tag{2.88}$$

where θ is the natural parameter and Ψ the dispersion parameter.

For example, if Y follows a Poisson distribution with PMF

$$f(y; \lambda) = e^{(y \ln \lambda - \lambda - \ln y!)}, \tag{2.89}$$

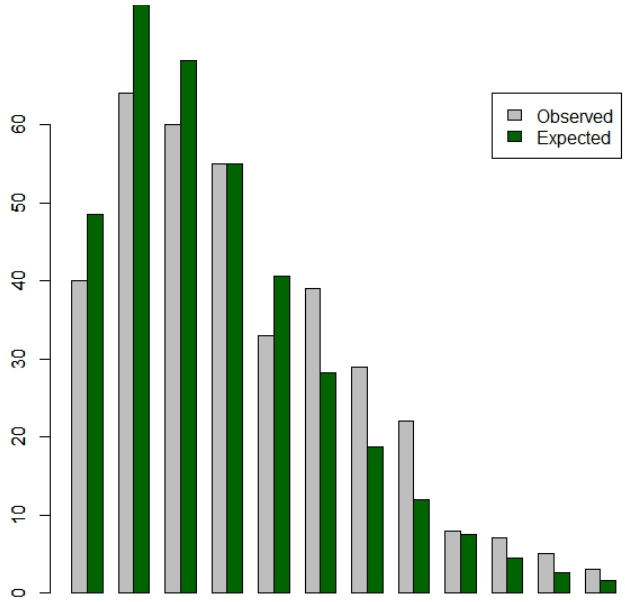


Figure 2.5: The graph of the fitted Schröter model to the accident dataset

then $y \ln \lambda = y\theta$ and $\lambda = e^{\theta}$ is the Poisson link. Setting $\theta = X^T\beta$ gives

$$\lambda = e^{X^T\beta}. \tag{2.90}$$

This representation connects the recursion-based model with regression structures, allowing for the inclusion of explanatory factors in the Schröter framework.

2.6 Numerical Evaluation of the Extended Schröter Recursive Formula

The previous section introduced an extension of the Schröter recursive formula that incorporates explanatory factors through a systematic component, thereby linking the recursion to generalized linear models. In this section, we present the numerical evaluation of the extended model. The goal of this section is twofold: first, to demonstrate how real-world insurance data can be incorporated into the Schröter recursive framework; and second, to illustrate the role of discretization and parameter estimation in ensuring the model's practical applicability.

We apply the methodology to the "InsClaim | EDA | Modeling" dataset consisting of 1341 observations obtained from https://www.kaggle.com/code/ravivarmaodugu/insclaim-eda-modeling/notebook. For this analysis, we focus only on the *claim observations* as the response variable and consider two covariates as explanatory factors:

- body mass index (BMI), denoted by X_1,
- blood pressure, denoted by X_2.

2.6.1 Discretization of Continuous Data

In actuarial practice, claim data are often modeled using continuous distributions (see, e.g., Hogg and Klugman [57], Panjer and Willmot [90]). However, recursive methods such as the Schröter recursive formula inherently operate on discrete distributions. This necessitates replacing continuous distributions with discrete analogues that preserve the key statistical properties of the original data.

Several discretization techniques have been proposed in the literature. In this study, we adopt the simple yet effective *crude rounding method* discussed by Panjer and Lutek [89], whereby claim observations are rounded to the nearest integer values. Although more sophisticated discretization techniques exist, such as flooring, ceiling, moment-matching, or quantile-based approaches, the crude rounding method offers several advantages that make it particularly suitable in our setting. First, it preserves the natural scale and distributional shape of the original claim data without imposing artificial shifts or distortions. Second, it ensures that the discretized data remain close to the observed values, which is important for retaining interpretability when modeling real claims. Third, the recursive formulas we employ, such as equations (2.49)–(2.51), are highly sensitive to the choice of discretization method: rounding minimizes approximation error when computing empirical probabilities P_n, P_{n-1}, and P_{n-2}, thereby improving numerical stability in recursion. Finally, given the large sample size of our dataset, the law of large numbers ensures that rounding yields results that closely approximate those obtained from more complex discretization schemes, but with considerably less computational burden. For these reasons, the crude rounding method provides a suitable, practical, and reliable balance between theoretical rigor, computational efficiency, and empirical accuracy in the present study.

The histogram of the discretized claims data is presented in Figure 2.6, with an empirical mean of approximately $1,325$ and a variance of $1,466$.

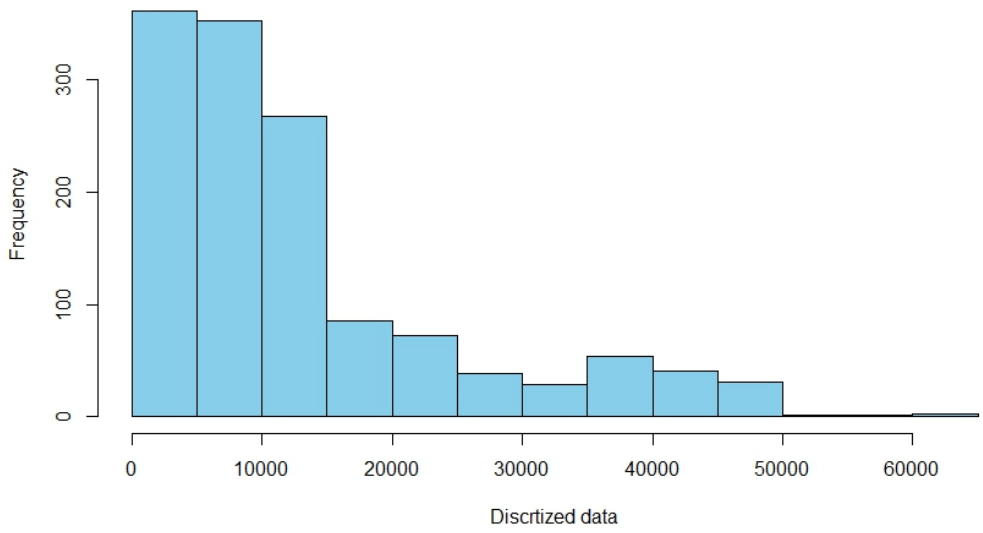


Figure 2.6: Histogram of the discretized claim observations.

From the discretized dataset, the empirical probabilities P_n, P_{n-1}, and P_{n-2} were

computed using a moving average method. Specifically, P_{n-2} was based on the first $(n-2)$ periods, P_{n-1} on the first $(n-1)$ periods, and P_n on the first n periods. Normalizing such that $P_n + P_{n-1} + P_{n-2} = 1$, we obtained $P_n = 0.3345$, $P_{n-1} = 0.3333$, and $P_{n-2} = 0.3322$.

Using the discretized data and equations (2.81)–(2.83) (with k replaced by n), we obtained the parameter estimates

$$\hat{a} = 0.9999, \qquad \hat{b} = 1375.603, \qquad \hat{c} = -1375.53.$$

Additionally, we employed the generalized linear model framework with a Poisson link function for the parameter estimation of the regression component. We obtained the estimated regression coefficients for the covariates as:

$$\hat{\beta}_0 = 5.9856, \qquad \hat{\beta}_1 = 0.0183, \qquad \hat{\beta}_2 = 0.0304.$$

The extended recursion with covariates is therefore expressed as

$$q_n \;=\; \vartheta \left[\left(\hat{a} + \frac{\hat{b}}{n} \right) \hat{P}_{n-1} + \frac{\hat{c}}{n} \hat{P}_{n-2} + \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 \right], \tag{2.91}$$

where ϑ is the normalization constant ensuring that q_n remains a probability for each contract period n.

To ensure that q_n defines a valid probability in $[0, 1]$, a natural choice of ϑ is obtained by mapping the linear predictor onto the unit interval via a logistic transformation. In this case, we set

$$\vartheta \;=\; \frac{1}{1 + \left[\left(\hat{a} + \frac{\hat{b}}{n} \right) \hat{P}_{n-1} + \frac{\hat{c}}{n} \hat{P}_{n-2} + \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 \right]},$$

so that

$$q_n \;=\; \frac{\left(\hat{a} + \frac{\hat{b}}{n} \right) \hat{P}_{n-1} + \frac{\hat{c}}{n} \hat{P}_{n-2} + \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2}{1 + \left(\hat{a} + \frac{\hat{b}}{n} \right) \hat{P}_{n-1} + \frac{\hat{c}}{n} \hat{P}_{n-2} + \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2}.$$

This construction guarantees that $q_n \in (0, 1)$ for all n, ensuring a well-defined probability of reporting or filing a claim in each contract period.

2.6.2 Numerical Illustration

To illustrate, substituting $X_1 = 19.5$ (BMI) and $X_2 = 120$ (blood pressure) into equation (2.89) for $n = 1$, and applying normalization, yields the probability of a claim being reported as: $\vartheta = 0.0778$ and $L = \left(\hat{a} + \frac{\hat{b}}{n} \right) \hat{P}_{n-1} + \frac{\hat{c}}{n} \hat{P}_{n-2} + \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 = 11.8611$. Then,

$$q_1 = 0.9222.$$

This demonstrates how the extended model can incorporate individual risk factors into the recursive formula, thereby providing a more realistic assessment of claim likelihood.

Finally, Figure 2.7 presents the sequence of probabilities q_n for $n = 1{:}200$. It is evident that as the contract period n increases, the probability of reporting a claim grows proportionally, reflecting the accumulation of risk over time.

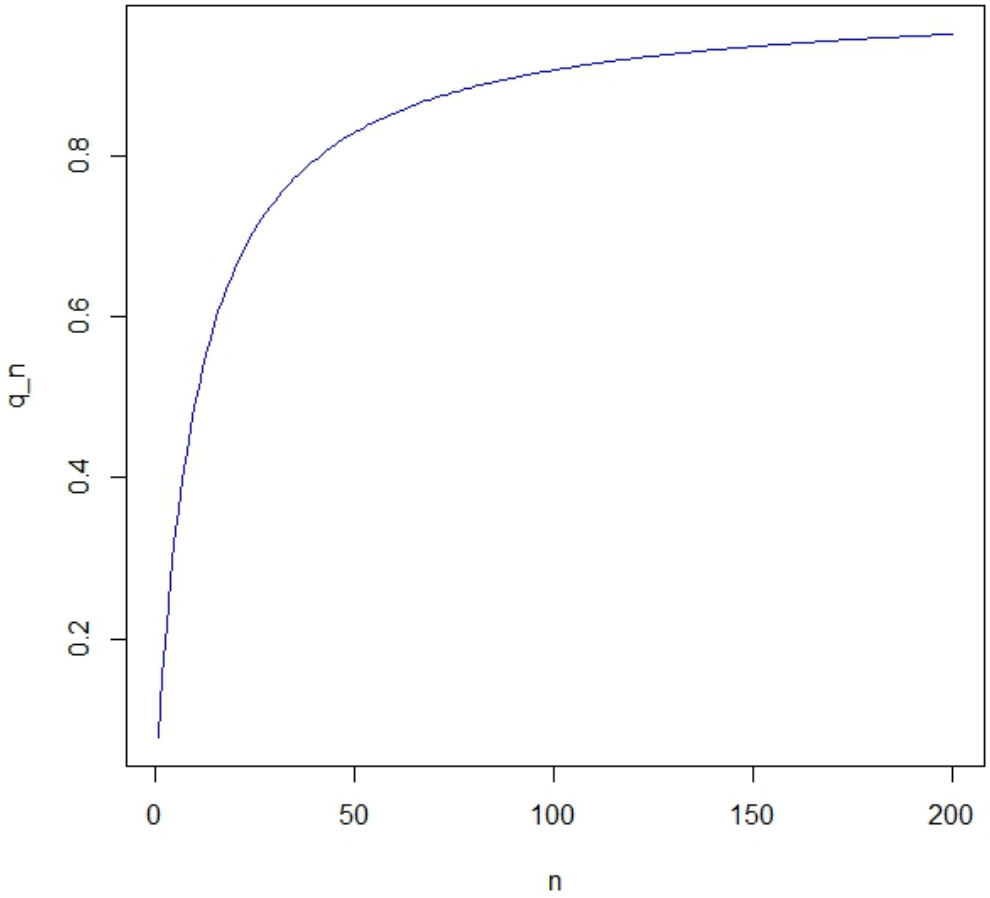


Figure 2.7: Probability sequence q_n of reporting a claim as contract periods n increase.

The extended Schröter recursion with covariates successfully integrates explanatory factors into the recursive structure while ensuring probabilistic validity through a logistic normalization. Applied to the insurance claims dataset, the approach yielded consistent parameter estimates for the recursive and regression components. For $n = 1$, the evaluation yielded a normalized probability of $q_1 = 0.9222$, indicating that the model can accurately capture the high likelihood of claims when individual risk factors (BMI and blood pressure) are considered. The probability sequence q_n further demonstrates a monotonic growth with n, reflecting the nature of risk exposure over contract periods.

This section contributes by bridging classical recursion with modern regression modeling:

1. It establishes a framework that allows recursive claim count models to explicitly incorporate policyholder characteristics and external factors.

2. It introduces a normalization scheme that guarantees well-defined probabilities across finite contract periods.

3. It provides an empirical illustration showing that the model not only fits real insurance data but also improves interpretability and predictive capacity compared to the recursion alone.

This extension enhances the Schröter recursive formula from a purely structural model into a flexible, data-informed tool for actuarial risk assessment and prediction.

2.7 Final Remarks

In summary, this chapter has presented a comprehensive discussion of recursive approaches in collective risk theory, with a particular focus on the Panjer and Schröter recursive formulas. Beginning with the Panjer recursion, the chapter highlighted its importance as an exact and computationally efficient alternative to convolution-based methods. The analysis then extended to the Schröter recursive formula, a second-order generalization of the Panjer class, which provides a richer framework for modeling claim count distributions and the major focus of the chapter.

Several new theoretical and practical results were developed. The feasible regions for the parameters of the Schröter recursion were investigated in both finite- and infinite-n cases, and the relationship between these regions was established. The asymptotic behavior of the Schröter recursion as $n \to \infty$ was also examined, showing that the probabilities decrease geometrically and converge to zero, with a tail behavior equivalent to that of the geometric distribution. This finding not only clarifies the long-run stability of the recursion but also connects it to a well-known family of distributions.

The chapter further introduced new results on parameter estimation for the Schröter family, drawing on our joint work in Agu et al. [4]. A simple moment-matching approach was proposed, and its performance was evaluated using both simulated and real-world data, including car accident injury counts. The results confirmed that the method provides reliable and interpretable parameter estimates, while also highlighting areas where further refinements could improve precision, particularly for certain parameters.

Beyond these contributions, the chapter also advanced the field by extending the Schröter recursion into a more flexible, data-informed model. By embedding the recursion within the GLM framework and applying a logistic normalization scheme, the extended model was demonstrated to incorporate explanatory factors, such as policyholder characteristics and external influences, while ensuring that the resulting probabilities remain valid across finite contract periods. Applied to real insurance datasets, this approach improved interpretability, predictive accuracy, and practical relevance compared to the structural recursion alone.

Taken together, the developments presented in this chapter emphasize the central role of recursion in modern actuarial modeling. They show that recursive formulas are not only powerful computational devices for aggregate claims but can also serve as adaptable tools for capturing real-world complexities in insurance data. These insights provide a strong foundation for the subsequent chapter, where further refinements and applications of recursive methods will be of interest.

CHAPTER 3

THE TRUNCATED SCHRÖTER RECURSIVE FORMULA

Historically, before the advent of modern computing, actuaries relied primarily on estimation and approximation techniques that lacked a rigorous theoretical foundation for determining aggregate claim amounts. These methods were limited in accuracy and reliability, making data-driven decision-making in insurance a challenging task.

In risk theory, truncated distributions play vital roles not only in modeling claim severity but also in modeling the inter-arrival times between successive claims and excess-of-zero claims. These distributions enable insurers and actuaries to gain a comprehensive understanding and accurate estimation of the frequency and severity of potential losses, making them vital for policy pricing and risk management. Applications of truncated distributions to modeling insurance problems have been extensively utilized in Kreer et al. [67], Zhang et al. [133], Hocquard et al. [55], Hamedani et al. [51], and Ghaddab et al. [46].

Despite the utility of the Schröter recursive formula, it does not fully capture the dynamics of claim amounts truncated at one and the utilization its recursion algorithm for computation of aggregate claim amounts still requires the convolution component which demand extensive computational time.. This practice holds significant practical relevance in real-world insurance settings. In many cases, insurers are primarily concerned with the number of events that generate claims, rather than the exact amounts. Once a claim is reported, the minimum observed claim amount is often truncated at one, effectively implying a zero probability for a claim amount of zero. This reflects typical policy structures that include deductibles, where insured individuals are responsible for losses below a certain threshold, and only the excess is reimbursed. Consequently, minor losses below the deductible are frequently unreported, making one the effective lower bound for observed claim amounts. This truncation has a substantial impact on the modeling of risk exposure, influencing both the accuracy of risk assessment and the determination of premium rates. In risk theory, truncated distributions are essential for modeling claim severities and inter-arrival times, providing insurers and actuaries with critical tools to better understand the frequency and magnitude of losses. As such, accurately modeling the number of claims truncated at one is vital for capturing the true nature of insurance liabilities.

3.1 Truncated Schröter Recursive Formula

In this section, our goal is to introduce an alternative, less time-consuming truncated Schröter recursive algorithm for the computation of aggregate claim amounts. To achieve this, we first present the truncated Schröter recursive formula and introduce the corresponding PMF and the maximum likelihood function.

We expressed the truncated Schröter recursive formula as:

$$P_n = \left(a + \frac{b}{n}\right) P_{n-1} + \frac{c}{n} P_{n-2}, \qquad n = 2, 3, ... \tag{3.1}$$

Equation (3.1) is truncated at 1 and a, b, and c are the parameters.

First, let N be a discrete random variable taking non-negative integer values as defined in equation (3.1) and using the fact that the PGF is defined as:

$$G_N(s) = \sum_{n=0}^{\infty} s^n P_n$$

for $s \in [0, 1]$ such that $G(s) \geq 0$ and P_n is the recursive probability defined in equation (3.1) and $\sum P_n = 1$.

Thus, the PGF corresponding to equation (3.1) is:

$$G_N(s) = e^{\frac{-c}{a}(s-1)} \left(\frac{1-a}{1-as}\right)^{\delta},$$

where $\delta = \left(\frac{a(a+b)+c}{a^2}\right)$ for $|as| \neq 1$ and Schröter [103] derived the truncated PMF $\{f_n\}_{n=0,1,...}$ in the case $a \neq 0$ corresponding to equation (2.27) with the general expression given as:

$$f_n = e^{\frac{c}{a}(1-a)\delta} \sum_{i=0}^{n} \binom{\delta + i - 1}{i} \frac{\left(-\frac{c}{a}\right)^{n-i} a^i}{(n-i)!} \tag{3.2}$$

and for $a = 0$, equation (3.2) becomes more cumbersome. Additionally, for $n = 0$, the sum has only the $i = 0$ term. That is:

$$\binom{\delta - 1}{0} = 1, \quad \left(-\tfrac{c}{a}\right)^0 = 1, \quad (0)! = 1;$$

hence,

$$f_0 = e^{\frac{c}{a}(1-a)\delta}.$$

Using the fact that:

$$Q_n = \frac{f_n}{1 - f_0},$$

we derived the truncated PMF $\{Q_n\}_{n=1,2,...}$ corresponding to equation (3.1) as:

$$Q_n = \frac{e^{\frac{c}{a}}(1-a)^{\delta} \sum_{i=0}^{n} \binom{\delta+i-1}{i} \frac{\left(-\frac{c}{a}\right)^{n-i} a^i}{(n-i)!}}{1 - e^{\frac{c}{a}}(1-a)^{\delta}}, \qquad n = 1, 2, \ldots, 0 < a < 1, b, c \in \mathbb{R}. \tag{3.3}$$

Let

$$x = \frac{c}{a},$$

and define

$$S_n = \sum_{i=0}^{n} \binom{\delta + i - 1}{i} \frac{(-x)^{n-i} a^i}{(n-i)!}. \tag{3.4}$$

Then, the PMF can be written in a simple form as:

$$Q_n = \frac{e^x (1-a)^\delta \, S_n}{1 - e^x (1-a)^\delta}.$$

The Log-likelihood for the truncated Schröter recursive PMF

Let

$$Q_n = \frac{e^x (1-a)^\delta \, S_n}{1 - e^x (1-a)^\delta}, \qquad x = \frac{c}{a},$$

where S_n is any known function of n, a, δ, x as defined in equation (3.4) (for example, the finite sum or hypergeometric representation).

Assume $n_1, \ldots, n_k$ are independent observations from Q_n. The likelihood is

$$L(a, \delta, x \mid n_1, \ldots, n_k) = \prod_{j=1}^{k} Q_{n_j},$$

so the log-likelihood is

$$\ell(a, \delta, x \mid n_1, \ldots, n_k) = \sum_{j=1}^{k} \log Q_{n_j}.$$

Substituting Q_n gives

$$\begin{aligned}
\ell(a, \delta, x) &= \sum_{j=1}^{k} \Big\{ \log \left(e^x (1-a)^\delta\right) + \log S_{n_j} - \log \left(1 - e^x (1-a)^\delta\right) \Big\} \\
&= \sum_{j=1}^{k} \left(x + \delta \log(1-a) - \log(1 - e^x (1-a)^\delta)\right) + \sum_{j=1}^{k} \log S_{n_j} \\
&= k\big[x + \delta \log(1-a) - \log(1 - e^x (1-a)^\delta)\big] + \sum_{j=1}^{k} \log S_{n_j}.
\end{aligned}$$

Equivalently, for a single observation n:

$$\ell^{(1)}(a, \delta, x \mid n) = x + \delta \log(1-a) - \log(1 - e^x (1-a)^\delta) + \log S_n.$$

3.1.1 The Truncated Schröter Recursive Algorithm

Let $X_1, X_2, \ldots X_n$ be independent and identitically distributed claim severities over the non-negative integers with probability density $f_k = P(X_i = k)$ for $i = 1, 2, \ldots n; k = 0, 1, 2, \ldots$, and $f^{k*} = P(X_1 + X_2 + \cdots + X_n = k)$ denotes the n-fold convolution of f_k.

Additionally, let N be a discrete random variable representing the number of claims with a discrete probability mass function defined as $P_n = P(N = n)$, such that X_i are stochastically independent of N, and $S = \sum_{i=1}^{N} X_i$ is the aggregate claim.

For all severity distribution f^n, we derived the truncated Schröter recursive algorithm as:

$$g(s) = \sum_{n=2}^{\infty} P_n f^{n*}(s) \tag{3.5}$$

where P_n is the truncated Schröter recursive formula defined in equation (3.1).

The Panjer recursion formula is based on the expression:

$$f^{n*}(s) = \frac{n}{s} \sum_{i=1}^{s} i f_i f_{s-i}^{n-1}, \quad s = 1, 2, \ldots; n = 1, 2, \ldots$$

(see Schröter [103]; page 164).

Suppose that $f_n^0 = 1$ for all $n > 0$, We can write that:

$$f(s) = \frac{1}{s} \sum_{i=1}^{s} i f_i. \tag{3.6}$$

Based on the truncation convention $f_{s-i}^0 = 1$ for all i. This convention is chosen to avoid the collapse of the sum to the single $i = s$ term.

$$g(s) = \sum_{n=2}^{\infty} \left[(a + \frac{b}{n}) P_{n-1} + \frac{c}{n} P_{n-2} \right] f^{n*}(s) \tag{3.7}$$

$$g(s) = a \sum_{n=2}^{\infty} P_{n-1} f^{n*}(s) + b \sum_{n=2}^{\infty} \frac{1}{n} P_{n-1} f^{n*}(s) + c \sum_{n=2}^{\infty} \frac{1}{n} P_{n-2} f^{n*}(s)$$

$$g(s) = a \sum_{n=1}^{\infty} P_n f^{n*}(s) + b \sum_{n=1}^{\infty} \frac{1}{n} P_n f^{n*}(s) + c \sum_{n=2}^{\infty} \frac{1}{n} P_{n-2} f^{n*}(s)$$

$$g(s) = a \sum_{n=1}^{\infty} P_n \left(\frac{n}{s} \sum_{i=1}^{s} i f_i f_{s-i}^{n-1} \right) + b \sum_{n=1}^{\infty} \frac{P_n}{n} \left(\frac{n}{s} \sum_{i=1}^{s} i f_i f_{s-i}^{n-1} \right) + c \sum_{n=2}^{\infty} \frac{P_{n-2}}{n} \left(\frac{n}{s} \sum_{i=1}^{s} i f_i f_{s-i}^{n-1} \right) \tag{3.8}$$

We set $\left(\sum_{i=0}^{s} f_i f_{s-i}^{n-1} \right)$ as $\sum_{i=0}^{s} f_i g_{s-i}$ since g_{s-i} is computed recursively and as $P_1 = 0$, we avoid any term involving it. This expression facilitates easier and more efficient computation, avoiding the need for complicated convolutional components.

We have that

$$g(s) = a\sum_{n=0}^{\infty} P_n\left(\sum_{i=0}^{s} f_i g_{s-i}\right) + b\sum_{n=0}^{\infty} P_n\left(\frac{i}{s}\sum_{i=0}^{s} f_i g_{s-i}\right) + c\sum_{n=1}^{\infty} P_{n-1}\left(\frac{i}{s}\sum_{i=0}^{s} f_i g_{s-i}\right)$$

We can write the first two terms as:

$$a\sum_{i=0}^{s} f_i g_{s-i}$$

and

$$\sum_{i=0}^{s} \frac{bi}{s} f_i g_{s-i}$$

From the third term, we re-indexed as follows: $n \mapsto m = n-1$ and set $m = n-1 \implies n = m+1$ and if $n = 1$, we have $1 = m+1 \implies m = 0$.

Note that the sum over $m = -1$ to ∞ is:

$$c\sum_{m=-1}^{\infty} P_m\left(\frac{i}{s}\sum_{i=0}^{s} f_i g_{s-i}\right)$$

which corresponds to the P_{-1}, which is 0. Thus, we have:

$$c\sum_{m=0}^{\infty} P_m\left(\frac{i}{s}\sum_{i=0}^{s} f_i g_{s-i}\right)$$

Since $m \implies n$, we write:

$$c\sum_{n=0}^{\infty} P_m\left(\frac{i}{s}\sum_{i=0}^{s} f_i g_{s-i}\right)$$

Hence,

$$g(s) = a\sum_{i=0}^{s} f_i g_{s-i} + \sum_{i=0}^{s} \frac{bi}{s} f_i g_{s-i} + \sum_{i=0}^{s} \frac{ci}{s} f_i g_{s-i} \tag{3.9}$$

Applying the weight $\frac{1}{1-af_0}$, the truncated Schröter recursive algorithm is expressed as:

$$g(s) = \frac{1}{1-af_0}\sum_{i=0}^{s}\left[a + \frac{i}{s}(b+c)\right] f_i g_{s-i} \tag{3.10}$$

for $s \neq 0$ and a, b, and c are the parameters. Additionally, $f_0 = \mathbb{P}(S=0) = 0$ and $g(0) = P(N=0)$ is the initial probability. If $c = 0$ in equation (3.10), we obtain the Panjer recursive algorithm defined in equation (2.3), and if we define $f^{n*}(s)$ as:

$$f^{n*}(s) = \frac{n}{ks}\sum_{i=1}^{s} i\, f_i^{k*}\, f_{s-i}^{(n-k)*}, \qquad \text{for } k \in \{1,2,\ldots,n\} \qquad n > k \tag{3.11}$$

in equation (3.7), equation (2.25) becomes a special case of equation (3.10). Equation (3.10) does not require the evaluation of the convolution component $f^{n*}(s)$ as in the Schröter recursive algorithm defined in Schröter [103].

To execute equation (3.10), we treat f_i as the claim frequencies per policy. To ensure numerical stability and convergence of equation (3.10), the parameters a, b, and c were estimated via maximum likelihood of equation (3.3) using the `nlminb()` optimizer with box constraints: $0 < a < 1$, $b \geq 0$ and $c \in \mathbb{R}$. These constraints prevent instability in the recursion weights and guarantee the validity of the logarithmic expressions in the likelihood function.

Theoretically, unlike (2.25), the recursion algorithm defined in (3.10) eliminates the need for any form of convolution.

Additionally, we consider the Negative binomial distribution as the count distribution for the number of claims (see Section 3.1.3, Table 3.4).

The probability mass function for the Negative binomial distribution is defined as

$$h(S = s) = \binom{s + r - 1}{s}(1 - p)^s p^r, \qquad s = 0, 1, 2, \ldots, \ r > 0, \ p \in [0, 1].$$

We have that

$$h(S = 0) = \binom{r - 1}{0} p^r.$$

From (3.9), we define

$$g(0) = P_0 = h(S = 0) = p^r. \tag{3.12}$$

3.1.2 The Numerical Implementation Procedure for the Truncated Schröter Algorithm

The implementation of the truncated Schröter recursive algorithm involves several computational stages designed to estimate the distribution of aggregate claim amounts. The procedure can be summarized as follows:

1. **Data Preparation:** Obtain and clean claim count data (from real-world and simulations). Compute the empirical frequency distribution f_i and normalize it to ensure $\sum f_i = 1$. Furthermore, the distribution of the data is determined (see Table 4). The parameters in equation (3.10) are estimated using the MLE method based on the observed data. The log-likelihood function is constructed from the truncated probability mass function of the claim counts. Parameter estimation is carried out using the `nlminb()` optimizer in R, which is well-suited for bounded, nonlinear optimization problems. This approach ensures numerical stability and facilitates the explicit enforcement of parameter constraints that are critical to the recursive structure of the model. A similar approach is applied by truncating the corresponding probability mass functions of the Panjer and Schröter families (see Panjer [88]; Schröter [103]).

2. **Initialization:** Determine $g(0)$ as in equation (3.12) and initialize a numeric vector to store $g(s)$ for $s = 1, 2, \ldots, n$.

3. **Recursive computation:** For each $s = 2, \ldots, n$, compute $g(s)$ using equation (3.10).

4. **Performance evaluation:** Evaluate the sum of $g(s)$ values and record execution time per iteration to assess computational efficiency.

5. **Visualization:** Utilize graphical tools (e.g., bar plots, execution time plots) to display the algorithm's output and benchmark it against Panjer and standard Schröter methods.

3.1.3 Numerical Evaluation of the Truncated Schröter Recursive Algorithm

In this section, we examine the run-time computational efficiency of the introduced truncated Schröter algorithm using the Automobile UK Collision Claims (AutoCollision) data obtained from https://instruction.bus.wisc.edu/jfrees/jfreesbooks/RegressionFirst, we began by exploring the descriptive statistics of the dataset, analyzing average claim severity and average claim counts across various age groups and vehicle categories to identify patterns and determine how frequently each group files claims. Particular attention was given to combinations of age groups and vehicle use categories associated with high claim severity and frequency, as these represent higher risk factors for insurers and may necessitate adjustments in insurance coverage strategies. To model the claim count data, we fitted both the Negative Binomial and Generalized Poisson distributions, selected for their ability to handle overdispersion commonly observed in count data. The choice between these distributions was guided by model fit, using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) to select the model with the lowest values. Furthermore, to implement the truncated algorithm defined in equation (3.10), we employed the truncated probability mass function introduced in (3.3) to obtain numerical estimates of the parameters $\hat{a}, \hat{b}$, and $\hat{c}$ as $\hat{a} = 0.99070, \hat{b} = 1.29297$, and $\hat{c} = 0.29330$ using the MLE method.

In this section, we first present the descriptive analysis of the Automobile UK Collision Claims data and then present the results for the fitted Negative binomial and generalized Poisson distributions applied to the data.

Table 3.1: Descriptive Statistics

	Min.	Max.	Mean	Variance	Kurtosis	Skewness
Values	5.00	970.00	279.44	58374.38	4.08	1.25

The descriptive statistics offer a comprehensive summary of the dataset's distribution and central tendency. The minimum and maximum values define the data range, while the mean provides a central value around which the data are distributed. The high variance indicates substantial variability (overdispersion), and the positive skewness and kurtosis indicate a right-skewed distribution with the presence of outliers.

Table 3.2 shows that vehicles used for business purposes exhibit the highest average claim severity. Although the claim count in this category is relatively low compared to others, each claim carries a substantial financial impact, indicating that business use presents a higher risk of costly claims.

Vehicles used for long-distance driving exhibit a moderate average claim severity, which is substantially lower than that of business use but higher than for short drives

Table 3.2: Analysis of Average Claim Severity by Vehicle Use

Vehicle Use	Average Claim Severity	Claim Count
Business	395.21	1075
DriveLong	265.26	2710
DriveShort	231.74	3888
Pleasure	213.20	1269

and pleasure use. The relatively high claim count indicates that long drives are associated with frequent incidents, though each claim tends to be less severe than those in the business category.

Short drives register the highest claim count but a lower average claim severity. This indicates that while short trips result in more frequent claims, the financial impact of each is comparatively minor. The high frequency highlights a notable number of incidents with less severe consequences per occurrence.

Pleasure use is associated with the lowest average claim severity and a relatively low claim count, indicating that leisure driving poses the least risk. It results in both fewer claims and lower financial losses, making it the lowest-risk category in terms of both frequency and severity in the UK Automobile Collision Claims dataset.

Table 3.3: Analysis of Average Claim Severity by Age

Age	Average Claim Severity
17–20	391.80
21–24	293.17
25–29	284.84
30–34	279.73
35–39	212.43
40–49	249.99
50–59	251.11
60+	247.68

Table 3.3 shows that drivers aged 17–20 have the highest average claim severity, indicating that accidents involving the youngest drivers tend to result in greater financial losses and represent a substantial risk to insurers. A notable decrease in average claim severity is observed among drivers aged 21–24, indicating a reduced but still relatively high financial risk as drivers gain minimal experience. The trend of decreasing claim severity continues in the 25–29 age group, reflecting a further decline in financial impact as drivers mature and gain experience. This downward trend persists in the 30–34 age group, with a slight reduction in average claim severity compared to the previous cohort. A substantial drop is observed in the 35–39 age group, indicating a much lower severity of claims and a correspondingly reduced financial risk. Interestingly, the 40–49 age group sees a modest increase in average claim severity compared to the 35–39 group, though it remains lower than that of drivers under 30, indicating a moderate financial risk. Claim severity levels for the 50–59 age group are comparable to those of the 40–49 cohort, pointing to a stable level of financial risk among middle-aged drivers. Finally, drivers aged 60 and above exhibit slightly lower average claim severity than

the 50–59 group, indicating a consistent and moderate financial risk, marginally higher than that of the 35–39 group but lower than the younger cohorts.

Table 3.4: The fitting of Negative Binomial and Generalized Poisson distributions

	Negative Binomial	**Generalized Poisson**
Parameters Estimate	$\hat{p} = 0.00453,\ \hat{r} = 1.25042$	$\hat{\theta} = 12.78279,\ \hat{\lambda} = 0.95426$
N-Loglikelihood	-211.9633	-216.1883
AIC	427.9267	436.3767
BIC	430.8581	439.3081

As shown in Table 4, the Negative Binomial distribution yields a higher (i.e., less negative) log-likelihood and the lowest AIC and BIC values, clearly indicating a superior fit to the AutoCollision claim count data compared to the Generalized Poisson distribution. These results indicate that the Negative Binomial model is more appropriate for capturing the underlying data structure. Both AIC and BIC are essential for model selection, as they strike a balance between goodness-of-fit and model complexity, thereby mitigating the risk of overfitting, a particularly important consideration in actuarial modeling.

Beyond information criteria, residual diagnostics further validate this conclusion. A comparative analysis of the Negative Binomial (see Figure 5) and Generalized Poisson models (see Figure 6) reveals that the former produces Pearson and deviance residuals tightly clustered around zero, with minimal dispersion and no extreme outliers. The histogram of Pearson residuals is approximately symmetric and unimodal. In contrast, the Q–Q plot of deviance residuals aligns closely with the theoretical quantile line, indicating that the model assumptions are well met. In contrast, the Generalized Poisson model exhibits more dispersed residuals, noticeable outliers, a skewed residual histogram, and a Q–Q plot that substantially deviates from the reference line, indicating potential model misspecification. Taken together, these statistical and graphical diagnostics confirm that the Negative Binomial model provides a more accurate and reliable representation of the claim count data, establishing it as the preferred modeling choice for this analysis.

3.1.4 Computation of Aggregate Claim

In this section, we used the estimates of $\hat{a}, \hat{b}$, and $\hat{c}$ for equations (2.3), (2.25), and (3.10) to compute aggregate claim amounts for each recursive algorithm and their computational run time using claim count data from the `AutoCollision` data. Based on Table 3.5, the performance of the three recursive algorithms is analyzed in terms of computational run time and the computed aggregate claim amounts.

The truncated Schröter recursion algorithm demonstrates the fastest run time at 0.051093 seconds and yields the highest aggregate claim sum of 0.005483, indicating that it either captures more aspects of the claim data or incorporates a more comprehensive modeling approach (see Figure 3.1). This result implies a probability of approximately 0.55% that the total claim amount will not exceed 32 units, and conversely, a 99.45% probability that it will exceed this threshold.

The Panjer recursion algorithm, with a slightly longer runtime of 0.060898 seconds, computes an aggregate claim sum of 0.004887. Although still efficient, this result may

indicate a more conservative or less data-sensitive approach (see Figure 3.2). The corresponding probability that the total claim amount does not exceed 32 is approximately 0.49%, implying a 99.51% chance of exceeding this amount.

The standard Schröter recursion algorithm, which has the longest runtime at 0.173438 seconds, produces an aggregate claim sum of 0.004930. This outcome implies a balance between sensitivity and comprehensiveness; however, it comes with higher computational demands due to the convolution component involved in the algorithm (see Figure 3.3). The probability that the total claim amount will not exceed 32 units is approximately 0.45%. In contrast, the probability that it will exceed 32 units is around 99.55%.

The general interpretation of these results is that the likelihood of the total claim amount being less than or equal to 32 units is very low, with probabilities ranging from approximately 0.45% to 0.55%. Consequently, the probability that the total claim amount will exceed 32 units is extremely high, ranging from 99.45% to 99.55% for the `AutoCollision` dataset. These findings indicate that, across all recursion algorithms evaluated, it is almost certain that total claims will surpass 32 units, underscoring the high-risk nature of the claims being modeled.

These results provide valuable insights for effective risk management and premium setting in the insurance sector. The high probability of large aggregate claims indicates that insurers must prepare for substantial payouts. Understanding this risk landscape allows insurers to more accurately assess claim distributions and frequencies, leading to more informed pricing strategies that ensure financial sustainability. Insurers can utilize these insights to allocate adequate reserves for high-expectation claims, thereby reducing the risk of insolvency. Moreover, policy designs can incorporate deductibles, limits, and exclusions that align with the high likelihood of large claims, striking a balance between customer affordability and insurer profitability. These findings also support the development of targeted reinsurance strategies, allowing insurers to transfer a portion of high-risk exposures and minimize the financial impact of large claims.

The observed differences in computational run times and aggregate claim sums across algorithms are attributable to the inherent complexity and structural differences of the recursion methods. The truncated Schröter algorithm, with its three-parameter structure, strikes an efficient balance between model complexity and computational speed, yielding both fast run times and higher aggregate claims. The Panjer recursion algorithm, while simpler with only two parameters, offers efficient computation but may not capture as many underlying data features. In contrast, the Schröter recursion algorithm, which incorporates an additional convolution term, requires more computation time but provides a nuanced perspective on aggregate claim modeling.

3.1.5 Simulation Study

Here, we generate random claim amounts data from the Negative binomial distribution by setting $r = 100$ and $p = 0.05$. We varied the sample size to examine the computational efficiency and run time of the truncated Schröter, Panjer, and Schröter recursion algorithms for the aggregate claim. Initially, we generated 5000 random numbers from the Negative binomial distribution and fit equation (3.3) to the data to obtain the estimate of the parameters $\hat{a}, \hat{b}$, and $\hat{c}$ as $\hat{a} = 0.9905$, $\hat{b} = 18.3096$, $\hat{c} = -1.2840$, and compute $g(0) = 0.00117$ to implement the algorithms. Tables 6, 7, and 8 present the sample sizes, aggregate claim amounts, and the execution time in seconds for each

Table 3.5: Aggregate claim for the truncated Schröter, Panjer, and Schröter algorithms

s	$g(s)$ Truncated Schröter	$g(s)$ Panjer	$g(s)$ Schröter
1	711000×10^{-3}	711000×10^{-3}	711000×10^{-3}
2	13488×10^{-5}	11953×10^{-5}	11958×10^{-5}
3	78790×10^{-6}	69718×10^{-6}	69763×10^{-6}
4	18245×10^{-6}	16054×10^{-6}	16132×10^{-6}
5	21313×10^{-5}	18881×10^{-5}	18908×10^{-5}
6	58042×10^{-5}	51406×10^{-5}	51450×10^{-5}
7	32214×10^{-5}	28454×10^{-5}	28493×10^{-5}
8	15979×10^{-5}	14070×10^{-5}	14113×10^{-5}
9	48176×10^{-5}	42617×10^{-5}	42727×10^{-5}
10	11819×10^{-4}	10454×10^{-4}	10474×10^{-4}
11	11254×10^{-4}	99323×10^{-5}	99514×10^{-5}
12	49170×10^{-5}	43135×10^{-5}	43306×10^{-5}
13	46463×10^{-5}	40787×10^{-5}	41095×10^{-5}
14	15944×10^{-4}	14065×10^{-4}	14119×10^{-4}
15	13782×10^{-4}	12090×10^{-4}	12149×10^{-4}
16	74915×10^{-5}	65005×10^{-5}	65502×10^{-5}
17	65619×10^{-5}	57003×10^{-5}	57681×10^{-5}
18	18092×10^{-4}	15882×10^{-4}	15999×10^{-4}
19	16315×10^{-4}	14191×10^{-4}	14324×10^{-4}
20	95516×10^{-5}	81572×10^{-5}	82706×10^{-5}
21	11549×10^{-4}	99748×10^{-5}	10108×10^{-4}
22	36393×10^{-4}	31961×10^{-4}	32173×10^{-4}
23	30416×10^{-4}	26473×10^{-4}	26719×10^{-4}
24	17428×10^{-4}	14879×10^{-4}	15094×10^{-4}
25	15023×10^{-4}	12823×10^{-4}	13062×10^{-4}
26	35570×10^{-4}	30989×10^{-4}	31358×10^{-4}
27	27630×10^{-4}	23653×10^{-4}	24074×10^{-4}
28	17519×10^{-4}	14558×10^{-4}	14920×10^{-4}
29	18952×10^{-4}	15950×10^{-4}	16353×10^{-4}
30	30447×10^{-4}	26076×10^{-4}	26676×10^{-4}
31	27511×10^{-4}	23012×10^{-4}	23706×10^{-4}
32	22584×10^{-4}	18434×10^{-4}	19048×10^{-4}
Sum of probabilities	**0.005483**	**0.004887**	**0.004553**
Execution time (s)	**0.051093**	**0.060898**	**0.173438**

recursion algorithm.

Figure 3.4 illustrates the execution times of the truncated Schröter recursion, Panjer recursion, and Schröter recursion algorithms for varying values of n, highlighting significant differences in computational efficiency as the sample size increases. The truncated Schröter recursion algorithm consistently demonstrates the lowest execution times across all values of n, starting at 0.0006919 s for $n = 20$ and increasing to 1.5046701 s for $n = 5000$ (see Figure 3.4).

Table 3.6: Efficiency of the truncated Schröter algorithm on simulated claim data

Recursion Algorithm	**Sample (n)**	**sum of $g(s)$**	**Execution time (s)**
The truncated Schröter algorithm	20	2.7080348	0.0006919
	50	1.5211150	0.0036724
	100	1.1730166	0.0188396
	150	1.1747598	0.0335643
	200	1.1296751	0.0586591
	300	0.9208012	0.1158113
	600	0.8504156	0.3374069
	1500	0.6127028	0.8211629
	2000	0.5397614	1.0198436
	3000	0.4756758	1.2040498
	4000	0.4180231	1.4369745
	5000	0.3979199	1.5046701

Table 3.7: Efficiency of the Panjer algorithm on simulated claim data

Recursion Algorithm	**Sample (n)**	**sum of $g(s)$**	**Execution time (s)**
The Panjer algorithm	20	0.0075441	0.0014127
	50	0.0074026	0.0047266
	100	0.0074002	0.0212255
	150	0.0074619	0.0378468
	200	0.0074672	0.0692441
	300	0.0072761	0.1391730
	600	0.0073219	0.4021811
	1500	0.0073130	1.0212069
	2000	0.0072111	1.0606146
	3000	0.0072833	1.3638492
	4000	0.0072209	1.6406126
	5000	0.0072351	1.7650454

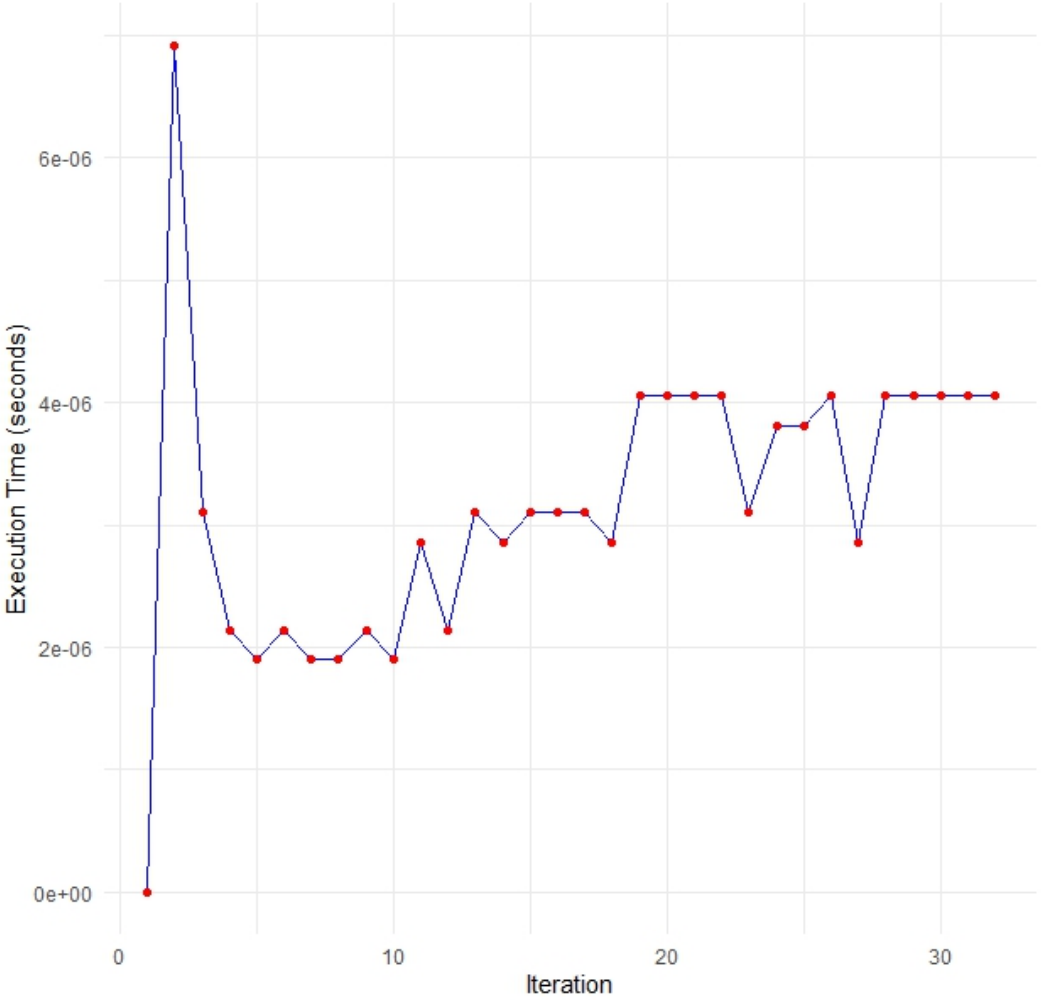


Figure 3.1: The Execution Time Plot of the Truncated Schröter Recursive Algorithm for Each Iteration.

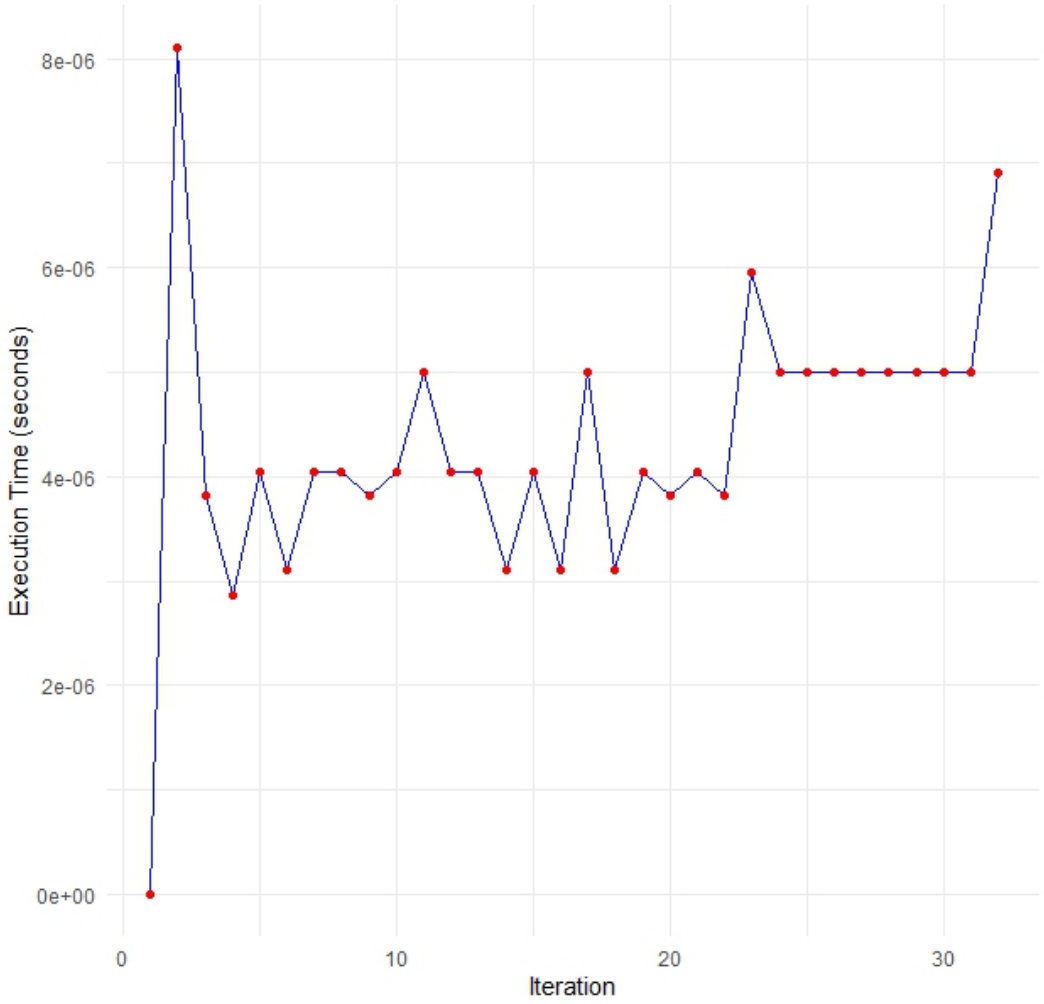


Figure 3.2: The Execution Time Plot of the Panjer Recursive Algorithm for Each Iteration.

Table 3.8: Efficiency of the Schröter algorithm on simulated claim data

Recursion Algorithm	**Sample (n)**	**sum of $g(s)$**	**Execution time (s)**
The Schröter algorithm	20	0.0062785	0.0028598
	50	0.0063757	0.0195415
	100	0.0064331	0.0765483
	150	0.0064909	0.1735694
	200	0.0065073	0.2585254
	300	0.0063946	0.6078202
	600	0.0064442	1.8534706
	1500	0.0064868	4.6577935
	2000	0.0064729	5.4729755
	3000	0.0065035	6.8906786
	4000	0.0064749	8.0277340
	5000	0.0064949	8.7507973

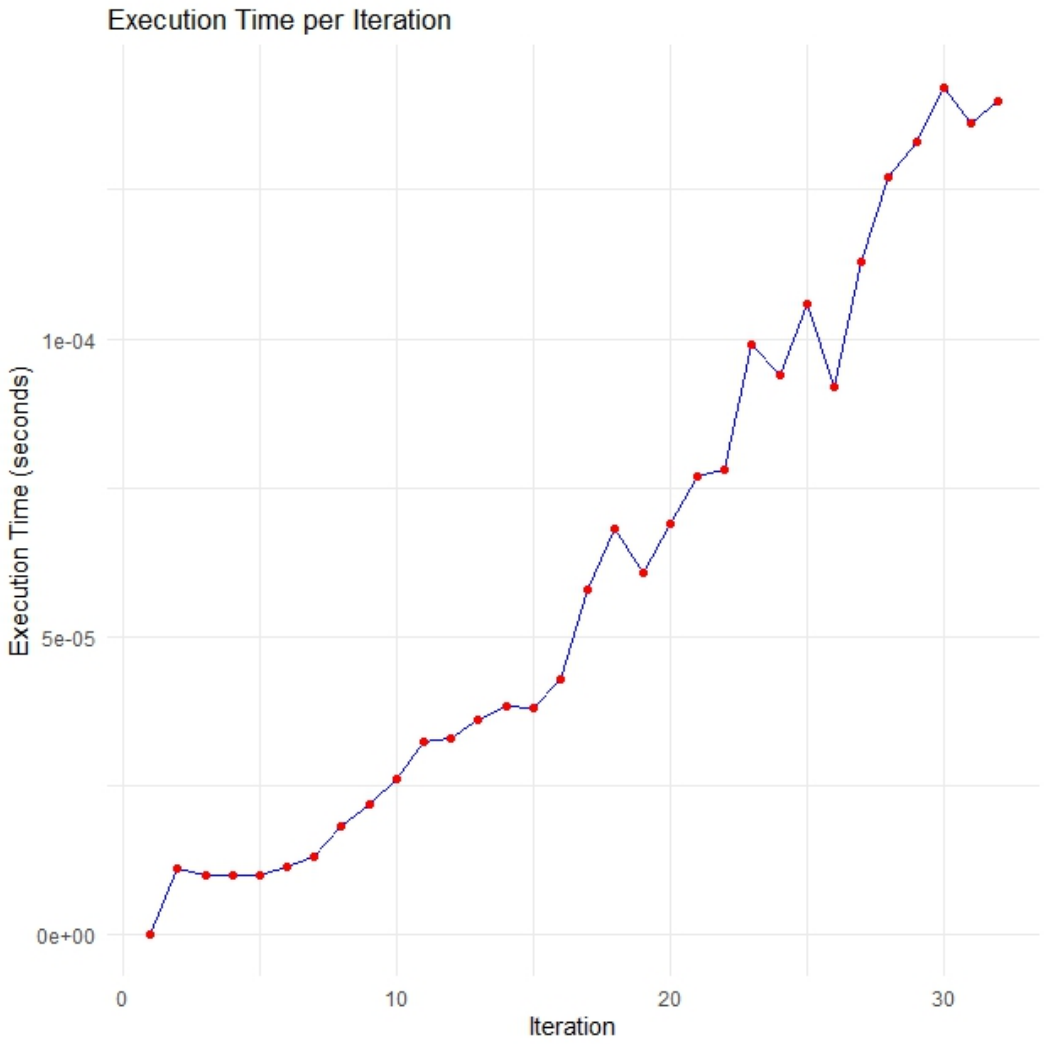


Figure 3.3: The Execution Time Plot of the Schröter Recursive Algorithm for Each Iteration.

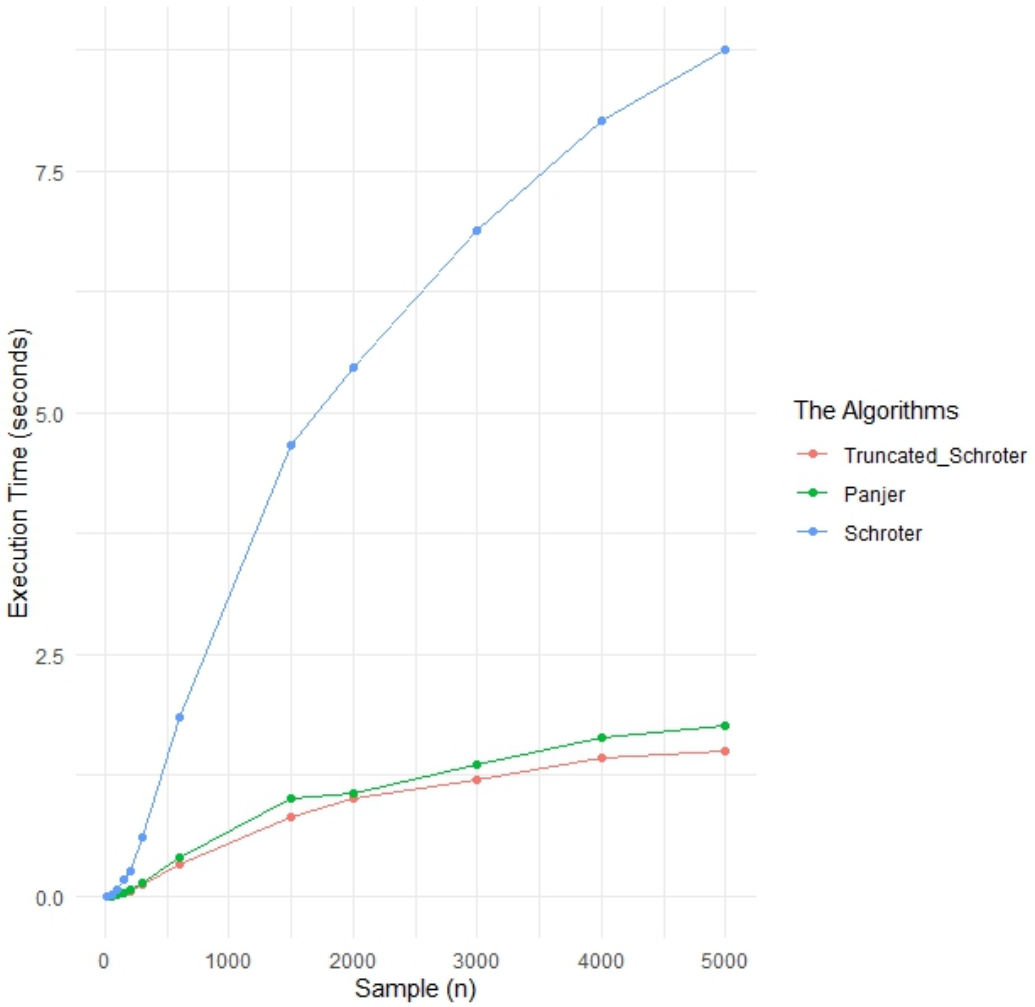


Figure 3.4: Visual Representation of the Execution Time of the Truncated Schröter, Panjer, and Schröter recursive Algorithms.

Discussion

The performance, presented in Tables 3.5–3.8 and Figures 3.1–3.4, demonstrates that the proposed algorithm is both highly efficient and scalable, capable of handling larger datasets with minimal computational burden. Among the benchmarked methods, the Panjer recursive algorithm exhibits increasing execution times as n grows, beginning at 0.0014127 s for $n = 20$ and rising to 1.7650454 s for $n = 5000$. Although reasonably efficient, it shows lower scalability compared to the truncated Schröter recursion algorithm (see Figure 3.4).

In contrast, the standard Schröter recursive algorithm, which incorporates a convolution component f_i^{n*}, displays substantially higher execution times, starting at 0.0028598 s for $n = 20$ and escalating sharply to 8.7507973 s for $n = 5000$ (see Figure 3.4). This steep increase reflects poor scalability and reduced efficiency, particularly for larger sample sizes, making it the least optimal option among the three algorithms.

Thus, in general, the truncated Schröter recursion algorithm emerges as the most efficient and scalable approach, consistently outperforming both the Panjer and standard Schröter algorithms. While the latter may offer greater modeling flexibility, its computational complexity limits practical applicability. In contrast, the truncated version achieves superior balance by combining computational speed with reliable accuracy.

The consistency of execution times was also assessed by repeating each sample size across five independent runs. The negligible variation observed indicates that the algorithms are stable and reproducible, though minor fluctuations may still arise from the operational state of the computing system.

More broadly, this section investigated the computation of aggregate claim amounts using multiple recursive algorithms, with particular emphasis on the newly introduced truncated Schröter recursion. The central objective was to enhance both the accuracy and computational efficiency of aggregate claim estimation, a cornerstone of effective risk management and premium setting in the insurance industry.

When applied to the `AutoCollision` dataset, the truncated Schröter algorithm consistently delivered the fastest execution times and the highest aggregate claim sums, reflecting both computational efficiency and modeling comprehensiveness. For modeling claim count data, the Negative Binomial distribution was preferred over the Generalized Poisson distribution, owing to its capacity to accommodate overdispersion, as confirmed by AIC and BIC selection criteria.

Simulation studies further validated the advantages of the truncated Schröter algorithm across varying sample sizes. It consistently outperformed the Panjer and standard Schröter algorithms in terms of execution time and scalability, while effectively capturing data variability for refined analysis. These findings highlight the importance of selecting suitable counting distributions and recursive methods when modeling aggregate claims.

In conclusion, the truncated Schröter recursion algorithm proves to be a robust and reliable tool for estimating aggregate claim amounts, striking an effective balance between computational speed and modeling accuracy. Its adoption can substantially improve risk assessment and premium pricing strategies, ultimately benefiting both insurers and policyholders. Future research may focus on enhancing the algorithm further, for example by incorporating machine learning techniques to optimize parameter estimation based on evolving claim patterns, or by applying it to other insurance domains beyond automobile claims to assess its generalizability and identify domain-specific refinements.

3.1.6 Generalized Schröter Recursive Formula

This section presents a new result on a more general class of the Schröter recursive formula and shows how it determines counting distributions.

We define the generalized Schröter recursive formula as

$$P_n = \sum_{i=1}^{t} \left[\left(a_i + \frac{b_i}{n} \right) P_{n-i} + \frac{c_i}{n} P_{n-i-1} \right], \qquad n = 1, 2, 3, \ldots, \tag{3.13}$$

for some t, where t is an integer. Furthermore, by definition, $P_{-1} = 0$ and $P_0 > 0$. If $c_i = 0$, we obtain the generalized class of the Panjer formula defined in Sundt [113], and for $t = i = 1$, we have the classical Schröter recursive formula as a special case. We will refer to this class as the $R_t(a, b, c)$ class of Schröter's family. The generalized class has $3t$ parameters.

Let $G(s)$ be a probability generating function (PGF) defined as

$$G(s) = \sum_{n=0}^{\infty} P_n s^n, \tag{3.14}$$

where P_n denotes the PMF of the $R_t(a, b, c)$ class. Differentiating,

$$G^{'}(s) = \sum_{n=1}^{\infty} n P_n s^{n-1}. \tag{3.15}$$

Substituting the recurrence relation into $G'(s)$ gives

$$G^{'}(s) = \sum_{n=1}^{\infty} n s^{n-1} \left[\sum_{i=1}^{t} \left(\left(a_i + \frac{b_i}{n} \right) P_{n-i} + \frac{c_i}{n} P_{n-i-1} \right) \right]. \tag{3.16}$$

Rearranging summations,

$$G^{'}(s) = \sum_{i=1}^{t} \left[\sum_{n=i}^{\infty} ((n a_i + b_i) P_{n-i} + c_i P_{n-i-1}) \right] s^{n-1}. \tag{3.17}$$

Changing indices of summation yields

$$G^{'}(s) = \sum_{i=1}^{t} \left[a_i \sum_{n=0}^{\infty} (n + i) s^{n+i-1} P_n + b_i \sum_{n=0}^{\infty} s^{n+i-1} P_n + c_i \sum_{n=0}^{\infty} s^{n+i-1} P_{n-1} \right]. \tag{3.18}$$

This becomes

$$G^{'}(s) = \sum_{i=1}^{t} \left[a_i s^i \sum_{n=0}^{\infty} n s^{n-1} P_n + (i a_i + b_i) s^{i-1} \sum_{n=0}^{\infty} s^n P_n + c_i s^i \sum_{n=0}^{\infty} s^{n-1} P_{n-1} \right]. \tag{3.19}$$

Recognizing the series as $G(s)$ and $G'(s)$, we obtain

$$G^{'}(s) = \sum_{i=1}^{t} \left[a_i s^i G^{'}(s) + (i a_i + b_i) s^{i-1} G(s) + c_i s^i G(s) \right]. \tag{3.20}$$

Hence,

$$\frac{G'(s)}{G(s)} = \frac{\sum_{i=1}^{t}(ia_i + b_i + sc_i)s^{i-1}}{1 - \sum_{i=1}^{t} a_i s^i} \tag{3.21}$$

Equation (3.21) determines the counting distributions of the $R_t(a, b, c)$ class. In particular, if a distribution with PMF $P(N = n)$ satisfies (3.21), then it belongs to the $R_t(a, b, c)$ family, and the parameters (a_i, b_i, c_i) can be uniquely determined.

Theorem 6 (Characterization of the $R_t(a, b, c)$ Class)**.** *Let $(P_n)_{n\geq 0}$ be a PMF with $P_0 > 0$ and PGF*

$$G(s) = \sum_{n=0}^{\infty} P_n s^n,$$

analytic in a neighborhood of $s = 0$. Fix an integer $t \geq 1$. Then the following are equivalent:

1. *The sequence (P_n) satisfies the generalized Schröter recurrence*

$$P_n = \sum_{i=1}^{t} \left[\left(a_i + \frac{b_i}{n} \right) P_{n-i} + \frac{c_i}{n} P_{n-i-1} \right], \qquad n \geq 1,$$

   with $P_{-1} = 0$, for some real parameters a_i, b_i, c_i $(i = 1, \ldots, t)$.

2. *The PGF $G(s)$ satisfies the differential identity*

$$\frac{G'(s)}{G(s)} = \frac{\sum_{i=1}^{t}(ia_i + b_i + sc_i)\, s^{\,i-1}}{1 - \sum_{i=1}^{t} a_i s^{\,i}}.$$

Moreover, the parameters a_i, b_i, c_i are uniquely determined by the rational function on the right-hand side of (2)*.*

Proof. We prove both directions.

(i) ⇒ (ii). This has already been derived in the steps leading to equation (3.21). Assuming the recurrence holds, substitution into the PGF derivative $G'(s)$ yields (2).

(ii) ⇒ (i). Assume the PGF identity (2) holds. Multiplying through gives

$$\left(1 - \sum_{i=1}^{t} a_i s^i\right) G'(s) = \left(\sum_{i=1}^{t} (ia_i + b_i) s^{i-1} + \sum_{i=1}^{t} c_i s^i \right) G(s).$$

Expanding both sides in power series, the coefficient of s^{n-1} on the left-hand side is

$$nP_n - \sum_{i=1}^{t} a_i (n - i) P_{n-i}, \qquad n \geq 1.$$

The coefficient on the right-hand side is

$$\sum_{i=1}^{t} (ia_i + b_i) P_{n-i} + \sum_{i=1}^{t} c_i P_{n-i-1}.$$

Equating coefficients yields

$$nP_n = \sum_{i=1}^{t}(na_i + b_i)P_{n-i} + \sum_{i=1}^{t} c_i P_{n-i-1}.$$

Dividing by n gives

$$P_n = \sum_{i=1}^{t}\left[\left(a_i + \frac{b_i}{n}\right)P_{n-i} + \frac{c_i}{n}P_{n-i-1}\right], \qquad n \geq 1,$$

which is exactly the recurrence. Thus, (i) holds.

Uniqueness. The rational form in (2) determines the a_i from the denominator, and then b_i, c_i from the numerator. Thus, the parameters are uniquely determined. □

3.1.7 Properties of the Members of the Schröter $R_t(a, b, c)$ Class of Distributions

This section examines the properties of the members of the more general class of the Schröter formula defined in equation (3.21).

Let P_n be the PMF of the random discrete variable N in $R_t(a, b, c)$. From equation (3.21), we can write that $\mu = E(N) = G^{'}(1)$. Hence,

$$\mu = \frac{\sum_{i=1}^{t}(ia_i + b_i + c_i)}{1 - \sum_{i=1}^{t} a_i}. \tag{3.22}$$

Using the fact that $G'(s) = \sum_{n=0}^{\infty} ns^{n-1}P_n$ and $G"(1) = \sum_{n=0}^{\infty} n(n-1)P_n = E(n(n-1)) = E(n^2) - E(n)$, we have that $\sigma = Var(N) = G"(1) + G'(1) - [G'(1)]^2$. By applying the quotient rule of differentiation to equation (3.21), we have

$$\sigma = \frac{[1 - \sum_{i=1}^{t} a_i][\sum_{i=1}^{t}(ia_i + b_i + c_i)(i-1) + \sum_{i=1}^{t}(ia_i + b_i + sc_i)c_i] + D}{[1 - \sum_{i=1}^{t} a_i]^2} + E \tag{3.23}$$

where $D = \sum_{i=1}^{t}(ia_i + b_i + c_i)\sum_{i=1}^{t} ia_i$ and $E = \frac{[1-\sum_{i=1}^{t} a_i][\sum_{i=1}^{t}(ia_i+b_i+c_i)] - [\sum_{i=1}^{t}(ia_i+b_i+c_i)]^2}{[1-\sum_{i=1}^{t} a_i]^2}$. These results generalized the proposition 2 of the Schröter [103] results (see Schröter [103]; page 167).

3.2 Final Remarks

This chapter establishes a rigorous and computationally efficient framework for estimating aggregate claims under truncation at one, directly motivated by real policy structures (deductibles and unreported small losses). From the truncated Schröter recurrence, we derive the PGF, the associated PMF, and a closed-form log-likelihood, which together enable constrained maximum likelihood estimation of (a, b, c) with numerically stable implementation. We then introduce a truncated Schröter recursive algorithm that eliminates convolution, resulting in a simple weighted update for $g(s)$ that retains Panjer and the classical Schröter recursive formulas as special cases. The implementation workflow is end-to-end. It starts by initialization via $g(0)$ from the

chosen count model (i.e., Negative Binomial), estimates (a, b, c) by MLE under box constraints, and computes the aggregate distribution by the new recursive algorithm.

Empirically, the AutoCollision application shows that the truncated Schröter method achieves the fastest runtimes and yields the largest aggregate-probability mass over the tail range examined, while model selection for claim counts favors the Negative Binomial over the Generalized Poisson by AIC/BIC and residual diagnostics. A simulation study confirms scalability: execution times for the truncated Schröter recursive algorithm grow slowly with sample size and remain substantially lower than those of Panjer algorithm and the classical Schröter algorithm, which relies on convolution. Finally, we generalize the Schröter recursive formula to the $R_t(a, b, c)$ family, characterize its counting distributions via PGF differential identity, and provide closed-form expressions for the mean and variance. Collectively, these results provide a robust, truncation-aware approach that enhances modeling accuracy and operational speed for aggregate claim estimation.

CHAPTER 4

RESULTS ON INFINITE-SUM REPRESENTATION OF AGGREGATE APPROVED CLAIM AMOUNT DISTRIBUTION: ZERO-TRUNCATED POISSON AND SHIFTED EXPONENTIAL DISTRIBUTIONS

In this chapter, we study the distribution of aggregate approved claim amounts when the claim frequency follows a zero-truncated Poisson (ZTP) distribution and the claim severity follows a shifted exponential (SE) distribution. This combination captures two essential features of insurance contracts: the exclusion of zero frequencies, since only positive numbers of approved claims are meaningful, and the presence of deductibles, which shift claim severities away from zero. The resulting compound ZTP-SE model reflects practical realities of insurance data and provides a more realistic basis for actuarial modeling of aggregate approved claims. A central contribution of this chapter is the derivation of an infinite-sum representation for the aggregate distribution, which facilitates theoretical analysis and efficient numerical evaluation. The study also shows that the derived approved claim distribution is a member of the Panjer and Schröter families. To complement this, the chapter employs saddlepoint approximation techniques to obtain accurate and computationally efficient approximations for the density and distribution functions of the aggregate model. The results presented in this chapter is based on my joint work with Agu and Wald [3]

4.1 Motivation

The motivation for this chapter arises from the limitations of traditional compound models in reflecting the operational structure of insurance data. Classical formulations often assume that severities can take values arbitrarily close to zero, yet in practice, deductibles ensure that insurer payouts begin only above a positive threshold. Standard exponential severities, therefore, misrepresent observed claims, whereas the shifted exponential distribution provides a natural correction by incorporating the deductible directly into the model. Likewise, frequency distributions such as the Poisson allocate positive mass at zero, which is inconsistent with datasets of approved claims that nec-

essarily involve at least one claim. The zero-truncated Poisson distribution eliminates this inconsistency while retaining the tractability of the Poisson law.

Although these modeling adjustments improve realism, the convolution of ZTP frequencies with SE severities does not yield closed-form aggregate distributions. Direct numerical evaluation becomes computationally expensive as the number of claims increases. Previous studies, including recursive approaches examined in Chapter **3**, have demonstrated the value of recursive methods for compound models; however, such approaches are not always optimal in the presence of truncation and shifting. To address this, this study derives an infinite-sum representation for the ZTP–SE compound distribution, which provides a precise theoretical foundation for further analysis.

Beyond exact representations, efficient approximations are necessary for practical actuarial applications. Saddlepoint approximation methods, introduced by Daniels [27] and further developed by Reid [100] and Wang [125], have been widely recognized for their accuracy in approximating the distributions of sums of random variables. Their relevance to actuarial and reliability contexts has been highlighted in recent studies Alhejaili and AlGhamedi [8]; Tang and Reid [119]; Meng et al. [76]; Butler [21]. By combining an exact infinite-sum representation with saddlepoint approximations, this chapter develops a framework that is theoretically rigorous and computationally feasible for evaluating the distribution of aggregate approved claim amounts.

The main aim of this chapter is to construct and analyze the compound ZTP-SE model, derive its infinite-sum representation, and integrate saddlepoint methods to provide efficient and accurate approximations. In doing so, the chapter advances the recursive framework established in Chapter **3** toward a more practically motivated actuarial modeling setting that aligns with deductible-inclusive and truncation-adjusted insurance data.

4.1.1 Model Setup

This section establishes the probabilistic framework for modeling aggregate approved claim amounts. We specify the claim frequency distribution using the ZTP law to ensure that only positive claim counts are considered, and we model claim severities with the SE distribution to capture the effect of deductibles. The aggregate approved claim amount is then formulated as a compound sum of SE severities, with the counting variable derived from the truncated claim frequency. This setup provides the foundation for deriving exact distributional properties, moment-based characterizations, and saddlepoint approximations in subsequent sections.

Let N denote the total number of claims made by policyholders within a fixed period. We assume that N follows the ZTP distribution with parameter $\lambda > 0$, so that $N \in \{1, 2, 3, \dots\}$. The PMF of N is defined as:

$$P(N = n) = \frac{\lambda^n e^{-\lambda}}{n!(1 - e^{-\lambda})}, \quad n = 1, 2, 3, \dots \tag{4.1}$$

and its corresponding MGF is expressed as:

$$M_N(t) = \frac{e^{\lambda(e^t - 1)} - 1}{1 - e^{-\lambda}} \tag{4.2}$$

Out of the N total claims, let L denote the number of approved claims in a fixed time, where $L \leq N$. We assume that $L \geq 1$ almost surely.

Consequently, the number of non-approved claims is $N - L$. The random variable L may depend on a probabilistic approval process, which is specified separately. The distribution of $N - L$ can thus be derived from the joint or conditional distribution of L given N.

Let $T_1, T_2, \ldots, T_L$ denote the claim amounts paid out by the insurance company for the L approved claims.

We assume that $T_j \sim \text{SE}(\theta, \beta)$ for $j = 1, 2, \ldots, L$, all T_j are independent and identically distributed (IID), and the severities T_j are stochastically independent of N.

The SE distribution incorporates a deductible threshold parameter $\theta > 0$ and scale parameter $\beta > 0$. Its probability density function (PDF) is expressed as:

$$f_T(t) = \beta e^{-\beta(t-\theta)}, \quad t > \theta \tag{4.3}$$

The aggregate amount paid out by the insurance company in the fixed time period is defined as:

$$X = \sum_{j=1}^{L} T_j = T_1 + T_2 + \cdots + T_L \tag{4.4}$$

which reflects the number of approved claims via L and the amount of each claim through T_j.

Thus, X represents the total approved payout, which depends on the number of approved claims (L) and the amounts paid for each of those claims ($T_1, \ldots, T_L$).

The random variable X is a compound sum where the counting variable L is derived from the zero-truncated claim count N, and the summands follow the SE distribution. This formulation allows for the realistic modeling of deductible thresholds through the SE distribution, the exclusion of zero-claim cases via the ZTP assumption on N, and the derivation of the distribution of approved claims L and the unapproved claim count $N - L$.

In the following sections, we introduce the distribution of approved claims, along with its corresponding mean, variance, moment generating function (MGF), and cumulant generating function (CGF), as well as the distribution of unapproved claims. We then develop the PMF and CGF of X and its related properties, and derive the saddlepoint approximation to its density and cumulative distribution functions.

4.1.2 The distribution of approved claims

This section presents the probabilistic framework for modeling the number of approved claims out of the total reported claims. Since not every reported claim results in payment, only those that exceed the deductible and meet policy conditions are approved.

Therefore, it is necessary to derive the distribution of approved claims separately. We model the approval process as a thinning mechanism applied to the ZTP claim frequency, which naturally leads to a compound structure. This formulation enables us to determine the distribution of approved and unapproved claims, along with their corresponding generating functions and moments.

In insurance practice, not all reported claims result in payment; only those that exceed a specified deductible and meet the contract conditions are approved for compensation. Consequently, it is essential to model the distribution of approved claims separately from the total reported claims.

Let each reported claim have an independent probability $q = \mathbb{P}(T > \theta)$ of being approved, with $T \sim \text{SE}(\theta, \beta)$.

Conditional distribution. Given that $N = n \sim \text{ZTP}(\lambda)$, the approved claim count $L \mid N = n \sim \text{Bin}(n, q)$ with the PMF defined in equation (4.5), and the marginal distribution of L becomes a compound distribution that integrates the effects of claim frequency and approval mechanism, and $L \in \{1, 2, 3, \ldots\}$, where we are only interested in positive approvals. So, the PMF of $L \mid N = n$ is:

$$P(L = l \mid N = n) = \binom{n}{l} q^l (1-q)^{n-l}, \quad l = 0, 1, 2, \ldots, n, \;\; q \in [0, 1]. \tag{4.5}$$

This forms the conditional distribution of L.

Mixture over the ZTP frequency (unconditional given $N \geq 1$). The marginal distribution of L (still allowing $L = 0$) is derived by marginalizing over N:

$$\begin{aligned} P(L = l) &= \sum_{n=l}^{\infty} P(L = l \mid N = n) P(N = n) \\ &= \sum_{n=l}^{\infty} \binom{n}{l} q^l (1-q)^{n-l} \frac{e^{-\lambda} \lambda^n}{n! \, (1 - e^{-\lambda})} \\ &= \frac{e^{-\lambda}}{1 - e^{-\lambda}} \frac{(q\lambda)^l}{l!} \sum_{n=l}^{\infty} \frac{\big((1-q)\lambda\big)^{n-l}}{(n-l)!}. \end{aligned}$$

Letting $m = n - l$ gives

$$P(L = l) = \frac{e^{-\lambda}}{1 - e^{-\lambda}} \frac{(q\lambda)^l}{l!} \sum_{m=0}^{\infty} \frac{\big((1-q)\lambda\big)^m}{m!} = \frac{e^{-\lambda}}{1 - e^{-\lambda}} \frac{(q\lambda)^l}{l!} e^{(1-q)\lambda} = \frac{e^{-q\lambda} (q\lambda)^l}{l! \, (1 - e^{-\lambda})},$$

valid for all $l = 0, 1, 2, \ldots$. In particular, the positive-l probabilities are

$$P(L = l) = \frac{e^{-q\lambda} (q\lambda)^l}{l! \, (1 - e^{-\lambda})}, \qquad l = 1, 2, \ldots, \tag{4.6}$$

and there remains nonzero mass at zero approvals unless $q = 1$:

$$P(L = 0) = \frac{e^{-q\lambda} - e^{-\lambda}}{1 - e^{-\lambda}}.$$

Restriction to positive approvals ($L \geq 1$). Because we are only considering episodes with at least one approved claim, we condition on $L \geq 1$. From equation (4.6),

$$P(L \geq 1) = \sum_{l=1}^{\infty} P(L = l) = \frac{1 - e^{-q\lambda}}{1 - e^{-\lambda}}.$$

Therefore, for $l = 1, 2, \ldots,$

$$P(L = l \mid L \geq 1) = \frac{e^{-q\lambda} (q\lambda)^l}{l! \left(1 - e^{-q\lambda}\right)} \tag{4.7}$$

which is the zero-truncated Poisson distribution with parameter $q\lambda$. We will refer to equation (4.7) as the zero-truncated Poisson-Binomial (ZTPB) distribution, induced by thinning a (zero-truncated) Poisson frequency with approval probability q.

Furthermore, it can be easily shown that the ZTPB belongs to the Panjer recursive family of discrete distributions with parameters $a = 0$ and $b = q\lambda$; indeed, for $l \geq 2$,

$$\frac{P(L = l \mid L \geq 1)}{P(L = l - 1 \mid L \geq 1)} = \frac{q\lambda}{l}.$$

(We note $P(L = 0 \mid L \geq 1) = 0$, so the standard Panjer recursion needs the usual adjustment for zero-truncation; the ratio above holds for $l \geq 2$.) Additionally, one can *show* that the convolution of equation (4.7) with the Panjer class belongs to the Schröter family, as studied in Chapter **3**.

The plot of the ZTPB for varying parameter values is shown in Figure 1. Similarly, the distribution of an unapproved claim is simply the distribution of $U = N - L$.

4.1.3 The Distribution of Unapproved Claims

The distribution of an unapproved claim is simply the distribution of $U = N - L$. Given that $N = n$, $U \mid N = n \sim \text{Bin}(n, 1 - q)$, where $1 - q$ is the probability of the claim not being approved.

Unconditional given $N \geq 1$. By the same mixture argument as above,

$$P(U = u) = \frac{e^{-(1-q)\lambda}\big((1-q)\lambda\big)^u}{u!\,(1 - e^{-\lambda})}, \qquad u = 0, 1, 2, \ldots.$$

If one also focuses on strictly positive unapproved counts ($U \geq 1$), then

$$P(U = u \mid U \geq 1) = \frac{e^{-(1-q)\lambda}\big((1-q)\lambda\big)^u}{u!\,\big(1 - e^{-(1-q)\lambda}\big)}, \qquad u = 1, 2, \ldots. \tag{4.8}$$

Equation (4.8) is the ZTPB distribution of an unapproved claim in this positive-u sense.

4.1.4 The Moments of the approved claim count

Since $N = n$, $L \mid N = n \sim \text{Bin(n, q)}$; so, by definition, the conditional moments are:

$$E(L \mid N = n) = nq, \qquad \text{Var}(L \mid N = n) = nq(1 - q).$$

Using the fact that

$$E(L) = E_N\big[E(L \mid N)\big],$$

and that $L \mid N = n \sim \text{Binomial}(n, q)$, we have

$$E(L \mid N = n) = nq.$$

Therefore,

$$E(L) = E_N\left[E(L \mid N)\right] = E_N\left[qN\right] = q\,E(N) = \frac{q\lambda}{1 - e^{-\lambda}},$$

since

$$E(N) = \frac{\lambda}{1 - e^{-\lambda}} \quad \text{for } N \sim \text{ZTP}(\lambda).$$

Furthermore, using the law of total variance,

$$\begin{aligned}\operatorname{Var}(L) &= \mathbb{E}\big[\operatorname{Var}(L \mid N)\big] + \operatorname{Var}\big(\mathbb{E}[L \mid N]\big) \\ &= \mathbb{E}\big[Nq(1-q)\big] + \operatorname{Var}(qN) \\ &= q(1-q)\,\mathbb{E}(N) + q^2\,\operatorname{Var}(N).\end{aligned}$$

For $N \sim \mathrm{ZTP}(\lambda)$ it holds that

$$\mathbb{E}(N) = \frac{\lambda}{1-e^{-\lambda}}, \qquad \operatorname{Var}(N) = \frac{\lambda}{1-e^{-\lambda}} - \frac{\lambda^2 e^{-\lambda}}{(1-e^{-\lambda})^2}.$$

Hence,

$$\operatorname{Var}(L) = q(1-q)\frac{\lambda}{1-e^{-\lambda}} + q^2\left(\frac{\lambda}{1-e^{-\lambda}} - \frac{\lambda^2 e^{-\lambda}}{(1-e^{-\lambda})^2}\right) = \frac{q\lambda}{1-e^{-\lambda}} - \frac{q^2\lambda^2 e^{-\lambda}}{(1-e^{-\lambda})^2}\,.$$

If, as in this chapter, we *restrict to positive approvals* ($L \geq 1$), then $L \mid L \geq 1 \sim \mathrm{ZTP}(q\lambda)$, and its mean and variance are

$$\mathbb{E}(L \mid L \geq 1) = \frac{q\lambda}{1-e^{-q\lambda}}, \qquad \operatorname{Var}(L \mid L \geq 1) = \frac{q\lambda}{1-e^{-q\lambda}} - \frac{(q\lambda)^2 e^{-q\lambda}}{(1-e^{-q\lambda})^2}\,.$$

4.1.5 The MGF of the approved claim count

Let $M_L(t) = \mathbb{E}(e^{tL})$. Using (4.5),

$$M_L(t) = \sum_{n=1}^{\infty} P(N=n)\,\mathbb{E}\big(e^{tL} \mid N=n\big) = \sum_{n=1}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!(1-e^{-\lambda})}\big(1-q+qe^t\big)^n.$$

By the exponential series,

$$M_L(t) = \frac{e^{-\lambda}}{1-e^{-\lambda}}\Big(e^{\lambda(1-q+qe^t)} - 1\Big) = \frac{e^{\lambda q(e^t-1)} - e^{-\lambda}}{1-e^{-\lambda}}\,,$$

which satisfies $M_L(0) = 1$.

If we condition on positive approvals ($L \geq 1$), then

$$M_{L|L\geq 1}(t) = \frac{e^{q\lambda(e^t-1)} - e^{-q\lambda}}{1-e^{-q\lambda}}\,.$$

4.1.6 The PGF of the approved claim count

Let $G_L(s) = \mathbb{E}(s^L)$. As above,

$$G_L(s) = \sum_{n=1}^{\infty}(1-q+qs)^n \frac{\lambda^n e^{-\lambda}}{n!(1-e^{-\lambda})} = \frac{e^{-\lambda}}{1-e^{-\lambda}}\Big(e^{\lambda(1-q+qs)} - 1\Big) = \frac{e^{\lambda q(s-1)} - e^{-\lambda}}{1-e^{-\lambda}}\,.$$

Under $L \geq 1$,

$$G_{L|L\geq 1}(s) = \frac{e^{q\lambda(s-1)} - e^{-q\lambda}}{1-e^{-q\lambda}}\,.$$

4.1.7 Approval probability from the SE

Let the payment approvals depend on the claim amounts exceeding the deductible threshold parameter θ, and $T \sim \text{SE}(\theta, \beta)$. Writing the deductible threshold as $\theta_0 > \theta$, using equation (4.3), we evaluate:

$$q = P(T > \theta_0) = \int_{\theta_0}^{\infty} f_T(t)\, dt = e^{-\beta(\theta_0 - \theta)}, \quad \theta_0 > \theta.$$

For $\theta_0 = \theta$, we obtain $q = 1$, which implies that all claims are approved. Hence, $q = P(T_j > \theta_0)$ is the probability of approval.

This section introduces the distribution of approved claims, showing that it follows the ZTPB distribution. The section derived the PMF of ZTPB, mean, variance, and generating functions, and also described the distribution of unapproved claims. This foundation is significant for the next subsection, where we extend the analysis to the aggregate amount of approved claims and derive its distributional properties using CGFs and saddlepoint approximations.

4.1.8 The aggregate approved claim amounts distribution

Having established the distribution of the number of approved claims in the previous subsection, we now turn to the distribution of the corresponding aggregate approved claim amounts. This section develops the probabilistic formulation of the total amount paid, considering only approved claims. By conditioning on the number of approvals and using the SE distribution for claim severities, we derive the density of the aggregate approved claims through compounding. This characterization not only highlights the role of the approval mechanism but also provides the analytical foundation for understanding how deductibles, approval probabilities, and claim arrival rates jointly shape the aggregate distribution.

The distribution of aggregate approved claim amounts

Let X denote the aggregate approved claim amount. Conditional on $L = l$ approved claims, we write

$$X \mid L = l = \sum_{j=1}^{l} T_j,$$

where each $T_j \sim \text{SE}(\theta, \beta)$. Recall that the shifted exponential distribution can be expressed as

$$T_j = \theta + Y_j, \quad Y_j \sim \text{Exp}(\beta),$$

where $\theta > 0$ is the deductible threshold and $\beta > 0$ is the *rate* parameter.

Substituting, we have

$$X \mid L = l = \sum_{j=1}^{l} T_j = \sum_{j=1}^{l} (\theta + Y_j) = l\theta + \sum_{j=1}^{l} Y_j.$$

Let

$$S_l = \sum_{j=1}^{l} Y_j.$$

Since the Y_j's are independent and exponentially distributed, S_l follows a gamma distribution:

$$S_l \sim \text{Gamma}(l, \beta),$$

with shape parameter l and rate parameter β. Therefore, the conditional distribution of $X \mid L = l$ is a shifted gamma:

$$X \mid L = l \sim l\theta + \text{Gamma}(l, \beta).$$

4.1.9 The unconditional PDF of the aggregate approved claim amounts

We start with the conditional distribution of X given $L = l$:

$$X \mid L = l \sim l\theta + \text{Gamma}(l, \beta),$$

so its conditional PDF is

$$f_{X|L=l}(x) = \frac{\beta^l (x - l\theta)^{l-1} e^{-\beta(x-l\theta)}}{\Gamma(l)} \mathbf{1}\{x \geq l\theta\}, \qquad l = 1, 2, \ldots \tag{4.9}$$

Because we are only considering *positive approvals* ($L \geq 1$), the approved-count PMF is the zero-truncated Poisson with parameter $q\lambda$:

$$P(L = l) = \frac{(\lambda q)^l e^{-q\lambda}}{l! \left(1 - e^{-q\lambda}\right)}, \qquad l = 1, 2, 3, \ldots$$

The unconditional PDF of X is then obtained by mixing over L:

$$\begin{aligned} f_X(x) &= \sum_{l=1}^{\infty} P(L = l)\, f_{X|L=l}(x) \\ &= \sum_{l=1}^{\infty} \left[\frac{\beta^l (x - l\theta)^{l-1} e^{-\beta(x-l\theta)}}{\Gamma(l)} \cdot \frac{(\lambda q)^l e^{-q\lambda}}{l! \left(1 - e^{-q\lambda}\right)} \right] \mathbf{1}\{x \geq l\theta\}. \end{aligned}$$

Equivalently, since only terms with $l \leq \lfloor x/\theta \rfloor$ contribute,

$$f_X(x) = \frac{e^{-q\lambda}}{1 - e^{-q\lambda}} \sum_{l=1}^{\lfloor x/\theta \rfloor} \frac{(\lambda q)^l}{l!} \frac{\beta^l (x - l\theta)^{l-1} e^{-\beta(x-l\theta)}}{\Gamma(l)}, \qquad x \geq \theta \tag{4.10}$$

with parameters $\lambda, \beta, \theta > 0$ and $q \in [0, 1]$. The summation accounts for all possible numbers of approved claims L, treating it as a latent count variable.

The term $\frac{\beta^l (x-l\theta)^{l-1} e^{-\beta(x-l\theta)}}{\Gamma(l)}$ represents the conditional gamma density of the aggregate claim amount given $L = l$, and the term $\frac{(\lambda q)^l e^{-q\lambda}}{l!\,(1-e^{-q\lambda})}$ is the probability of exactly l approved claims under the ZTPB/ZTP($q\lambda$) model. Summing over $l = 1, 2, \ldots$ produces the unconditional PDF of the aggregate approved claim amount X.

Remark. If one *does not* condition on positive approvals (i.e., allows $L = 0$), then there is an additional point mass at $x = 0$ of size $\Pr(L = 0 \mid N \geq 1) = \frac{e^{-q\lambda} - e^{-\lambda}}{1 - e^{-\lambda}}$; the continuous part for $x > 0$ is given by the same series with the mixing PMF $\frac{e^{-q\lambda}(q\lambda)^l}{l!\,(1-e^{-\lambda})}$.

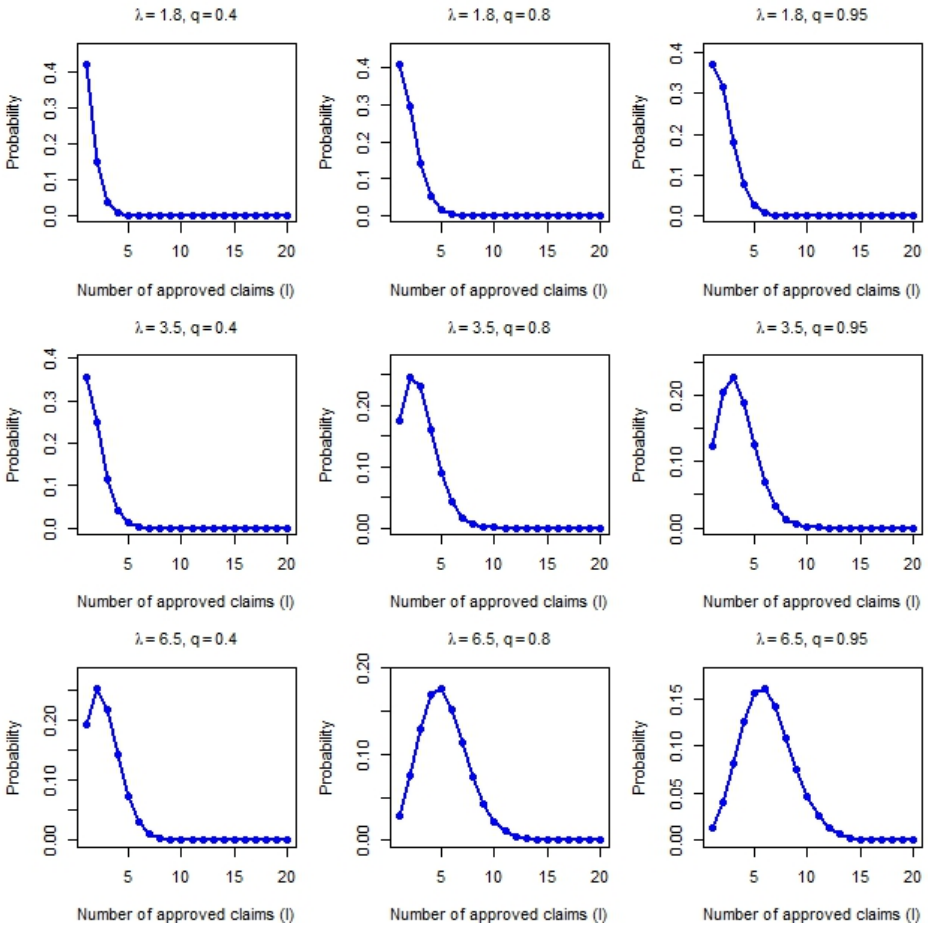


Figure 4.1: PMF for various values of λ and q.

Figure 1 illustrates how the PMF of the number of approved claims L varies with different values of the claim frequency parameter λ and the approval probability q. For smaller λ values (e.g., $\lambda = 1.8$), the distribution is heavily right-skewed, with most of its mass concentrated at $L = 1$–3, and the probability declines sharply as L increases. As λ increases (e.g., $\lambda = 3.5$ or $\lambda = 6.5$), the PMF shifts rightward and becomes more dispersed, resulting in a higher probability of larger numbers of approved claims. Similarly, larger q values (e.g., $q = 0.95$) increase the likelihood of approving more claims, flattening the distribution compared to smaller q values (e.g., $q = 0.4$), which produce shorter right tails. This figure shows that higher claim arrival rates and higher approval probabilities lead to larger expected numbers of approved claims and heavier right tails in the PMF.

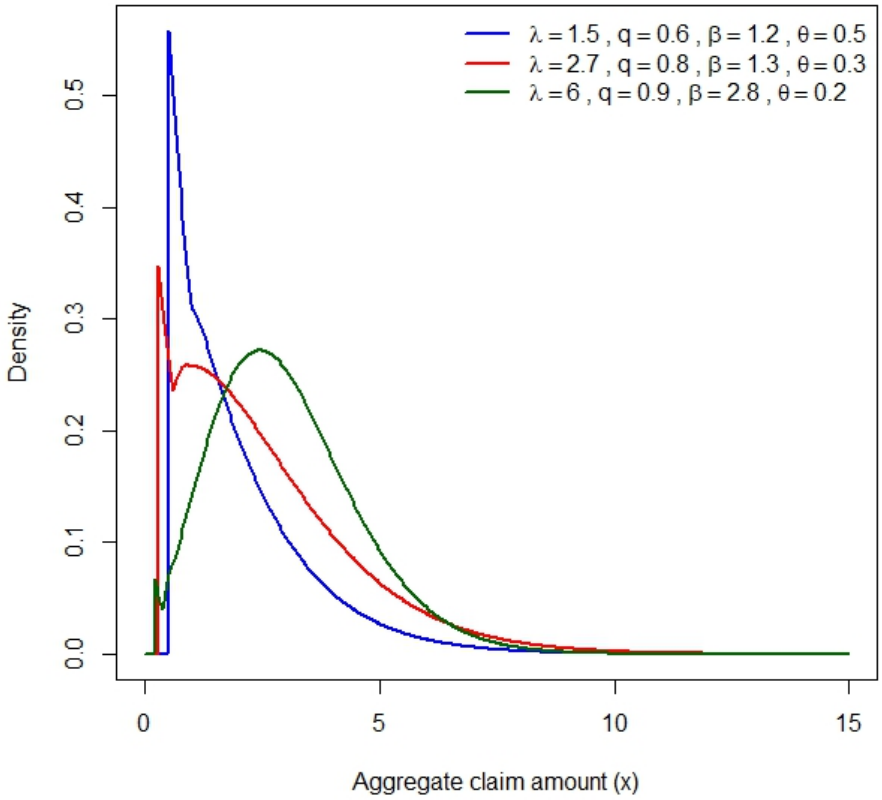


Figure 4.2: The plot corresponding to equation (4.11) for different parameter values.

Figure 2 presents the PDF of the aggregate approved claim amounts X for different parameter combinations. When λ and q are small (e.g., $\lambda = 1.5$, $q = 0.6$, $\beta = 1.2$, $\theta = 0.5$; blue curve), the density is sharply peaked near the deductible threshold θ, indicating that small aggregate claim amounts dominate, and the right tail decays rapidly. For moderate parameter values (e.g., $\lambda = 2.7$, $q = 0.8$, $\beta = 1.3$, $\theta = 0.3$; red curve), the peak shifts slightly rightward and flattens, reflecting a greater spread in possible aggregate amounts. For larger λ and q (e.g., $\lambda = 6$, $q = 0.9$, $\beta = 2.8$, $\theta = 0.2$; green curve), the distribution becomes more dispersed and unimodal with its peak further away from θ, showing higher probability for larger aggregate claims. This means that higher claim arrival rates (λ) and approval probabilities (q) lead to broader

distributions with heavier right tails, while β and θ influence the concentration and location of the density around smaller claim amounts.

As illustrated in Figure 2, the PDF of the aggregate approved claim amount is not smooth at the lower boundary. This behavior is a structural feature of the ZTP-SE compound model. Since each approved claim must exceed the deductible *theta*, the aggregate distribution has support starting at $x = \theta$. For $L = 1$, the conditional distribution $\theta + \text{Exp}(\beta)$ induces a positive jump at $x = \theta$ of size $\beta P(L = 1)$, while for $L \geq 2$ the shifted-gamma components enter at $x = l\theta$ with density initially zero but positive slope, producing visible kinks at multiples of θ. The apparent peak near the origin is therefore a reflection of the theoretical structure of the model.

4.1.10 The parameter estimation via the expectation-maximization (EM) approach

Having derived the distribution of the aggregate approved claim amounts, we now turn to the problem of parameter estimation. Direct maximization of the likelihood associated with equation (4.11) is analytically intractable because of the summation over the latent claim count inside the log-likelihood. To address this, we apply the expectation–maximization (EM) algorithm, which is well-suited for models involving latent variables and missing data. The EM framework provides numerical stability, reliable convergence, and ease of implementation, making it widely applicable in statistics and related fields. This section formulates the observed-data likelihood, derives the E- and M-steps for the key parameters, and establishes update rules for λ, q, β, and θ.

From equation (4.11), we maximize the observed-data log-likelihood, summing over all observations and marginalizing over the latent count l, since direct maximization is cumbersome due to the summation inside the logarithm. We write

$$\mathcal{L}(\lambda, q, \beta, \theta) = \sum_{i=1}^{n} \log \left[\sum_{l=1}^{\infty} P_L(l; \lambda, q)\, f_{X|L}(X_i \mid l; \beta, \theta) \right],$$

where $P_L(l; \lambda, q)$ is defined in equation (4.6) and $f_{X|L}(x \mid l; \beta, \theta)$ in equation (4.10).

Posterior weights. We define the posterior weight for observation X_i and count l:

$$W_{il} = \mathbb{P}\big(L = l \mid X_i; \lambda, q, \beta, \theta\big).$$

By Bayes' rule,

$$W_{il} = \frac{f_{X|L}(X_i \mid l; \beta, \theta)\, P_L(l; \lambda, q)}{\displaystyle\sum_{k=1}^{\infty} f_{X|L}(X_i \mid k; \beta, \theta)\, P_L(k; \lambda, q)}. \tag{4.11}$$

The numerator is the joint density of $(X_i, L) = (X_i, l)$,

$$\underbrace{f_{X|L}(X_i \mid l)}_{\text{equation (4.10)}}\ \underbrace{P_L(l; \lambda, q)}_{\text{equation (4.6)}},$$

and the denominator is the marginal $f_X(X_i) = \sum_{k \geq 1} f_{X|L}(X_i \mid k)\, P_L(k; \lambda, q)$. For numerical work, we truncate the infinite sum at a large M:

$$W_{il} = \frac{f_{X|L}(X_i \mid l; \beta, \theta)\, P_L(l; \lambda, q)}{\displaystyle\sum_{k=1}^{M} f_{X|L}(X_i \mid k; \beta, \theta)\, P_L(k; \lambda, q)}.$$

(Recall $f_{X|L}(X_i \mid l)$ carries the indicator $\mathbf{1}\{X_i \geq l\theta\}$.)

Expected complete-data log-likelihood. The EM Q-function is

$$\psi(\lambda, q, \beta, \theta) = \sum_{i=1}^{n} \sum_{l=1}^{M} W_{i,l} \Big\{ \log P_L(l; \lambda, q) + \log f_{X|L}(X_i \mid l; \beta, \theta) \Big\}.$$

With the (positive-approvals) mixing law $P_L(l; \lambda, q) = \dfrac{(q\lambda)^l e^{-q\lambda}}{l!\,(1 - e^{-q\lambda})}$ (eq. (4.6)) and the shifted-gamma density (eq. (4.10)),

$$\psi(\lambda, q, \beta, \theta) = \sum_{i,l} W_{i,l} \Big[l \log(q\lambda) - q\lambda - \log l! - \log\big(1 - e^{-q\lambda}\big) + \; l \log \beta + (l-1) \log\big(X_i - l\theta\big) - \beta(X_i - l\theta) - \log \Gamma(l) \Big].$$

Define the summaries

$$S_1 = \sum_{i,l} W_{i,l}\, l, \qquad S_y = \sum_{i,l} W_{i,l}\, (X_i - l\theta), \qquad \sum_{l} W_{i,l} = 1 \;\Rightarrow\; \sum_{i,l} W_{i,l} = n.$$

Update for β.

$$\psi(\beta) = \sum_{i,l} W_{i,l} \big[l \log \beta - \beta(X_i - l\theta) \big], \quad \frac{\partial \psi}{\partial \beta} = \frac{S_1}{\beta} - S_y = 0 \;\Rightarrow\; \beta^{\text{new}} = \frac{S_1}{S_y}\,.$$

Since $\partial^2 \psi / \partial \beta^2 = -S_1/\beta^2 < 0$, the update is the unique maximizer.

Update for θ (implicit). With $y_{i,l} = X_i - l\theta$,

$$\frac{\partial \psi}{\partial \theta} = \sum_{i,l} W_{i,l} \left[-\frac{l(l-1)}{y_{i,l}} + \beta\, l \right] = 0 \;\Rightarrow\; \beta\, S_1 = \sum_{i,l} W_{i,l}\, \frac{l(l-1)}{X_i - l\theta}\,.$$

This defines θ^{new} implicitly (solve numerically on $\{\theta : X_i > l\theta\}$). Moreover,

$$\frac{\partial^2 \psi}{\partial \theta^2} = -\sum_{i,l} W_{i,l} \frac{l^2(l-1)}{(X_i - l\theta)^2} < 0,$$

so the solution is the unique maximizer on the feasible domain.

Updates for q and λ. The (λ, q)-part of ψ is

$$\psi_{\lambda,q} = \sum_{i,l} W_{i,l} \Big[l \log(q\lambda) - q\lambda - \log(1 - e^{-q\lambda}) \Big].$$

Partial derivatives (holding the other parameter fixed):

$$\frac{\partial \psi_{\lambda,q}}{\partial q} = \frac{S_1}{q} - n\lambda - n\, \frac{\lambda e^{-q\lambda}}{1 - e^{-q\lambda}}, \qquad \frac{\partial \psi_{\lambda,q}}{\partial \lambda} = \frac{S_1}{\lambda} - nq - n\, \frac{q e^{-q\lambda}}{1 - e^{-q\lambda}}.$$

It is convenient to reparametrize by $\mu := q\lambda$. Then

$$\psi_\mu = \sum_{i,l} W_{i,l}\Big[l \log \mu - \log l!\Big] \; - \; n\,\mu \; - \; n \, \log(1 - e^{-\mu}) \; + \; \text{const},$$

so

$$\frac{\partial \psi_\mu}{\partial \mu} = \frac{S_1}{\mu} - n - \frac{n\, e^{-\mu}}{1 - e^{-\mu}} = 0 \iff \frac{S_1}{\mu} = \frac{n}{1 - e^{-\mu}} \, ,$$

with

$$\frac{\partial^2 \psi_\mu}{\partial \mu^2} = -\frac{S_1}{\mu^2} - \frac{n\, e^{-\mu}}{(1 - e^{-\mu})^2} < 0,$$

so the root is unique and ψ_μ is strictly concave. Solve this scalar equation for μ^{new} (e.g., Newton or `uniroot`), then recover (λ, q) using the chosen identifiability link (e.g., $q = e^{-\beta(\theta_0 - \theta)}$ from the shifted-exponential approval with known θ_0, hence $\lambda = \mu/q$).

Substitution of explicit forms. Using equation (4.6)

$$P_L(l; \lambda, q) = \frac{(q\lambda)^l e^{-q\lambda}}{l!\,(1 - e^{-q\lambda})},$$

and equation (4.10) for $f_{X|L}$, the weight (4.11) becomes

$$W_{il} = \frac{\dfrac{(q\lambda)^l e^{-q\lambda}}{l!\,(1 - e^{-q\lambda})} \, \dfrac{\beta^l (X_i - l\theta)^{l-1} e^{-\beta(X_i - l\theta)}}{\Gamma(l)} \, \mathbf{1}\{X_i \geq l\theta\}}{\displaystyle\sum_{k=1}^{M} \dfrac{(q\lambda)^k e^{-q\lambda}}{k!\,(1 - e^{-q\lambda})} \, \dfrac{\beta^k (X_i - k\theta)^{k-1} e^{-\beta(X_i - k\theta)}}{\Gamma(k)} \, \mathbf{1}\{X_i \geq k\theta\}} \, ,$$

which is the expression used in the EM implementation.

4.1.11 Theoretical properties of the aggregate approved claim amounts

This section develops the theoretical foundations of the aggregate approved claim amounts by deriving their MGF, CGF, and related higher-order derivatives. These properties are crucial for understanding the distributional behavior of the compound model, particularly in the context of saddlepoint approximation, where accurate characterizations of moments and cumulants are vital. Leveraging the independence of individual claim severities and the structure of the ZTPB frequency distribution, the section establishes closed-form expressions that facilitate analytical insights and numerical implementations.

Since $T_j \sim \text{SE}(\theta, \beta)$, the MFG is given as:

$$M_T(t) = \frac{\beta e^{\theta t}}{\theta - t}, \; t < \theta. \tag{4.12}$$

Similarly, $L \sim \text{ZTPB}(l, q)$ with the MGF ($M_L(t) = E(e^{Lt})$) defined in equation (4.8).

4.1.12 MGF of the aggregate approved claim count

Let $X = \sum_{j=1}^{L} T_j$ be the aggregate approved claim amounts. To derive the MGF, we start as:

$$M_X(t) = E\left[e^{tX}\right].$$

Since L is a random variable representing the number of approved claims, we apply the law of total expectation:

$$M_X(t) = E\left[e^{tX}\right] = E_L\left[E\left(e^{tX} \mid L\right)\right],$$

where $E_L[\cdot]$ denotes expectation over the distribution of L.

Conditional MGF given $L = l$. That is, if $L = l$, then

$$X \mid L = l = \sum_{j=1}^{l} T_j.$$

with $T_1, ..., T_l$ i.i.d. Thus, the conditional MGF of X given $L = l$ is

$$M_{X|L=l}(t) = E\left[e^{tX} \mid L = l\right] = E\left[e^{t\sum_{j=1}^{l} T_j}\right].$$

Using independence of T_j and since $T_1, \ldots, T_l$ are independent and identically distributed (i.i.d.), we can factorize the expectation as:

$$e^{t\sum_{j=1}^{l} T_j} = \prod_{j=1}^{l} e^{tT_j} \quad \Rightarrow \quad E\left[e^{t\sum_{j=1}^{l} T_j}\right] = E\left[\prod_{j=1}^{l} e^{tT_j}\right] = \prod_{j=1}^{l} E\left[e^{tT_j}\right].$$

Let us express this via the MGF of T and by definition, the MGF of each T_j is:

$$M_T(t) = E\left[e^{tT_j}\right].$$

Hence,

$$M_{X|L=l}(t) = \prod_{j=1}^{l} M_T(t) = \left[M_T(t)\right]^l.$$

Using the law of total expectation again, the unconditional MGF of X is:

$$M_X(t) = E_L\left[M_{X|L}(t)\right] = E_L\left[\left(M_T(t)\right)^L\right].$$

This indicates that the MGF of the aggregate claim amount X can be expressed as the expectation of the L-th power of the MGF of the individual claim severity T. That is:

$$M_X(t) = E\left[\left(M_T(t)\right)^L\right].$$

$$M_X(t) = E\left[\left(M_T(t)\right)^L\right] = E_L\left[M_T(t)\right].$$

where $M_T(t)$ is defined in equation (4.8). This form enables us to easily derive the CGF of X based on its identity and subsequently derive saddlepoint approximations.

Unconditional MGF via the PGF of L**.** Using the fact that

$$M_X(t) = E_L\left[\left(M_T(t)\right)^L\right]$$

Equivalently,

$$M_X(t) = G_L(M_T(t))$$

where $G_L(s) = E(s^L)$ is the PGF of L.

However, we are interested in positive approvals only; hence, we evaluate the PGF of L and denote it as $G_X(s) = E(X^L)$ and recall $X = \sum_{j=1}^{L} T_j$.

We already know that:

$$M_X(t) = E\left[e^{tX}\right] = E\left[(M_T(t))^L\right],$$

where $M_T(t) = E[e^{tT_j}]$ is the MGF of a single claim severity T_j.

Now, let $G_L(s)$ denote the PGF of L as already stated:

$$G_L(s) = E[s^L].$$

By setting $s = M_T(t)$, we immediately obtain:

$$M_X(t) = E\left[(M_T(t))^L\right] = G_L\big(M_T(t)\big).$$

For $L \sim \text{ZTPB}(\lambda q)$, the PGF is:

$$G_L(s) = \frac{e^{\lambda(1-q+qs)} - 1}{1 - e^{-q\lambda}}.$$

Substituting $s = M_T(t)$ and using $T_j \sim \text{SE}(\theta, \beta)$ with MGF:

$$M_T(t) = \frac{\beta e^{\theta t}}{\theta - t}, \quad t < \beta,$$

we obtain the MGF of X:

$$M_X(t) = \frac{e^{\lambda\left[1-q+q\left(\frac{\beta e^{\theta t}}{\theta - t}\right)\right]} - 1}{1 - e^{-q\lambda}}. \tag{4.13}$$

Finally, the cumulant generating function (CGF) of X is simply

$$K_X(t) = \log\left[\frac{e^{\lambda\left[1-q+q\left(\frac{\beta e^{\theta t}}{\theta - t}\right)\right]} - 1}{1 - e^{-q\lambda}}\right]. \tag{4.14}$$

We define $K_X^{'}(t_0) = x$, and we will call this the saddlepoint equation such that $K_X^{'}(t_0)$ could be solved numerically to obtain t_0 estimate.

The estimate of the saddlepoint t_0 is simply the root where the CGF's slope matches the median of the observation.

Subsequently, t_0 is defined as the saddlepoint satisfying $K_X^{'}(t_0) = x$ and will t_0 the introduced saddpoint estimator.

Let

$$A(t) = \lambda\left(1 - q + qg(t)\right)$$

where $g(t) = \frac{\beta e^{\theta t}}{\theta - t}$ and $S(t) = e^{A(t)}$ so $S(t) > 0$ and $B = 1 - e^{-q\lambda}$

$$g'(t) = \frac{\beta e^{\theta t}\left[\theta(\theta - t) + 1\right]}{(\theta - t)^2}$$

$$g"(t) = \frac{\beta e^{\theta t}\left[\theta^2(\theta-t)^2 + 2\theta(\theta-t) + 2\right]}{(\theta-t)^3}$$

$$g'''(t) = \frac{\beta e^{\theta t}\left[\theta^2(\theta-t)^3 + 3\theta^2(\theta-t)^2 + 6\theta(\theta-t) + 6\right]}{(\theta-t)^4}$$

$$A'(t) = q\lambda g'(t)$$

$$A''(t) = q\lambda g"(t)$$

$$A'''(t) = q\lambda g'''(t)$$

.

Hence,

$$K_X^{'}(t_0) = \frac{S(t_0)A'(t_0)}{S(t_0)-1} = x \tag{4.15}$$

Similarly,

$$K_X^{"}(t_0) = \frac{S(t_0)\left[(S(t_0)-1)\,A"(t_0) - (A'(t_0))^2\right]}{(S(t_0)-1)^2}$$

and

$$K_X^{'''}(t_0) = \frac{S(t_0)\left[(S(t_0)-1)^3 A'''(t_0) + (S(t_0)-1)(e^{2A(t_0)} - D\right]}{(S(t_0)-1)^4}$$

, where

$$D = 4S(t_0)+1)A^{'}(t_0)A^{''}(t_0) + (S(t_0)+1)(A^{'}(t_0))^3$$

Notes. (i) Domain: all formulas hold for $t < \beta$. (ii) If one were to work with the *untruncated* mixing law (allowing $L = 0$), replace $q\lambda(g(t)-1)$ by $\lambda(1-q+q\,g(t))$ and c by 1, reproducing $M_X(t) = \dfrac{e^{\lambda(1-q+qg(t))}-1}{1-e^{-\lambda}}$.

4.1.13 The saddlepoint approximation

In this section, we employed two saddlepoint approximation techniques for our modeling problem, specifically for estimating the saddlepoint of equation (4.16): the classical method proposed by Daniels [27] and its extended form introduced by Wang [125], which builds upon the foundational work of Daniels. Considering both methods allows for a comprehensive comparison and ensures robustness in our approximation results.

4.1.14 The Daniels [27] approximation approach

Let

$$\omega = \text{sign}(t_0)\sqrt{2\,(t_0 x - K_X(t_0))}.$$

Then, the CDF of X is approximated by

$$P(X \le x) \approx \Phi(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\omega} e^{-z^2/2}\,dz, \tag{4.16}$$

where $\Phi(\omega)$ is the standard normal CDF.

Using the error function, we can also write

$$P(X \leq x) \approx \Phi(\omega) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{\omega}{\sqrt{2}}\right)\right].$$

The corresponding PDF is

$$f_X(x) \approx \frac{d}{dx}\Phi(\omega) = \phi(\omega)\frac{d\omega}{dx},$$

where $\phi(\omega)$ is the standard normal PDF.

From the definition of ω and using $\Phi'(\omega) = \phi(\omega)$, we have

$$\frac{d\omega}{dx} = \frac{1}{2\omega}\left(2t_0 + 2x\frac{dt_0}{dx} - 2K'_X(t_0)\frac{dt_0}{dx}\right).$$

By the saddlepoint equation $K'_X(t_0) = x$, it follows that $x - K'_X(t_0) = 0$, hence

$$\frac{d\omega}{dx} = \frac{t_0}{\omega}.$$

Therefore, the approximate PDF is

$$f_X(x) \approx \phi(\omega)\frac{d\omega}{dx} = \frac{1}{\sqrt{2\pi}}e^{-\omega^2/2}\frac{t_0}{\omega}. \tag{4.17}$$

4.1.15 The Wang [125] second-order correction approximation approach

Let

$$u = t_0\sqrt{K''_X(t_0)}.$$

Then, the Wang [125] second-order corrected CDF approximation is

$$P(X \leq x) = F_X(x) \approx \Phi(\omega) + \phi(\omega)\left(\frac{1}{\omega} - \frac{1}{u}\right), \tag{4.18}$$

where

$$\omega = \operatorname{sign}(t_0)\sqrt{2\left(t_0 x - K_X(t_0)\right)},$$

and $\Phi(\cdot)$ and $\phi(\cdot)$ are the standard normal CDF and PDF, respectively.

We also have

$$\phi(\omega) = \frac{1}{\sqrt{2\pi}}e^{-\omega^2/2}, \quad \phi'(\omega) = -\omega\phi(\omega). \tag{4.19}$$

Differentiating the CDF to obtain the PDF gives

$$\begin{aligned} f_X(x) &= \frac{dF_X}{dx} \\ &= \frac{d\Phi(\omega)}{dx} + \frac{d}{dx}\left[\phi(\omega)\left(\frac{1}{\omega} - \frac{1}{u}\right)\right]. \end{aligned} \tag{4.20}$$

The first term is

$$\frac{d\Phi(\omega)}{dx} = \phi(\omega)\frac{d\omega}{dx} = \phi(\omega)\frac{t_0}{\omega}. \tag{4.21}$$

For the second term, let

$$A(\omega) = \frac{1}{\omega} - \frac{1}{u}.$$

Then,

$$\begin{aligned}
\frac{d}{dx}\left[\phi(\omega)A(\omega)\right] &= \phi'(\omega)\frac{d\omega}{dx}A(\omega) + \phi(\omega)\frac{dA}{dx} \\
&= \left[-\omega\phi(\omega)\frac{t_0}{\omega}\right]A(\omega) + \phi(\omega)\left(-\frac{t_0}{\omega^3} + \frac{1}{u^2}\frac{du}{dx}\right) \\
&= -t_0\phi(\omega)A(\omega) + \phi(\omega)\left(-\frac{t_0}{\omega^3} + \frac{1}{u^2}\frac{du}{dx}\right).
\end{aligned}$$

Adding the contributions, the exact PDF is

$$\begin{aligned}
f_X(x) &= \phi(\omega)\frac{t_0}{\omega} - t_0\phi(\omega)A(\omega) + \phi(\omega)\left(-\frac{t_0}{\omega^3} + \frac{1}{u^2}\frac{du}{dx}\right) \\
&= \phi(\omega)\left[\frac{t_0}{u} - \frac{t_0}{\omega^3} + \frac{1}{u^2}\frac{du}{dx}\right]. \qquad (4.22)
\end{aligned}$$

Since

$$u = t_0\sqrt{K''_X(t_0)},$$

differentiating and using $\frac{dt_0}{dx} = \frac{1}{K''_X(t_0)}$ (from $K'_X(t_0) = x$) gives

$$\frac{du}{dx} = \frac{1}{\sqrt{K''_X(t_0)}} + \frac{t_0 K'''_X(t_0)}{2(K''_X(t_0))^{3/2}}. \qquad (4.23)$$

If the curvature variation is small, we may approximate $\frac{du}{dx} \approx 0$. Then, equation (4.22) simplifies to

$$f_X(x) \approx t_0\,\phi(\omega)\left(\frac{1}{u} - \frac{1}{\omega^3}\right). \qquad (4.24)$$

Equations (4.19) and (4.22) present the saddlepoint approximate PDFs derived using the two respective approaches.

4.2 Numerical results

This section presents simulation studies and real-life numerical evaluations of our model, the saddlepoint approximation, and the utilization of the EM approach for the parameter estimations, assessing the fitness of the derived PDFs. For the real-life application, we utilize the "PAID" variable, representing the insurance payout amount in USD, from the `AutoClaims` dataset with a mean of 1853.035 and a variance of 7006129, indicating overdispersed data. This data contains 6,773 observations. These include 4,191 males (aged 50–97 years) and 2,582 females (aged 50–95 years). The data was obtained in `R` using the commands: `library(insuranceData)`, `data(AutoClaims)`, and `print(AutoClaims)`. This dataset represents closed claim experience from a large Midwestern U.S. property and casualty insurer for private passenger automobile insurance. The dependent variable "PAID" corresponds to the actual payout (in USD) made by the insurer for claims that were approved and closed within the reporting calendar year.

4.2.1 Saddlepoint estimation via Monte Carlo simulation with ZTPB count and SE

Since the saddlepoint equation $K'_X(t_0) = x$ has no closed-form solution, we estimate t_0 numerically using simulated data from the compound model. We generate observations of

$$X = \sum_{j=1}^{L} T_j, \qquad T_j = \theta + Y_j, \;\; Y_j \sim \text{Exp}(\theta),$$

where the approved-claim count L follows the ZTPB distribution with parameters (q, λ).

We solve $K'_X(t_0) = x$ numerically for t_0, taking x as the sample median of X.

We implemented a direct Monte Carlo simulation of the compound model in R. The discrete component L (ZTPB) is sampled via a thin-and-truncate approach that is marginally equivalent to a zero-truncated Poisson with rate λq: we draw from `rpois()` with mean λq and reject zeros. Severities are generated with `rexp()` and shifted by θ; aggregation uses `sum()`, repetition uses `replicate()`, and the nonlinear root t_0 is obtained with `uniroot()`. Bootstrap resampling (via `sample()`) provides variability and bias estimates for t_0.

4.2.2 Simulation procedure

- Draw an approved claim count $L \sim \text{ZTPB}(q, \lambda)$ (implemented as ZTP with mean λq).
- Conditional on L, generate i.i.d. severities $T_j = \theta + Y_j$ with $Y_j \sim \text{Exp}(\theta)$ (i.e., `rexp()` plus a constant shift θ).
- Compute the compound sum $X = \sum_{j=1}^{L} T_j$.
- Repeat steps 1–3 to obtain n i.i.d. samples of X.
- Set $x = \text{median}(X_1, \ldots, X_l)$ and solve $K'_X(t_0) = x$ for t_0 using `uniroot()`.
- Assess uncertainty of the estimate of t_0 via bootstrap: resample $\{X_i\}$ with replacement, recompute x and t_0, and summarize the bootstrap distribution (mean, SD, bias).

For the sensitivity analysis, a small perturbations of 0.1 and 0.4 were applied to the parameters λ, q, θ, and β, which are reasonable as these values avoid entering regions where the formula, such as the saddlepoint equation, may break down. The objective of these analyses is to assess how close t_0 is to the true value of the statistic based on X, and how much t_0 varies across repeated simulations.

From Table 4.1 and Figure 4.3, the bootstrap bias of t_0 decreases from 0.0013 at $n = 500$ to near zero at $n = 70{,}000$, while the bootstrap SD drops from 0.0111 to 0.0007, confirming the asymptotic unbiasedness and increasing stability of the estimate of t_0. The SD curve declines sharply with n, and the bias approaches zero, indicating consistency.

Figure 4.4 shows the sensitivity of t_0 to parameter perturbations. The $(\beta + 0.1)$ produces the largest upward shift across all n, followed by the $(\theta + 0.1)$, both remaining stable as n increases. The $(\lambda + 0.1)$ lies just below the baseline (t_0 from X), indicating

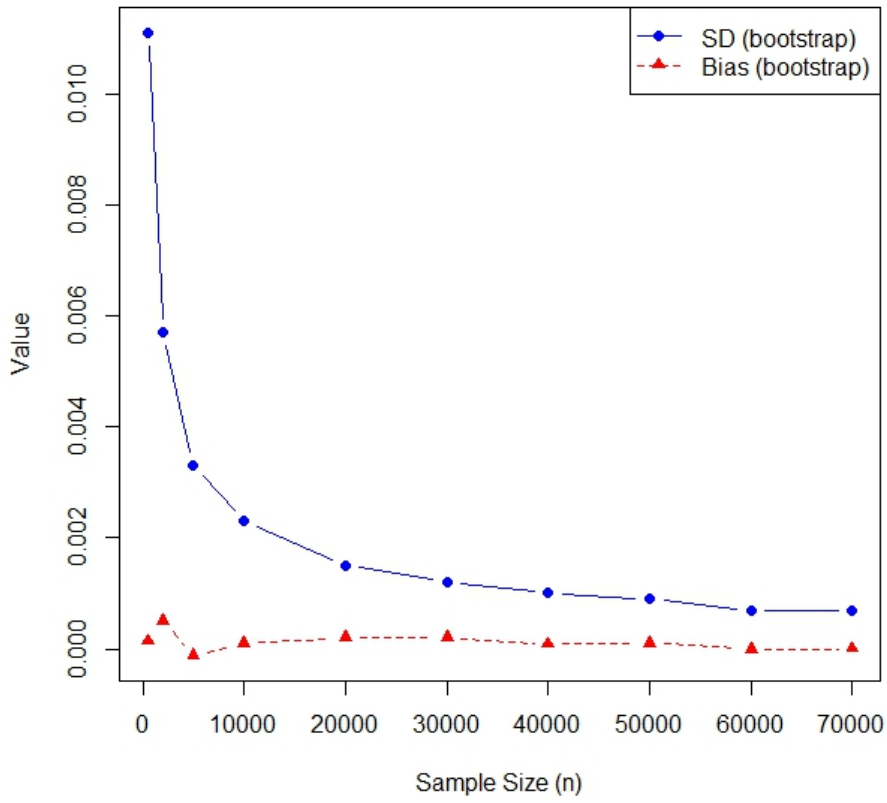


Figure 4.3: Bootstrap SD and bias of the saddlepoint t_0 vs. sample size (n) as shown in Table 4.1

Table 4.1: Simulation results for estimating t_0 using the parameter values: $\lambda = 0.8$, $q = 0.65$, $\theta = 2.5$, and $\beta = 1.2$, with perturbation of parameters: $\lambda + 0.1$, $q + 0.1$, $\theta + 0.1$, and $\beta + 0.1$.

n	t_0(from X)	t_0(bootstrap)	SD(bootstrap)	Bias(bootstrap)	$t_0(\lambda + 0.1)$	$t_0(q + 0.1)$	$t_0(\theta + 0.1)$	$t_0(\beta + 0.1)$
500	0.9946	0.9959	0.0111	0.0013	0.9825	0.8818	1.0115	1.0681
2000	0.9802	0.9806	0.0057	0.0005	0.9683	0.8660	0.9974	1.0530
5000	0.9816	0.9814	0.0033	-1.193×10^{-4}	0.9697	0.8676	0.9987	1.0554
10000	0.9817	0.9818	0.0023	9.351×10^{-5}	0.9698	0.8677	0.9988	1.0555
20000	0.9780	0.9783	0.0015	0.0002	0.9663	0.8637	0.9952	1.0519
30000	0.9776	0.9778	0.0012	0.0002	0.9659	0.8633	0.9949	1.0516
40000	0.9780	0.9781	0.0010	8.372×10^{-5}	0.9663	0.8637	0.9952	1.0519
50000	0.9777	0.9778	0.0009	0.0001	0.9659	0.8633	0.9949	1.0516
60000	0.9771	0.9770	0.0007	-1.011×10^{-5}	0.9653	0.8627	0.9943	1.0510
70000	0.9771	0.9771	0.0007	3.643×10^{-6}	0.9654	0.8627	0.9944	1.0511

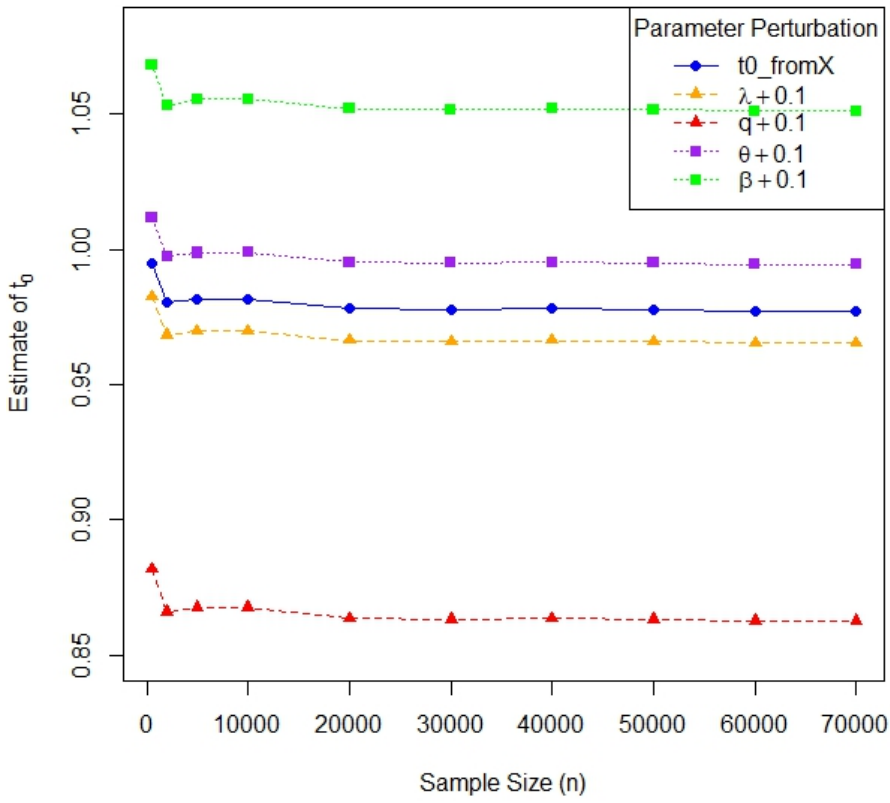


Figure 4.4: Sensitivity of parameter perturbations to t_0, as shown in Table 4.1

low positive sensitivity. In contrast, the $(q + 0.1)$ shows the largest negative shift, starting near 0.88 and stabilizing around 0.86. The sensitivity ranking is $\beta > \theta > \lambda$ (positive) and q (negative).

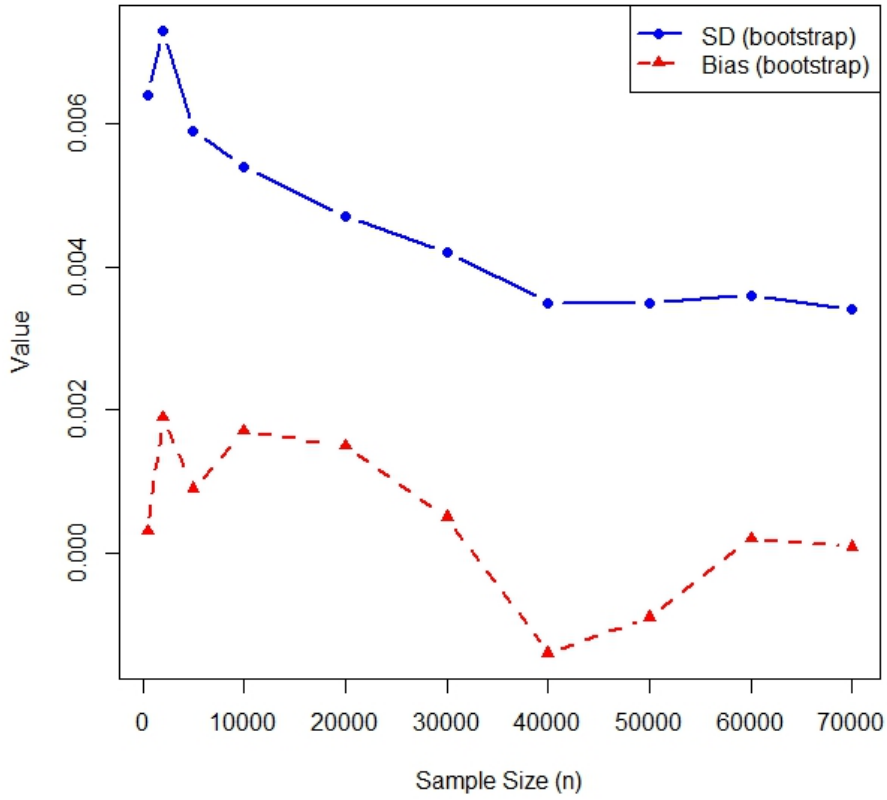


Figure 4.5: Bootstrap SD and bias of the saddlepoint t_0 vs. sample size (n) as shown in Table 4.2

From Table 4.2 and Figure 4.5, the bootstrap bias of t_0 fluctuates slightly around zero, showing no upward or downward trend, while the bootstrap SD steadily decreases, indicating improved estimator stability with larger samples. Figure 4.6 shows the sensitivity of t_0 to parameter perturbations. The $(\beta+0.4)$ produces the largest upward shift, followed by the $(\theta+0.4)$, both remaining stable across n. The $(\lambda+0.4)$ lies slightly below the blue baseline (t_0 from X), indicating low positive sensitivity. The $(q + 0.4)$ yields the strongest negative shift, consistently lowest across all n. The sensitivity ranking is $\beta > \theta > \lambda$ (positive) and q (negative). This result is consistent with the inference from Table 4.1, Figures 4.3 and 4.4.

4.2.3 Efficiency of our saddlepoint approximation estimate (t_0)

In this section, we compare the efficiency of our numerical estimator with that derived in Alhejaili and Alghamedi [7]. While Alhejaili and Alghamedi [7] derived the saddlepoint estimator in their study, they focused on approximating the cumulative distribution function based on the saddlepoint method. They observed that their approximation technique performs excellently in terms of accuracy and computational

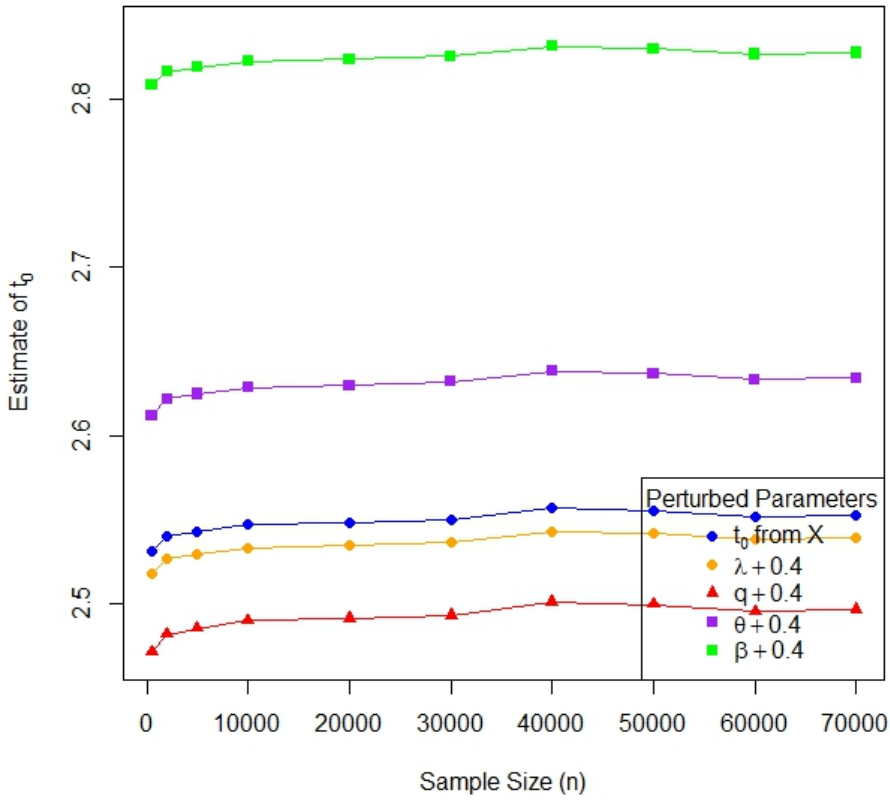


Figure 4.6: Sensitivity of parameter perturbations to t_0, as shown in Table 4.2

Table 4.2: Simulation results for estimating t_0 using the parameter values: $\lambda = 5.8$, $q = 0.8$, $\theta = 3.5$, and $\beta = 2.7$ with perturbation of parameters: $\lambda + 0.4$, $q + 0.4$, $\theta + 0.4$, and $\beta + 0.4$.

n	t_0(from X)	t_0(bootstrap)	SD(bootstrap)	Bias(bootstrap)	$t_0(\lambda + 0.4)$	$t_0(q + 0.4)$	$t_0(\theta + 0.4)$	$t_0(\beta + 0.4)$
500	2.5314	2.5317	0.0064	0.0003	2.5178	2.4714	2.6120	2.8084
2000	2.5401	2.5420	0.0073	0.0019	2.5265	2.4819	2.6217	2.8163
5000	2.5431	2.5430	0.0059	0.0009	2.5295	2.4854	2.6246	2.8189
10000	2.5471	2.5488	0.0054	0.0017	2.5335	2.4901	2.6286	2.8224
20000	2.5482	2.5497	0.0047	0.0015	2.5346	2.4914	2.6298	2.8235
30000	2.5504	2.5509	0.0042	0.0005	2.5368	2.4930	2.6319	2.8254
40000	2.5567	2.5553	0.0035	-0.0014	2.5431	2.5013	2.6382	2.8311
50000	2.5553	2.5544	0.0035	-0.0009	2.5417	2.4997	2.6368	2.8298
60000	2.5516	2.5518	0.0036	0.0002	2.5380	2.4954	2.6332	2.8265
70000	2.5527	2.5527	0.0034	8.621×10^{-5}	2.5391	2.4967	2.6342	2.8274

efficiency through simulation studies. However, the authors did not report the efficiency in their estimate of the derived saddlepoint estimator. Thus, the inefficiencies of their estimator observed here are empirical findings specific to the estimation task and parameter configuration used in this study.

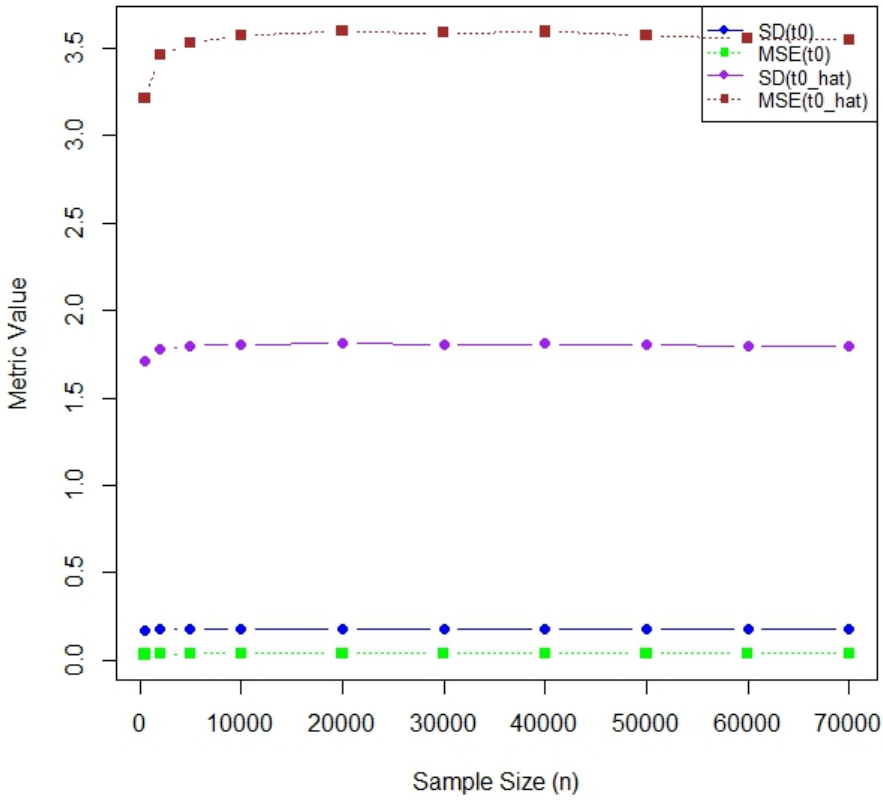


Figure 4.7: Comparison of metrics vs. sample size under the parameter settings $\lambda = 0.8$, $q = 0.65$, $\theta = 2.5$, and $\beta = 1.2$

The results in Table 4.3 and Figure 4.7 show that our saddlepoint estimator t_0 achieves consistently low SD (≈ 0.18) and very low MSE (≈ 0.038) across all sample sizes, indicating high stability and accuracy. In contrast, the Alhejaili and Alghamedi [7] estimator ($\hat{t}_0$) exhibits much larger variability (SD ≈ 1.8) and a much higher MSE (≈ 3.55), with little improvement as n increases. This confirms that our introduced saddlepoint estimator shows considerable efficiency, accuracy, and robustness.

The results in Table 4 and Figure 8 show that our saddlepoint estimator t_0 maintains consistently low standard deviation (≈ 0.121) and low MSE (≈ 0.015) across all sample sizes, indicating high accuracy and stability. In contrast, the Alhejaili and Alghamedi [7] estimator $\hat{t}_0$ exhibits extremely large variability (SD $\approx$ 1560) and much higher MSE ($\approx$ 25,900), with no notable improvement as n increases. This confirms the clear superiority and robustness of our introduced estimator.

Table 4.3: Comparative analysis of the efficiency of our saddlepoint estimator (t_0) and the one ($\hat{t}_0$) derived in Alhejaili and Alghamedi [7] using the parameter values: $\lambda = 0.8$, $q = 0.65$, $\theta = 2.5$, and $\beta = 1.2$.

n	t_0	SD_{t_0}	MSE_{t_0}	$\hat{t}_0$	$\mathrm{SD}_{\hat{t}_0}$	$\mathrm{MSE}_{\hat{t}_0}$
500	0.9740	0.1698	0.0329	32.8352	1.7109	3.2107
2000	0.9770	0.1770	0.0361	32.8003	1.7774	3.4602
5000	0.9782	0.1802	0.0369	32.7872	1.7991	3.5316
10000	0.9767	0.1812	0.0374	32.8041	1.8051	3.5716
20000	0.9766	0.1817	0.0377	32.8057	1.8093	3.5947
30000	0.9750	0.1813	0.0376	32.8123	1.8068	3.5861
40000	0.9750	0.1816	0.0377	32.8126	1.8087	3.5943
50000	0.9759	0.1810	0.0375	32.8139	1.8040	3.5745
60000	0.9763	0.1805	0.0372	32.8086	1.7992	3.5525
70000	0.9764	0.1805	0.0372	32.8075	1.7980	3.5498

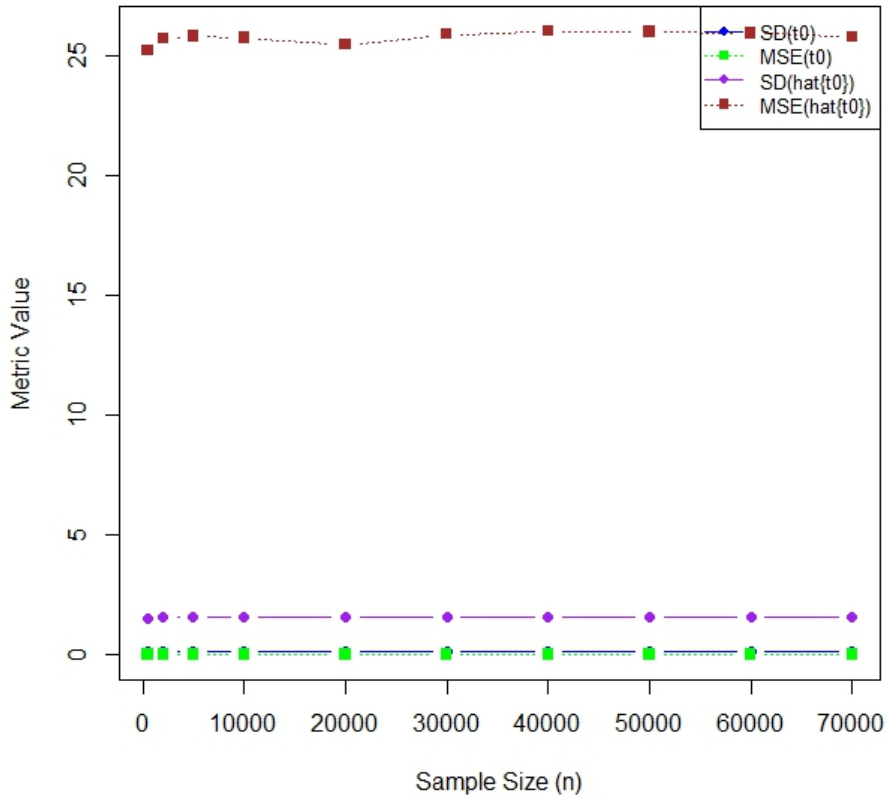


Figure 4.8: Comparison of metrics vs. sample size under the parameter settings $\lambda = 5.8$, $q = 0.8$, $\theta = 3.5$, and $\beta = 2.7$.

Table 4.4: Comparative analysis of the efficiency of our saddlepoint estimator (t_0) and the one ($\hat{t}_0$) derived in Alhejaili and AlGhamedi (2024) using the parameter values: $\lambda = 5.8$, $q = 0.8$, $\theta = 3.5$, and $\beta = 2.7$.

n	t_0	SD_{t_0}	MSE_{t_0}	$\hat{t}_0$	$\text{SD}_{\hat{t}_0}$	$\text{MSE}_{\hat{t}_0}$
500	2.5587	0.1185	0.0149	21990.62	1532.25	25193
2000	2.5462	0.1210	0.0152	22143.49	1576.48	25699
5000	2.5537	0.1211	0.0153	22051.41	1563.69	25802
10000	2.5485	0.1218	0.0153	22115.67	1572.30	25743
20000	2.5492	0.1210	0.0151	22106.72	1562.55	25445
30000	2.5545	0.1211	0.0154	22042.20	1562.80	25874
40000	2.5561	0.1210	0.0154	22021.69	1562.06	25997
50000	2.5568	0.1209	0.0154	22013.78	1559.93	25989
60000	2.5562	0.1209	0.0154	22021.64	1559.64	25920
70000	2.5569	0.1205	0.0153	22012.13	1553.80	25784

4.2.4 The real-life data application

This section presents the real-life numerical evaluation of our introduced model and assesses the fitness of the saddlepoint PDFs obtained when the model is derived using the approaches of Daniels [27] and Wang [125]. First, we applied the EM algorithm to obtain initial parameter estimates, yielding $\hat{\lambda} = 0.0409$, $\hat{q} = 0.9999$, $\hat{\beta} = 0.0006$, and $\hat{\theta} = 4.7499$. These values served as starting points for a refinement step based on maximum likelihood estimation (MLE) under the saddlepoint approximation framework. In this step, the saddlepoint-based likelihood function was numerically maximized with respect to all parameters using the observed "PAID" data, and the resulting saddlepoint loglikelihood (LL), Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), and Kolmogorov–Smirnov (KS) statistics were computed to assess the goodness-of-fit for each derivation of the model.

These statistical criteria were used to determine which derivation of our model provided the best fit to the data. Furthermore, the adequacy of each fitted PDF was visually evaluated by comparing its corresponding CDF to the empirical CDF (ECDF) of the "PAID" data. Table 4.5 presents a comparative analysis between the Daniels-derived and Wang-derived versions of our model.

The results in Table 4.5 indicate that our introduced model, derived using the Daniels [27] approach, provides a considerably better fit to the "PAID" data than when derived using the Wang [125] approach. Specifically, the Daniels-based version of our model achieves a less negative log likelihood and substantially lower AIC and BIC values, indicating a superior fit and greater parsimony. Furthermore, the KS statistic for the Daniels-based derivation is considerably smaller than that for the Wang-based derivation, indicating that the Daniels version of our model more closely matches the empirical distribution of the observed claim amounts data under consideration.

Figure 4.9 (Daniels approach) and Figure 4.10 (Wang approach) visually compare the ECDF of the "PAID" data with the CDFs fitted using each version of our model. In Figure 9, corresponding to the Daniels-derived model, the fitted CDF closely follows the

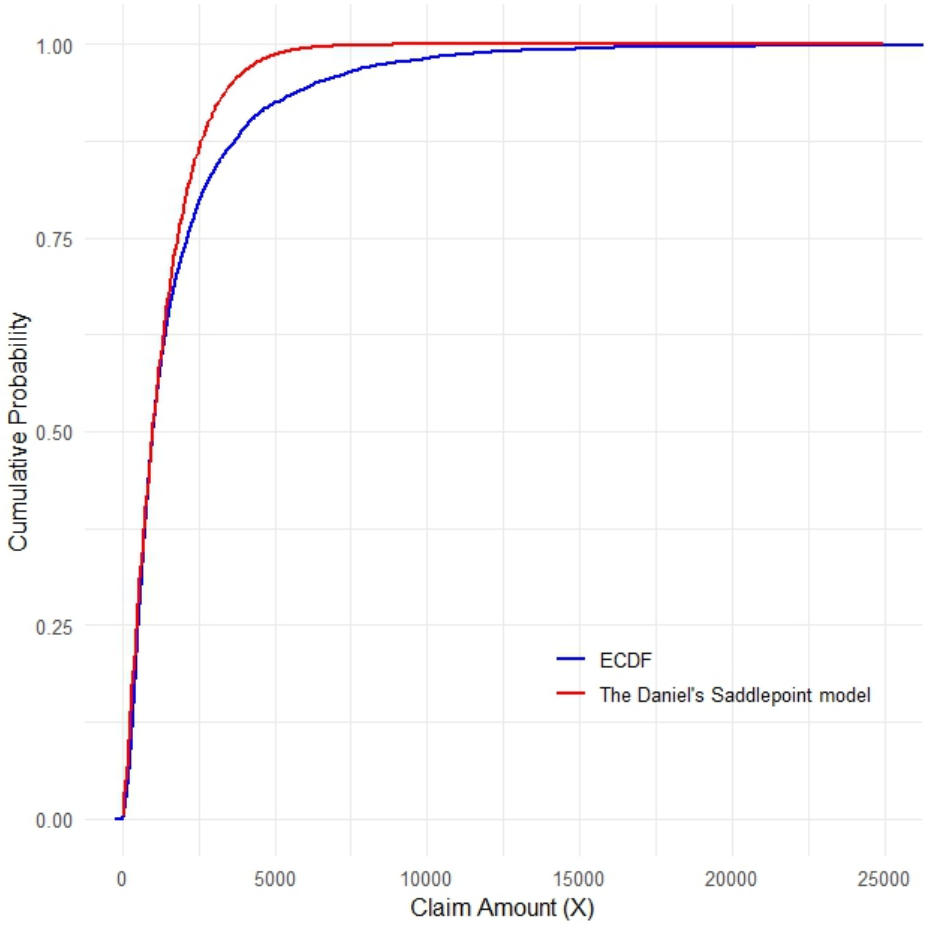


Figure 4.9: The fitting of the distribution derived from the Daniels [27] approach

Table 4.5: Comparative analysis of the fitted saddlepoint PDFs of our proposed model when derived using the Daniels [27] and Wang [125] approaches, applied to the "PAID" data

	Saddlepoint PDF derived from	
	Daniels [27] approach	Wang [125] approach
Saddlepoint LL	−5853	−4919
AIC	1171	9846
BIC	1174	9873
KS Statistic	0.0812	0.5671

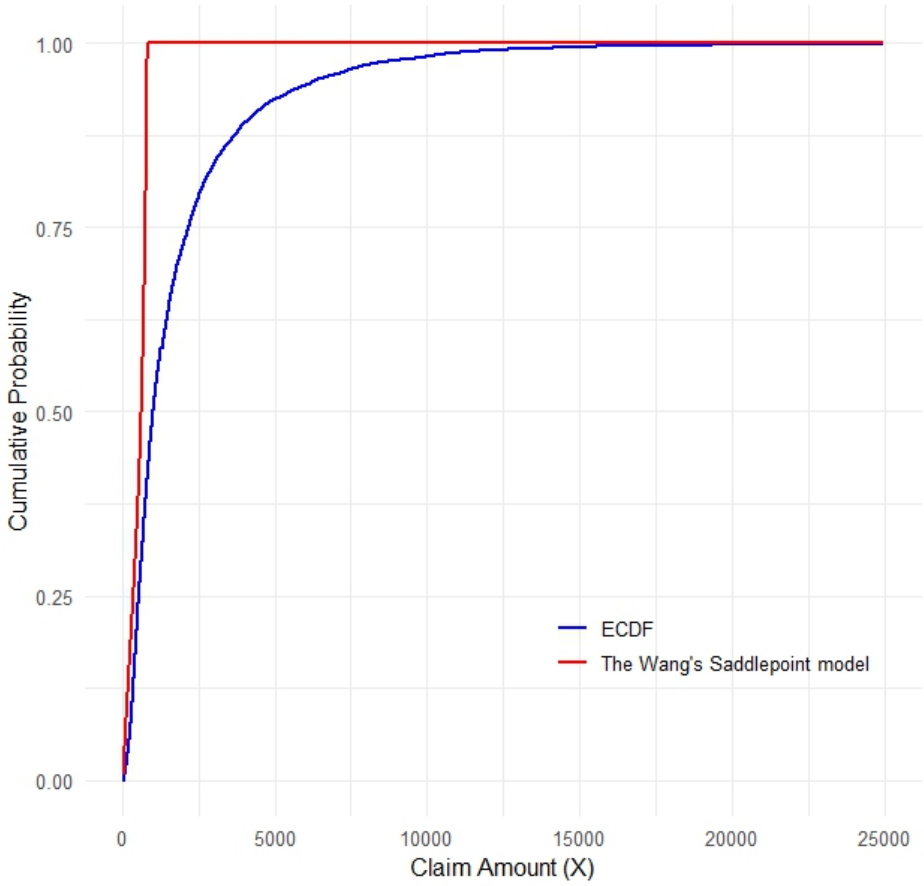


Figure 4.10: The fitting of the distribution derived from the Wang [125] approach

ECDF across the full range of the data, accurately capturing the central distribution and tail behavior of approved claims. By contrast, Figure 10, representing the Wang-derived model, exhibits more pronounced deviations from the ECDF, particularly in the upper tail, where it underestimates the cumulative probability associated with larger claim amounts. These results indicate that the Daniels-derived version of our model provides a more accurate representation of the underlying claim distribution. This considerable accuracy enhances the reliability of risk assessment, reserve estimation, and premium pricing, while reducing the likelihood of underestimating extreme claim sizes. This version of our model could be suitable for practical applications in the insurance sector, where precision in modelling approved claim amounts directly impacts financial stability and decision-making.

4.3 Final Remarks

The classical compound distributions often misalign with approved-claims datasets. For instance, Poisson frequencies place positive mass at zero, while real datasets of *approved* claims exclude zero counts, and the classical exponential severities allow arbitrary values close to zero, whereas deductibles shift payouts above a positive threshold. These mismatches make theoretical and computational efficiency challenging for the aggregate distribution, particularly when truncation and shifting render closed forms inapplicable and make direct numerical convolution computationally expensive. Practitioners, therefore, need a model that naturally incorporates zero truncation and deductibles, ensures theoretical representations for the aggregate law, and supports fast, accurate approximation and estimation at scale.

This chapter presents a practically motivated compound distribution for aggregate *approved* claims, combining a ZTP frequency with SE severities, which aligns with zero truncation in counts and deductible-induced shifts in severities. It establishes the approved-claims count L as follows a ZTPB distribution, derives its PMF, PGF, and moments, and shows that this count model belongs to the Panjer family and, by convolution, to the Schröter family. Additionally, the chapter derives an infinite-series representation for the aggregate approved-claims density, which permits exact evaluation via a tractable summation, and develops closed-form expressions for the aggregate moments and CGF through the composition $M_X(t) = G_L(M_T(t))$ with the associated $K_X(t) = \log M_X(t)$ and higher derivatives needed for asymptotic analysis.

For efficient inference and evaluation, the chapter constructs saddlepoint approximations based on the classical Daniels and Wang approaches for the CDF and PDF of the aggregate, with explicit working formulas for the saddlepoint t_0, curvature $K''_X(t_0)$, and related quantities. It introduces an EM estimation method for λ, q, β, and θ of the model. This yields closed-form updates for β and q, a concave objective in β and θ, and a single implicit equation for λ, resulting in a stable and scalable computational pipeline that integrates exact series evaluation, saddlepoint solvers, and EM iterations. Monte Carlo experiments with bootstrap demonstrate that the proposed saddlepoint estimator t_0 is consistent, with rapidly decreasing bias and SD as sample size increases, and a clear sensitivity ordering in which β and θ exert substantial positive influence, λ a smaller one, and q a negative one. Further comparison demonstrates orders-of-magnitude efficiency gains (lower SD and MSE) of the proposed t_0 estimator over a recent alternative across the sample sizes. Furthermore, an empirical application to the AutoClaims "PAID" data indicates that the Daniels-based saddlepoint model de-

livers the best overall fit in terms of log-likelihood, AIC, BIC, and KS statistics, closely tracking the empirical distribution, including the tails, and underscoring the practical value of the introduced model for actuarial risk measurement, reserving, and pricing under approval and deductible mechanisms.

CHAPTER 5

EXPLORING TRUNCATED DISTRIBUTIONS FROM THE SCHRÖTER FAMILY OF DISCRETE DISTRIBUTIONS

Building on the results of Chapter 4, which highlighted the limitations of classical compound distributions and developed a ZTP–SE–based model with efficient inference tools. This chapter extends this line of inquiry by focusing on the broader Schröter family of discrete distributions. In particular, it explores the truncated versions of these distributions and their convolutions, offering new insights into aggregate claims modeling and parameter estimation under a truncation mechanism.

5.0.1 Motivation

Within the insurance and risk management system, claim data are frequently subject to truncation, whereby only claims above a certain threshold are observed. This situation arises naturally due to policy deductibles, minimum reporting requirements, or regulatory constraints. Truncation alters both the observed claim counts and severities, leading to distortions in aggregate claim distributions if not adequately incorporated into the modeling framework. Classical techniques, such as the Panjer recursive method, have long been used for aggregate claim analysis; however, these methods are typically developed under non-truncated settings, thereby limiting their relevance in real-world applications where truncation is unavoidable.

The Schröter family of discrete distributions offers a natural extension to the Panjer class, providing convolution-based recursions that accommodate zero-inflated and truncated distributions. Despite its theoretical promise, relatively little attention has been given to studying the convolution of truncated members of the Schröter family and addressing the estimation challenges associated with them. This gap narrows the actuarial toolkit available for handling aggregate claims in situations where truncation plays a central role.

Addressing this methodological gap forms the core motivation of this chapter. Specifically, the aim is to explore the convolution of truncated distributions within the Schröter framework, to rigorously establish their probabilistic properties, and to develop practical parameter estimation strategies. By combining theoretical analysis with simulation-based validation, this work seeks to demonstrate the effectiveness of

truncated Schröter distributions in representing aggregate claims under truncation. Ultimately, the chapter contributes both to actuarial theory by extending the class of convolution-based models and to practice by providing computationally feasible tools for real-life claim modeling.

Several studies underscore the importance of truncated probability distributions as more flexible alternatives to their classical counterparts. Unlike standard distributions, truncated models allow flexible support. That is, their support can be adjusted in accordance with the natural range of the dataset, thereby improving their empirical fit and broadening their applications. For example, Prodhani and Shanker [97] introduced the truncated Sujatha distribution and demonstrated its applicability in engineering and medical sciences. By examining its statistical properties through moment-based measures and reliability features, and by validating its performance on two real datasets, they showed that the truncated Sujatha distribution outperforms classical alternatives, including the truncated exponential, truncated Lindley, and exponential models.

Similarly, Singh et al. [108] highlighted the relevance of the truncated Lindley distribution in modeling the strength data of aircraft window glass. Truncated exponential and normal distributions, first examined by Bain and Weeks [11] and further detailed in Johnson et al. [61], have also been widely applied across fields, particularly in modeling truncated lifetime data. More recently, Feng [42] introduced the GEM distribution, providing a comprehensive theoretical framework covering stochastic dynamics, sampling formulas, and the asymptotic behavior of homozygosity. In another development, KRISHNAKUMARI and George [68] constructed a left-truncated distribution based on an Esscher-transformed Laplace distribution with heavy tails, establishing its reliability properties and illustrating its usefulness with failure-time data.

Within the insurance domain, truncated models have proven especially relevant. Gatti and Wüthrich [44] emphasized that claim amounts are often lower-truncated and right-censored due to deductibles and policy limits. Conventional models struggle in such contexts. Revisiting the MBBEFD family of distributions, originally introduced by Bernegger [18] for reinsurance pricing, they demonstrated that this family flexibly accommodates both unimodal skewed and monotonically decreasing densities, thus providing a powerful framework for modeling truncated and censored claims. They further examined how changes in deductibles or policy limits influence model choice and performance.

Beyond insurance, truncated count data have attracted growing attention across diverse fields such as agriculture, engineering, public health, sociology, psychology, and epidemiology. A particularly challenging case is the modeling of zero-truncated counts with inflated frequency at one. In this regard, Wani and Ahmad [126] proposed a one-inflated zero-truncated Poisson distribution, conceptualized as a two-part process. They derived key properties, including generating functions and moments, and developed maximum likelihood estimation procedures. Their analysis, supported by simulation studies and real data applications, demonstrated both the robustness and empirical relevance of the proposed model. Furthermore, the statistical significance of the one-inflation parameter was established using likelihood ratio, Wald, and Rao's score tests.

Taken together, this body of literature demonstrates the flexibility and practical importance of truncated models in applied probability and actuarial science. However, a clear research gap remains in systematically extending the Schröter family to trun-

cated settings and developing associated estimation techniques. This chapter addresses this gap, thereby advancing the theoretical foundations of truncated distributions while providing practically relevant tools for insurance and related fields.

5.1 The convolution of truncated Poisson and Poisson distributions from the Schröter Family

Building on the motivation outlined in the previous subsection, this section develops a new convolutional framework within the Schröter family of distributions. Specifically, we focus on the convolution of a Poisson-distributed random variable with a truncated Poisson-distributed random variable. Such a formulation not only extends the theoretical scope of the Schröter family but also provides a practically relevant model for aggregate claims where truncation naturally arises due to deductibles or reporting thresholds.

We begin by formally defining the distributional setup, where the Poisson and truncated Poisson random variables are assumed to be mutually independent. Under this assumption, the convolution is derived, and its key probabilistic characteristics are established. In particular, we obtain closed-form expressions for the PGF, the mean, and the variance. These results form the foundation for understanding the behavior of the proposed distribution and serve as essential tools for subsequent parameter estimation and simulation.

By embedding the truncated Poisson structure within the Schröter framework, this construction demonstrates how truncation can be systematically incorporated into recursive models of aggregate claims, thereby addressing a methodological gap in the statistical literature.

5.1.1 The convolution of Poisson and truncated Poisson distributions

The Poisson distribution has been shown to be a member of the Panjer and Schröter family of distributions as studied in [88] and Schröter [103].

Let K and K_{tr} be two independent random variables and $P(K = k)$ be the PMF of the Poisson distribution.

Similarly, let

$$P(K_{tr} = k) = \frac{P(K = k)}{1 - P(K < 1)} \tag{5.1}$$

be the PMF of the Poisson random variable truncated at zero and

$$Z = K + K_{tr} \tag{5.2}$$

be the convolution of the Poisson and truncated Poisson random variables.

5.1.2 Derivation of the PMF of $Z = K + K_{tr}$

Recall that K and K_{tr} are independent random variables and

$$K \sim \text{Poisson}(\lambda_1)$$

such that

$$P(K = i) = \frac{e^{-\lambda_1}\lambda_1^i}{i!},\ i = 0, 1, 2, \ldots, \tag{5.3}$$

and K_{tr} is the Poisson(λ_2) distribution truncated at zero (so its support is $\{1, 2, \ldots\}$). The PMF of the truncated Poisson is:

$$P(K_{tr} = j) = \frac{P_{\text{Pois}(\lambda_2)}(j)}{1 - P_{\text{Pois}(\lambda_2)}(0)} = \frac{e^{-\lambda_2}\lambda_2^j}{j!\,(1 - e^{-\lambda_2})}, \qquad j = 1, 2, \ldots \tag{5.4}$$

because $1 - P(K < 1) = 1 - P(K = 0) = 1 - e^{-\lambda_2}$.

Recall that we defined

$$Z = K + K_{tr}.$$

We derive $P(Z = k)$ for integer k.

Support. Since $K \geq 0$ and $K_{tr} \geq 1$, we have $Z \geq 1$. Hence $P(Z = 0) = 0$ and we consider $k = 1, 2, \ldots$.

Convolution. Using independence and the discrete convolution formula,

$$P(Z = k) = \sum_{i=0}^{k} P(K = i)\, P(K_{tr} = k - i). \tag{5.5}$$

However, $P(K_{tr} = k - i) > 0$ only when $k - i \geq 1$, i.e. $i \leq k - 1$. Therefore, the sum reduces to

$$P(Z = k) = \sum_{i=0}^{k-1} P(K = i)\, P(K_{tr} = k - i), \qquad k \geq 1. \tag{5.6}$$

Substitution. Substituting the Poisson and truncated-Poisson PMFs defined in equations (5.3) and (5.4), we have:

$$\begin{aligned} P(Z = k) &= \sum_{i=0}^{k-1} \left(\frac{e^{-\lambda_1}\lambda_1^i}{i!}\right)\left(\frac{e^{-\lambda_2}\lambda_2^{k-i}}{(k-i)!\,(1 - e^{-\lambda_2})}\right) \\ &= \frac{e^{-(\lambda_1+\lambda_2)}}{1 - e^{-\lambda_2}} \sum_{i=0}^{k-1} \frac{\lambda_1^i\,\lambda_2^{k-i}}{i!\,(k-i)!}. \end{aligned}$$

Combine factorials and use the binomial identity. Notice that:

$$\frac{\lambda_1^i \lambda_2^{k-i}}{i!\,(k-i)!} = \frac{1}{k!}\binom{k}{i}\lambda_1^i \lambda_2^{k-i}, \tag{5.7}$$

because $\binom{k}{i} = \dfrac{k!}{i!(k-i)!}$.

Thus,

$$\sum_{i=0}^{k-1} \frac{\lambda_1^i \lambda_2^{k-i}}{i!\,(k-i)!} = \frac{1}{k!} \sum_{i=0}^{k-1} \binom{k}{i} \lambda_1^i \lambda_2^{k-i}. \tag{5.8}$$

Applying the binomial theorem to the full sum $\sum_{i=0}^{k} \binom{k}{i} \lambda_1^i \lambda_2^{k-i} = (\lambda_1 + \lambda_2)^k$. The sum from $i = 0$ to $k - 1$ therefore equals the full sum minus the $i = k$ term:

$$\sum_{i=0}^{k-1} \binom{k}{i} \lambda_1^i \lambda_2^{k-i} = (\lambda_1 + \lambda_2)^k - \binom{k}{k} \lambda_1^k \lambda_2^0 = (\lambda_1 + \lambda_2)^k - \lambda_1^k. \tag{5.9}$$

Substituting back, for $k = 1, 2, \ldots$,

$$P(Z = k) \;=\; \frac{e^{-(\lambda_1+\lambda_2)}}{k!\,(1 - e^{-\lambda_2})} \Big[(\lambda_1 + \lambda_2)^k - \lambda_1^k \Big], \qquad k = 1, 2, \ldots, \tag{5.10}$$

such that

$$\begin{aligned} \sum_{k=1}^{\infty} P(Z = k) &= \frac{e^{-(\lambda_1+\lambda_2)}}{1 - e^{-\lambda_2}} \sum_{k=1}^{\infty} \frac{(\lambda_1 + \lambda_2)^k - \lambda_1^k}{k!} \\ &= \frac{e^{-(\lambda_1+\lambda_2)}}{1 - e^{-\lambda_2}} \big[(e^{\lambda_1+\lambda_2} - 1) - (e^{\lambda_1} - 1) \big] \\ &= \frac{e^{-(\lambda_1+\lambda_2)} \big(e^{\lambda_1+\lambda_2} - e^{\lambda_1} \big)}{1 - e^{-\lambda_2}} = \frac{1 - e^{-\lambda_2}}{1 - e^{-\lambda_2}} = 1; \end{aligned}$$

thus, the distribution is a PMF.

Therefore the PMF of $Z = K + K_{tr}$ is

$$P(Z = k) = \frac{e^{-(\lambda_1+\lambda_2)}}{k!\,(1 - e^{-\lambda_2})} \Big[(\lambda_1 + \lambda_2)^k - \lambda_1^k \Big], \qquad k = 1, 2, \ldots \text{‘}$$

and $P(Z = 0) = 0$.

5.1.3 The PGF of Z

Let

$$X \sim \text{Pois}(\lambda_1), \qquad Y \sim \text{Pois}(\lambda_2),$$

and define the ZTP random variable:

$$\hat{Y} \stackrel{d}{=} Y \mid (Y \geq 1). \tag{5.11}$$

Assume X and Y (hence X and $\hat{Y}$) are independent. We study the PGF of

$$Z \stackrel{d}{=} X + \hat{Y}. \tag{5.12}$$

1. PGF of the Poisson components. The PGF of X is

$$G_X(s) = E[s^X] = e^{\lambda_1(s-1)}. \tag{5.13}$$

The PGF of Y (untruncated Poisson) is

$$G_Y(s) = E[s^Y] = e^{\lambda_2(s-1)}, \tag{5.14}$$

and $P(Y = 0) = e^{-\lambda_2}$.

2. PGF of the ZTP $\hat{Y}$. By definition of conditional PGF (conditioning on the event $Y \geq 1$),

$$G_{\hat{Y}}(s) = E[s^{\hat{Y}}] = E\big[s^Y \mid Y \geq 1\big] = \frac{E[s^Y \mathbf{1}_{\{Y\geq 1\}}]}{P(Y \geq 1)} = \frac{E[s^Y] - P(Y = 0) \cdot s^0}{1 - P(Y = 0)}. \tag{5.15}$$

Since $s^0 = 1$ and $P(Y = 0) = e^{-\lambda_2}$, this becomes

$$G_{\hat{Y}}(s) = \frac{G_Y(s) - e^{-\lambda_2}}{1 - e^{-\lambda_2}} = \frac{e^{\lambda_2(s-1)} - e^{-\lambda_2}}{1 - e^{-\lambda_2}} \ . \tag{5.16}$$

3. PGF of $Z = X + \hat{Y}$. Independence implies the PGF of the sum is the product of PGFs:

$$\begin{aligned} G_Z(s) &= G_X(s)\, G_{\hat{Y}}(s) \\ &= e^{\lambda_1(s-1)}\, \frac{e^{\lambda_2(s-1)} - e^{-\lambda_2}}{1 - e^{-\lambda_2}} \\ &= \frac{e^{((\lambda_1+\lambda_2)(s-1))} - e^{(\lambda_1(s-1)-\lambda_2)}}{1 - e^{-\lambda_2}}. \end{aligned}$$

(We may keep the factorized form $G_Z(s) = e^{\lambda_1(s-1)} \dfrac{e^{\lambda_2(s-1)} - e^{-\lambda_2}}{1 - e^{-\lambda_2}}$ for differentiation convenience.)

5.1.4 The mean and variance of Z via the PGF approach

First derivative and the mean $E[Z]$. Recall that for a PGF $G(s)$, we have $G'(1) = E[Z]$.

Let us differentiate $G_Z(s) = A(s)B(s)$ by setting $A(s) = e^{\lambda_1(s-1)}$ and $B(s) = \dfrac{e^{\lambda_2(s-1)} - e^{-\lambda_2}}{1 - e^{-\lambda_2}}$.

We present the derivatives as:

$$A'(s) = \lambda_1 e^{\lambda_1(s-1)} = \lambda_1 A(s), \qquad B'(s) = \frac{\lambda_2 e^{\lambda_2(s-1)}}{1 - e^{-\lambda_2}}.$$

Thus, using the product rule,

$$G'_Z(s) = A'(s)B(s) + A(s)B'(s),$$

and evaluate at $s = 1$.

Note $A(1) = 1$, $B(1) = 1$; therefore,

$$\begin{aligned} G'_Z(1) &= A'(1)B(1) + A(1)B'(1) \\ &= \lambda_1 \cdot 1 + 1 \cdot \frac{\lambda_2}{1 - e^{-\lambda_2}} \end{aligned}$$

$$E[Z] = G'_Z(1) = \lambda_1 + \frac{\lambda_2}{1 - e^{-\lambda_2}}. \tag{5.17}$$

As a direct check, we have:

$E[Z] = E[X] + E[\hat{Y}] = \lambda_1 + E[Y \mid Y \geq 1]$ and $E[Y \mid Y \geq 1] = \dfrac{E[Y] - 0 \cdot P(Y = 0)}{1 - P(Y = 0)} = \dfrac{\lambda_2}{1 - e^{-\lambda_2}}$

Thus, the same result follows.

The Second derivative and $\mathrm{Var}(Z)$**.** Recall that $G''(1) = E[Z(Z-1)]$ and therefore,

$$\mathrm{Var}(Z) = G''_Z(1) + G'_Z(1) - \left(G'_Z(1)\right)^2. \tag{5.18}$$

Let us differentiate again using

$A''(s) = \lambda_1^2 e^{\lambda_1(s-1)} = \lambda_1^2 A(s)$ and $B''(s) = \dfrac{\lambda_2^2 e^{\lambda_2(s-1)}}{1 - e^{-\lambda_2}}$, applying the product rule (twice) gives:

$$G''_Z(s) = A''(s)B(s) + 2A'(s)B'(s) + A(s)B''(s). \tag{5.19}$$

By evaluating at $s = 1$ and using the fact that $A(1) = B(1) = 1$, $A'(1) = \lambda_1$, $A''(1) = \lambda_1^2$, $B'(1) = \dfrac{\lambda_2}{1 - e^{-\lambda_2}}$ and $B''(1) = \dfrac{\lambda_2^2}{1 - e^{-\lambda_2}}$:

$$\begin{aligned} G''_Z(1) &= \lambda_1^2 \cdot 1 + 2\lambda_1 \cdot \frac{\lambda_2}{1 - e^{-\lambda_2}} + 1 \cdot \frac{\lambda_2^2}{1 - e^{-\lambda_2}} \\ &= \lambda_1^2 + \frac{2\lambda_1\lambda_2 + \lambda_2^2}{1 - e^{-\lambda_2}}. \end{aligned}$$

Now, let us compute the variance and by definition, we have that:

$$G''_Z(1) + G'_Z(1) - \left(G'_Z(1)\right)^2 \tag{5.20}$$

$$\begin{aligned} \mathrm{Var}(Z) &= G''_Z(1) + G'_Z(1) - \left(G'_Z(1)\right)^2 \\ &= \left[\lambda_1^2 + \frac{2\lambda_1\lambda_2 + \lambda_2^2}{1 - e^{-\lambda_2}}\right] + \left[\lambda_1 + \frac{\lambda_2}{1 - e^{-\lambda_2}}\right] - \left[\lambda_1 + \frac{\lambda_2}{1 - e^{-\lambda_2}}\right]^2. \end{aligned}$$

By expanding and simplifying, the λ_1^2 terms cancel and the cross terms $2\lambda_1\lambda_2/(1-e^{-\lambda_2})$ cancel, leaving

$$\begin{aligned}\mathrm{Var}(Z) &= \lambda_1 + \left(\frac{\lambda_2^2+\lambda_2}{1-e^{-\lambda_2}} - \frac{\lambda_2^2}{(1-e^{-\lambda_2})^2}\right)\\ &= \lambda_1 + \frac{(\lambda_2^2+\lambda_2)(1-e^{-\lambda_2}) - \lambda_2^2}{(1-e^{-\lambda_2})^2}\\ &= \lambda_1 + \frac{\lambda_2 - (\lambda_2+\lambda_2^2)e^{-\lambda_2}}{(1-e^{-\lambda_2})^2}\\ &= \lambda_1 + \frac{\lambda_2\big[1-(1+\lambda_2)e^{-\lambda_2}\big]}{(1-e^{-\lambda_2})^2}.\end{aligned}$$

Hence,

$$\mathrm{Var}(Z) = \lambda_1 \; + \; \frac{\lambda_2\big[1-(1+\lambda_2)e^{-\lambda_2}\big]}{(1-e^{-\lambda_2})^2}. \tag{5.21}$$

We can easily notice that λ_2 is small, the truncation has a strong effect; that is, the $E[\hat{Y}]$ and $\mathrm{Var}(\hat{Y})$ are inflated compared to the untruncated Poisson moments. Additionally, if there were no truncation (i.e. formally $P(Y=0)=0$), the formulae reduce to those of a sum of independent Poisson(λ_1) and Poisson(λ_2) variables (so $E[Z]=\lambda_1+\lambda_2$ and $\mathrm{Var}(Z)=\lambda_1+\lambda_2$).

5.2 The Parameter Estimation

This section addresses estimation of the parameters λ_1 and λ_2 for the distribution

$$P(Z=k) = \frac{e^{-(\lambda_1+\lambda_2)}}{k!\,(1-e^{-\lambda_2})}\Big[(\lambda_1+\lambda_2)^k - \lambda_1^k\Big], \qquad k=1,2,\ldots$$

A direct maximum likelihood approach is natural, but because the PMF is nonlinear in λ_1, λ_2, the score equations do not admit closed form solutions, and numerical maximization may be unstable for small samples. Hence, we present a practical method-of-moments procedure that uses the mean, variance, and the ratio $P(Z=2)/P(Z=1)$.

Preliminaries and notation. Throughout this chapter, we set

$$y = \frac{1}{1-e^{-\lambda_2}} \quad\Longleftrightarrow\quad 1-e^{-\lambda_2} = \frac{1}{y}, \qquad e^{-\lambda_2} = 1-\frac{1}{y} = \frac{y-1}{y}.$$

Observe that $y>1$ for all $\lambda_2>0$, and as $\lambda_2\to\infty$, $y\to 1$.

From the PGF derivations, we have the closed-form moments

$$\mathbb{E}[Z] = \lambda_1 + \lambda_2 y, \tag{5.22}$$

$$\mathrm{Var}(Z) = \lambda_1 + \frac{\lambda_2\big[1-(1+\lambda_2)e^{-\lambda_2}\big]}{(1-e^{-\lambda_2})^2}. \tag{5.23}$$

Let us define auxiliary quantities K and L and Introduce

$$K = \mathrm{Var}(Z) - \mathbb{E}[Z], \qquad L = \mathbb{E}[Z] - \frac{P(Z=2)}{P(Z=1)}.$$

We will express K and L in terms of λ_2 and y and hence eliminate λ_1.

We simplified $K = \mathrm{Var}(Z) - \mathbb{E}[Z]$ by starting from equations (5.22) and (5.23):

$$\begin{aligned} K &= \left[\lambda_1 + \frac{\lambda_2\big[1-(1+\lambda_2)e^{-\lambda_2}\big]}{(1-e^{-\lambda_2})^2}\right] - [\lambda_1 + \lambda_2 y] \\ &= \lambda_2\left[\frac{1-(1+\lambda_2)e^{-\lambda_2}}{(1-e^{-\lambda_2})^2} - y\right]. \end{aligned}$$

Now, by substituting $1 - e^{-\lambda_2} = 1/y$ and $e^{-\lambda_2} = (y-1)/y$, we have:

$$\begin{aligned} \frac{1-(1+\lambda_2)e^{-\lambda_2}}{(1-e^{-\lambda_2})^2} &= \big(1-(1+\lambda_2)\tfrac{y-1}{y}\big)\, y^2 \\ &= \big(1-(1+\lambda_2)\frac{y-1}{y}\big)\, y^2 = \big(\frac{y-(1+\lambda_2)(y-1)}{y}\big)\, y^2 \\ &= \big(y-(1+\lambda_2)(y-1)\big)\, y. \end{aligned}$$

Therefore,

$$\begin{aligned} K &= \lambda_2\big[\big(y-(1+\lambda_2)(y-1)\big)y - y\big] \\ &= \lambda_2\big[y^2 - (1+\lambda_2)y(y-1) - y\big] \\ &= \lambda_2\big[y(y-1) - (1+\lambda_2)y(y-1)\big] \\ &= \lambda_2\big[-\lambda_2\, y(y-1)\big] \\ &= -\lambda_2^2\, y(y-1). \end{aligned}$$

Hence,

$$K = \mathrm{Var}(Z) - \mathbb{E}[Z] = -\lambda_2^2\, y(y-1) < 0. \tag{5.24}$$

Next, we compute the ratio $P(Z=2)/P(Z=1)$ and simplify L, using the PMF,

$$P(Z=1) = \frac{e^{-(\lambda_1+\lambda_2)}}{1-e^{-\lambda_2}}\,\lambda_2$$

$$P(Z=2) = \frac{e^{-(\lambda_1+\lambda_2)}}{2(1-e^{-\lambda_2})}\,\big[(\lambda_1+\lambda_2)^2 - \lambda_1^2\big] = \frac{e^{-(\lambda_1+\lambda_2)}}{2(1-e^{-\lambda_2})}\,\lambda_2(2\lambda_1+\lambda_2). \tag{5.25}$$

Now, divide to obtain the ratio $\frac{P(Z=2)}{P(Z=1)}$ and we have:

$$\frac{P(Z=2)}{P(Z=1)} = \frac{\lambda_2(2\lambda_1+\lambda_2)/(2(1-e^{-\lambda_2}))}{\lambda_2/(1-e^{-\lambda_2})} = \frac{2\lambda_1+\lambda_2}{2} = \lambda_1 + \frac{\lambda_2}{2}. \tag{5.26}$$

Hence,

$$L = \mathbb{E}[Z] - \frac{P(Z=2)}{P(Z=1)} = \big(\lambda_1+\lambda_2 y\big) - \big(\lambda_1 + \tfrac{\lambda_2}{2}\big) = \lambda_2\Big(y - \tfrac{1}{2}\Big).$$

So

$$L = \lambda_2\Big(y - \tfrac{1}{2}\Big). \tag{5.27}$$

Next, we defined the scale-free ratio as $v = K/L^2$ and solve for y as:

$$v = \frac{K}{L^2} = \frac{-\lambda_2^2 y(y-1)}{\lambda_2^2 (y-\frac{1}{2})^2} = -\frac{y(y-1)}{(y-\frac{1}{2})^2}.$$

Thus, $v \in (-1, 0)$ for admissible $y > 1$.

We solve the equation:

$$-\frac{y(y-1)}{(y-\frac{1}{2})^2} = v. \tag{5.28}$$

We multiply both sides by $(y-\frac{1}{2})^2$ and rearrange, we have that:

$$-(y^2 - y) = v\,(y-\tfrac{1}{2})^2 \quad \Longrightarrow \quad (1+v)(y^2-y) + \frac{v}{4} = 0.$$

This expression is quadratic in y. We have that:

$$(1+v)\,y^2 - (1+v)\,y + \frac{v}{4} = 0. \tag{5.29}$$

By applying the quadratic formula. The discriminant simplifies as:

$$\Delta = (1+v)^2 - 4(1+v)\frac{v}{4} = (1+v)\big[(1+v) - v\big] = 1 + v.$$

We write $1 + v = (\sqrt{(1+v)})^2$ and we have that:

$$y = \frac{1}{2}\left(\frac{1+v}{1+v} + \frac{1+v}{\left(\sqrt{(1+v)}\right)^2}\right).$$

Hence, we have:

$$y = \frac{(1+v) \pm \sqrt{1+v}}{2(1+v)} = \frac{1}{2}\Big(1 \pm \frac{1}{\sqrt{1+v}}\Big).$$

Here, we consider only the positive root (i.e., the one > 1):

$$y = \frac{1}{2}\Big(1 + \frac{1}{\sqrt{1+v}}\Big). \tag{5.30}$$

Next, we recover λ_2 from y by using the fact that $y = 1/(1 - e^{-\lambda_2})$, we have:

$$1 - \frac{1}{y} = e^{-\lambda_2} \quad \Longrightarrow \quad \lambda_2 = -\log\Big(1 - \frac{1}{y}\Big).$$

The recovery of λ_1: Note that once λ_2 (hence y) is known, λ_1 follows from the mean or from the ratio P_2/P_1:

$$\lambda_1 = \mathbb{E}[Z] - \lambda_2 y \quad \text{or} \quad \lambda_1 = \frac{P(Z=2)}{P(Z=1)} - \frac{\lambda_2}{2}.$$

5.2.1 Sample (method-of-moments) estimators

Given an i.i.d. sample $Z_1, \ldots, Z_n$ from the model, the sample summaries are computed as:

$$\overline{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i, \qquad s^2 = \frac{1}{n}\sum_{i=1}^{n}(Z_i - \overline{Z})^2,$$

and the empirical one- and two-count sample probabilities are computed as:

$$\widehat{p}_1 = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{Z_i = 1\}, \qquad \widehat{p}_2 = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{Z_i = 2\}.$$

These form the sample analogues:

$$\widehat{K} = s^2 - \overline{Z}, \qquad \widehat{L} = \overline{Z} - \frac{\widehat{p}_2}{\widehat{p}_1}, \qquad \widehat{v} = \frac{\widehat{K}}{\widehat{L}^2}.$$

Next, we proceed as in the population case:

$$\widehat{y} = \frac{1}{2}\Big(1 + \frac{1}{\sqrt{1+\widehat{v}}}\Big), \qquad \widehat{\lambda}_2 = -\log\Big(1 - \frac{1}{\widehat{y}}\Big), \qquad \widehat{\lambda}_1 = \overline{Z} - \widehat{\lambda}_2\,\widehat{y}.$$

Remark: It is worth noting that our approach for the parameter estimation requires $\widehat{p}_1 > 0$ (otherwise $\widehat{p}_2/\widehat{p}_1$ is undefined) and $\widehat{v} \in (-1, 0)$. If these conditions are not met, the introduced method-of-moments estimator above is not applicable. Thus, an alternative estimation approach may be necessary.

Furthermore, the choice of population vs. sample variance denominator (here we used n for s^2) is consistent with other parts of our results. Although using the unbiased sample variance with a denominator $n-1$ is possible, it requires adapting the moment equations accordingly.

The parameter estimates here could serve as starting values for numerical optimization of the log-likelihood to obtain asymptotically efficient estimates.

5.2.2 Numerical Applications

This section presents the numerical applications of the convolution of the truncated Poisson and Poisson distributions, as defined in equation (5.10). The focus is on evaluating the performance of the proposed parameter estimation approach through simulation studies.

5.3 Simulation Study

5.3.1 Simulation Procedure for the Probability Mass Function

To simulate from $P(Z = k)$, probabilities are first constructed for $k = 1, \ldots, k_{\max}$ using the log formulation

$$\log\big(P(Z = k)\big)$$

to prevent numerical underflow and improve stability when computing very small probabilities involving factorials and exponentials. The probabilities are then normalized and the sample Z is generated. From the simulated data, the method of moments

estimates for the parameters λ_1 and λ_2 are obtained by computing the sample mean, variance, and empirical probabilities for $Z = 1$ and $Z = 2$, and solving the resulting moment equations. Theoretical moments are derived from the truncated distribution structure to allow comparison with the empirical estimates.

To examine the effectiveness and accuracy of the proposed moment-based parameter estimation method outlined in Section 3.2, we conducted a simulation experiment. The goal was to determine how well the procedure recovers the true parameter values under controlled conditions. We began by setting the true parameter values as:

$$\lambda_1 = 1.5 \quad \text{and} \quad \lambda_2 = 2.0.$$

Using these parameters, we generated a synthetic dataset consisting of 100,000 random observations from the proposed probability mass function and repeated the procedure 10,000 times. This large sample size ensures statistical stability, allowing for meaningful inference. The generated data were then used to estimate the model parameters based on the moment-based approach described in Section 5.2.

After applying the estimation procedure, we obtained the theoretical and empirical means and variance for the generated data as 3.813 and 3.814, and 3.089 and 3.096, respectively. The estimated values for the parameters were:

$$\hat{\lambda}_1 = 1.44 \quad \text{and} \quad \hat{\lambda}_2 = 2.07.$$

Although the proposed approach is computationally appealing, its effectiveness is highly sensitive to the structure of the moment equations and the stability of the solution. The simulation findings underscore the importance of integrating constraint-aware mechanisms in moment-based procedures to ensure valid and interpretable estimates.

The parameter estimates obtained based on the simulated data indicate that the estimated values of the parameters are close to the true generating initial values. This demonstrates that the employed moment-based approach for the parameter estimates provides a consistent and accurate recovery of the parameters when applied to large sample sizes. Table 5.1 presents the descriptive statistics of the replicated estimates. From the table, the means of λ_1 and λ_2 confirm the unbiasedness of the estimation procedure. Additionally, the relatively small standard deviations for λ_1 and λ_2 indicate stable estimation across the 10,000 replications, while the minimum and maximum values show the natural variability of the estimates under repeated sampling.

Table 5.1: The Descriptive Statistics of the Estimates

	$\hat{\lambda}_1$	$\hat{\lambda}_2$
Mean	1.50141	1.99771
Standard Deviation	0.06566	0.08271
Minimum	1.21361	1.65037
Maximum	1.77657	2.34789

Figure 5.1 further illustrates the distribution of the estimates through the boxplots, indicating that both λ_1 and λ_2 are tightly centered around their true values with symmetric distributions and only a small proportion of mild outliers, suggesting the efficiency and robustness of the estimator.

Furthermore, Figure 5.2 compares the observed and expected frequencies under the fitted PMF using the simulated data. The close alignment between the bars demonstrates a good fit of the introduced PMF to the generated data, indicating that the model is capable of capturing the underlying distributional characteristics of the data. These results provide significant empirical evidence that the introduced PMF, combined with the moment-based estimation approach, performs well in accurately estimating parameters and reproducing the distributional features of the data under repeated large-sample-size simulation scenarios.

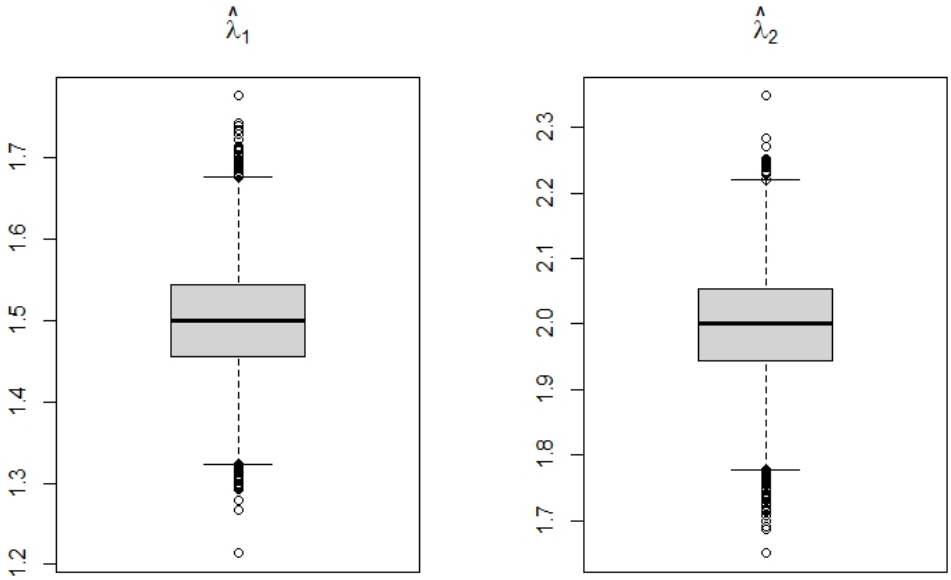


Figure 5.1: Boxplots of the Estimates

5.4 Final Remarks

This chapter advances the agenda opened in Chapter 4 by moving from a single ZTP–SE model to the broader Schröter family under truncation, with a focus on convolution structures that arise in aggregate claims. The work delivers three main outcomes that align with the stated aim of developing theory and practical inference for truncated Schröter distributions.

First, we construct the convolution of a Poisson and a zero–truncated Poisson random variable within the Schröter framework and derive a closed–form PMF for the sum. Using the PGF approach, we establish concise expressions for the mean and variance, prove normalization and support, and show explicitly how truncation inflates the moments relative to the untruncated case. This provides a rigorous and reusable building block for aggregate claims subject to deductibles and reporting thresholds.

Second, we develop a practical method–of–moments estimator tailored to the truncated setting. The approach introduces the quantities $K = \mathrm{Var}(Z) - \mathbb{E}[Z]$ and $L = \mathbb{E}[Z] - P(Z = 2)/P(Z = 1)$, leading to a scale–free ratio $v = K/L^2 \in (-1, 0)$, and yields

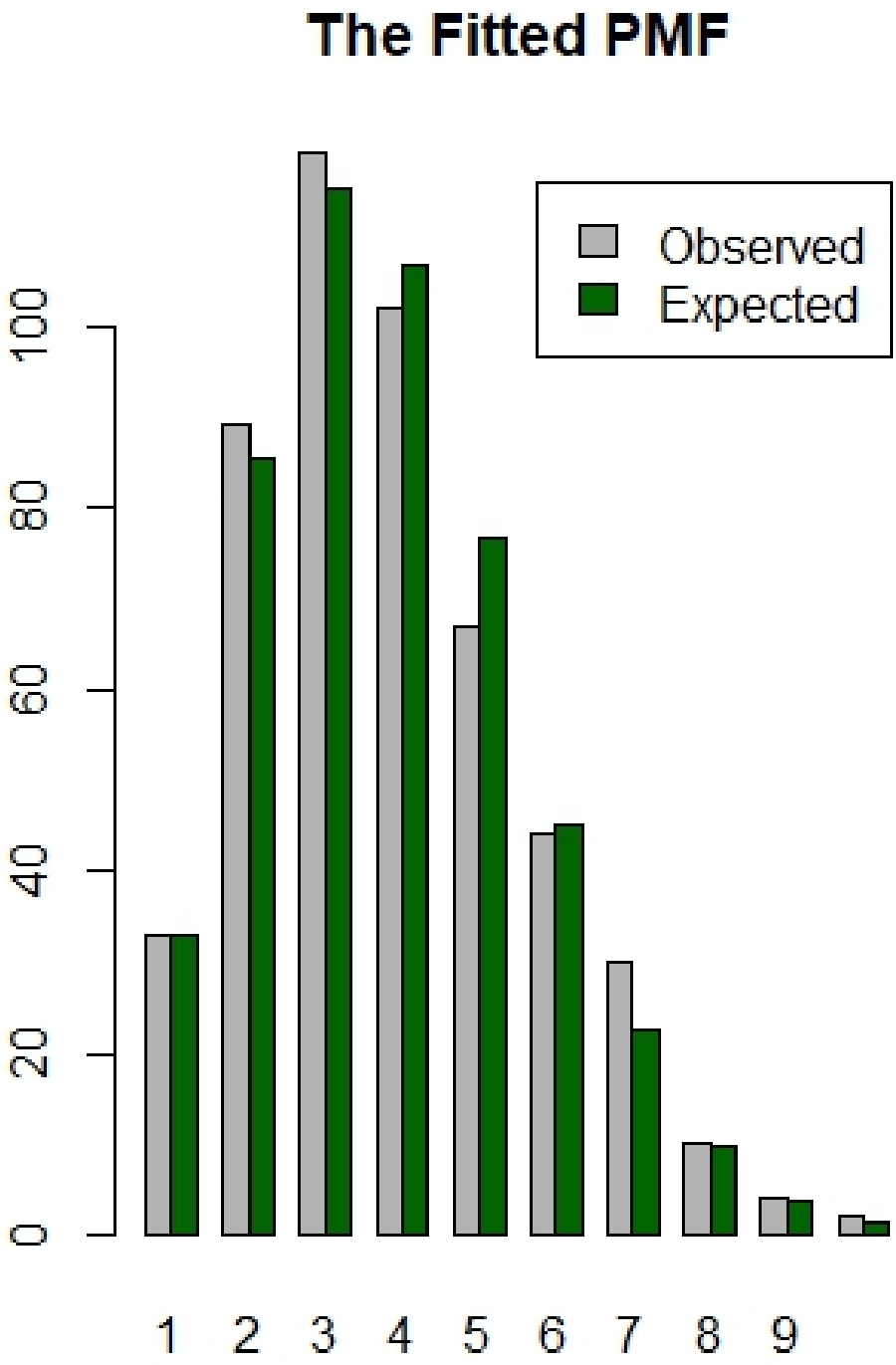


Figure 5.2: The Graph of the Fitted PMF model to the Simulated Data

closed–form recovery of the auxiliary parameter y, the truncation intensity λ_2, and subsequently λ_1. We state the feasibility conditions, provide sample analogues, and note the usefulness of the resulting estimates as stable starting values for likelihood–based refinement. This directly addresses the practical difficulty of solving nonlinear score equations in the truncated Schröter setting.

Finally, a large–scale simulation study based on a numerically stable log–PMF sampling routine confirms that the proposed estimators are accurate and stable for realistic sample sizes. Descriptive summaries and graphical diagnostics (boxplots and fitted–PMF comparisons) indicate a tight concentration around the true parameters and a good fidelity of the fitted distribution to the generated data. These findings demonstrate that the model captures significant distributional features induced by truncation and is therefore well-suited for actuarial applications where deductibles and reporting limits are intrinsic.

Together, these results extend the Schröter family to a practically important truncated regime, providing closed–form moment identities and a computationally light estimation scheme, and provide evidence of performance under simulation. The chapter provides theoretical clarity and implementable tools for modeling aggregate claims under truncation. In practice, the proposed framework can be embedded in convolution–based recursions for portfolio risk, used to stress–test deductible strategies, and serve as a gateway to richer generalizations (such as other truncated Schröter members). As such, this chapter strengthens the actuarial toolkit and creates a coherent bridge from the model in Chapter 4 to general truncation–aware compound models in subsequent chapters.

THESIS SUMMARY

In this thesis, we explored the recursive approach for computing aggregate claim amounts in the collective risk model, focusing specifically on the Schröter recursive family of discrete distributions and their associated theoretical and practical challenges. These challenges include identifying the feasible region of the parameters of the Schröter family, parameter estimation, derivation of new members of the family, development of truncated versions, modeling of approved claim amounts, and the convolution of truncated and non-truncated members of the Schröter family of discrete distributions.

First, we presented the background and theoretical foundation of the collective risk model and advanced the recursive approach by examining the Panjer and Schröter recursion techniques for computing aggregate claim amounts when the number of claims is large. We examined the characterization of the Panjer family of discrete distributions through the $R_0(a, b)$ and $R_1(a, b)$ classes and analyzed the Schröter recursive approach and its characterization. The study further explored the convolution of Poisson and $R_0(a, b)$ classes, introduced members of the Schröter $R_0(a, b, c)$ and $R_1(a, b, c)$ families, and presented new results on the derivation of an additional Schröter distribution family member. Furthermore, we investigated the feasible regions of parameters for both finite and infinite cases, proposed a novel parameter estimation method, and extended the Schröter model to a regression framework suitable for real-life insurance applications. Theoretical developments were validated using both simulated and empirical data.

Subsequently, we addressed the problem of capturing truncated claim amounts by introducing the *truncated Schröter recursive formula* and its corresponding recursion algorithm. The numerical evaluation of this new algorithm was conducted using simulated and real-world datasets. Moreover, we presented results on the generalized Schröter $R_t(a, b, c)$ family and derived the associated class of distributions.

We then investigated the distribution of aggregate approved claim amounts within the collective risk model. In particular, we derived the infinite-sum representation of the distribution by combining the zero-truncated Poisson and shifted exponential distributions. We also examined the distribution of unapproved claim amounts, derived the approval probability, and explored the mathematical properties of both the approved and aggregate approved claim amount distributions. To approximate the distribution of aggregate approved claim amounts, we employed the saddlepoint approximation methods proposed by Daniels [27] and Wang and Sobrero [124]. The proposed model

was further illustrated using simulation studies and real-life data.

Given the research gap in systematically extending the Schröter family of discrete distributions to truncated settings and developing associated estimation methods, we explored the convolution of truncated Poisson and standard Poisson distributions from the Schröter class. We derived the PMF of $Z = K + K_{\text{tr}}$, examined its statistical properties, and applied the method of moments for parameter estimation. The performance of the proposed estimation method was validated through simulation experiments.

In conclusion, this thesis includes material from the following papers and conference book of abstract:

1. Agu, F. I., Mačutek, J., & Szűcs, G. (2023). A Simple Estimation of Parameters for Discrete Distributions from the Schröter Family. *Statistika: Statistics & Economy Journal*, 103(2).

2. Agu, F. I., & Mačutek, J. (2024). Some Extensions of the Schröter Distribution Family. *PROBASTAT 2024*, 6.

3. Agu, F. I. (2025). The Truncated Schröter Recursive Algorithm for Computation of Aggregate Claim Amounts. *Statistics in Transition New Series*, (Accepted for publication in vol 26 No 4/2025- December issue 2025).

4. Agu, F. I. (2025). The Schröter Recursive Formula: The Parameters Feasible Region, Asymptotic Analysis, and Its Systematic Structure for Explanatory Number of Claim Amount Variables. (Manuscript submitted for publication). *Scandinavian Actuarial Journal*.

5. Agu, F. I., Mačutek, J., & Szűcs, G. (2025). Exploring Truncated Distributions from the Schröter Family of Discrete Distributions. (Manuscript in preparation).

6. Agu, F. I., & Wald, H. (2025). Modeling Insurance Claim Data: An Infinite-Sum Representation of Aggregate Approved Claim Amount Distribution and the Saddlepoint Model. (Manuscript in preparation).

Bibliography

[1] Adelson. Compound poisson distributions. *Journal of the Operational Research Society*, 17(1):73–75, 1966.

[2] Friday I. Agu. The truncated schröter recursive algorithm for computation of aggregate claim amounts. *Statistics in Transition New Series*, 2025. Forthcoming; accepted.

[3] Friday I. Agu and Helmut Wald. Modeling insurance claim data: An infinite-sum representation of aggregate approved claim amount distribution and the saddlepoint model. Manuscript in preparation, 2025.

[4] Friday I. Agu, Ján Mačutek, and Gábor Szűcs. A simple estimation of parameters for discrete distributions from the schröter family. *Statistika: Statistics and Economy Journal*, 103(2):246–251, 2023. doi: 10.54694/stat.2022.42.

[5] Friday I. Agu, Ján Mačutek, and Gábor Szűcs. Exploring truncated distributions from the schröter family of discrete distributions. Manuscript in preparation, 2025.

[6] S. Ejaz Ahmed, Claudia Castro-Kuriss, Esteban Flores, Víctor Leiva, and Antonio Sanhueza. A truncated version of the birnbaum–saunders distribution with an application in financial risk. *Pakistan Journal of Statistics*, 26(1), 2010.

[7] Amal Alhejaili and Ateq Alghamedi. Approximation methods for the bivariate compound truncated poisson gamma distribution. *Pakistan Journal of Statistics and Operation Research*, pages 285–299, 2024.

[8] Amal D Alhejaili and Ateq Ahmed AlGhamedi. A review of saddle-point approximation: Theory and applications. *Advances and Applications in Statistics*, 92(3):343–390, 2025.

[9] G. Arfwedson. Research in collective risk theory: The case of equal risk sums. *Scandinavian Actuarial Journal*, 1953(sup1):1–15, 1953.

[10] Gerhard Arfwedson. Research in collective risk theory: Part ii. *Scandinavian Actuarial Journal*, 1955(1–2):53–100, 1955.

[11] Lee J Bain and David L Weeks. A note on the truncated exponential distribution. *The Annals of Mathematical Statistics*, 35(3):1366–1367, 1964.

[12] Wioleta Baranowska-Zając. Freedom of establishment on the internal insurance market of the european union in the light of directive 2009/138/ec (solvency ii). *Mediterranean Journal of Social Sciences*, 5(19):661–670, 2014.

[13] R. E. Beard, T. Pentikäinen, and E. Pesonen. *Risk Theory*. Chapman and Hall, 2 edition, 1977.

[14] J. A. Beekman. The collective risk model and the total claim amount. *Skandinavisk Aktuarietidskrift*, 49:208–220, 1966.

[15] John A. Beekman. Risk convolution calculations. In *Premium Calculation in Insurance*, pages 19–30. Springer, 1984.

[16] Aleksandr Beknazaryan and Peter Adamic. On a stochastic order induced by an extension of panjer's family of discrete distributions. *Metrika*, 85(1):67–91, 2022.

[17] Vladimir E. Bening and Victor Yu Korolev. *Generalized Poisson Models and Their Applications in Insurance and Finance*. De Gruyter, 2002.

[18] Stefan Bernegger. The swiss re exposure curves and the mbbefd distribution class1. *ASTIN Bulletin: The Journal of the IAA*, 27(1):99–111, 1997.

[19] Suparna Biswas and Rituparna Sen. Estimation of spectral risk measure for left truncated and right censored data. arXiv:2402.14322, 2024.

[20] Christopher Blier-Wong, Hélène Cossette, and Étienne Marceau. Collective risk models with fgm dependence. *Scandinavian Actuarial Journal*, 2025(2):139–167, 2024.

[21] Ronald W Butler. Distribution theory for dependent renewal–reward processes and their first-passage times using saddlepoint and residue expansions. *Journal of Applied Probability*, 62(2):447–474, 2025.

[22] James W. Cooley and John W. Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of Computation*, 19(90):297–301, 1965.

[23] Harald Cramér. *On the Mathematical Theory of Risk*. Skandia Jubilee Volume, Stockholm, 1930.

[24] Harald Cramér. Collective risk theory: A survey from the point of view of the theory of stochastic processes. In *Collected Works II*, pages 1028–1115. Springer, Berlin, 1994.

[25] Harald Cramér. On some questions connected with mathematical risk. In *Collected Works II*, pages 1003–1027. Springer, 1994.

[26] Harald Cramér. *Collective Risk Theory: A Survey of the Theory from the Point of View of the Theory of Stochastic Processes*. Thomas Nelson & Sons, 1955.

[27] Henry E Daniels. Saddlepoint approximations in statistics. *The Annals of Mathematical Statistics*, pages 631–650, 1954.

[28] Nelson De Pril. Recursions for convolutions of arithmetic distributions. *ASTIN Bulletin: The Journal of the IAA*, 15(2):135–139, 1985.

[29] E. G. Déniz and E. C. Ojeda. The compound dgl/erlang distribution in the collective risk model. *Revista de Métodos Cuantitativos para la Economía y la Empresa*, 16:121–142, 2013.

[30] D. C. M. Dickson. A review of panjer's recursion formula and its applications. *British Actuarial Journal*, 1(1):107–124, 1995.

[31] David C. M. Dickson. *Insurance Risk and Ruin*. Cambridge University Press, 2016.

[32] David C. M. Dickson, Laura M. Tedesco, and Ben Zehnwirth. Predictive aggregate claims distributions. *Journal of Risk and Insurance*, 65(4):689–709, 1998. doi: 10.2307/253539.

[33] Zoran Djurić. Collective risk model in non-life insurance. *Economic Horizons*, 15(2):167–175, 2013. doi: 10.5937/ekonhor1302167D.

[34] Louis G. Doray and A. Haziza. Minimum distance inference for sundt's distribution. In *COMPSTAT 2004—Proceedings in Computational Statistics, Prague*, 2004.

[35] S. K.-B. Dzidzornu and R. Minkah. Assessing the performance of the discrete generalised pareto distribution in modelling non-life insurance claims. *Journal of Probability and Statistics*, 2021:1–8, 2021.

[36] M. M. Mohie El-Din, M. M. Amein, and A. M. Abd El-Raheem. On mid–truncated distributions and its applications. *Journal of Advanced Research in Applied Mathematics*, 5(2):20–38, 2013.

[37] Paul Embrechts and Marco Frei. Panjer recursion versus fft for compound distributions. *Mathematical Methods of Operations Research*, 69:497–508, 2009.

[38] Michael Fackler. Panjer class united: one formula for the probabilities of the poisson, binomial, and negative binomial distribution. In *Anales del Instituto de Actuarios Españoles*, volume 17, pages 1–12, 2011.

[39] Michael Fackler. Panjer class revisited: one formula for the distributions of the panjer (a, b, n) class. *Annals of Actuarial Science*, 17(1):145–169, 2023.

[40] Adam Farooq. *Modelling inference strategies and robust clustering topologies*. PhD thesis, Aston University, 2022.

[41] William Feller. *An Introduction to Probability Theory and Its Applications*. Wiley, 1968. Vol. 1, 3rd ed.

[42] Shui Feng. Truncated two-parameter gem distribution. *Discrete and Continuous Dynamical Systems-S*, pages 0–0, 2025.

[43] Edward Frees. Loss data analytics. arXiv:1808.06718, 2018.

[44] Selim Gatti and Mario V. Wüthrich. Modeling lower-truncated and right-censored insurance claims with an extension of the mbbefd class. *European Actuarial Journal*, 15(1):199–240, 2025.

[45] Hans Gerber. An introduction to mathematical risk theory. Lecture notes, 1979.

[46] Sarra Ghaddab, Manel Kacem, Christian de Peretti, and Lotfi Belkacem. Extreme severity modeling using a glm–gpd combination: application to an excess-of-loss reinsurance treaty. *Empirical Economics*, pages 1–23, 2023.

[47] Alois Gisler. The insurance risk in the sst and in solvency ii: modelling and parameter estimation. *SSRN eLibrary*, 2009. SSRN 2704364.

[48] P.-O. Goffard, S. Rao Jammalamadaka, and Simos G. Meintanis. Goodness-of-fit procedures for compound distributions with an application to insurance. *Journal of Statistical Theory and Practice*, 16(3):52, 2022.

[49] Roger J. Gray and Susan M. Pitts. *Risk Modelling in General Insurance: From Principles to Practice*. Cambridge University Press, 2012.

[50] John Gurland and Ram Tripathi. Estimation of parameters on some extensions of the katz family of discrete distributions involving hypergeometric functions. In *A Modern Course on Statistical Distributions in Scientific Work: Volume 1—Models and Structures*, pages 59–82. Springer, 1975.

[51] G. G. Hamedani, Hafida Goual, Walid Emam, Yusra Tashkandy, Fiaz Ahmad Bhatti, Mohamed Ibrahim, and Haitham M. Yousof. A new right-skewed one-parameter distribution with mathematical characterizations, distributional validation, and actuarial risk analysis, with applications. *Symmetry*, 15(7):1297, 2023.

[52] Philip E. Heckman and Glenn G. Meyers. The calculation of aggregate loss distributions from claim severity and claim count distributions. In *Proceedings of the Casualty Actuarial Society*, volume 70, pages 49–66. Casualty Actuarial Society, 1983.

[53] Klaus Th. Hess, Anett Liewald, and Klaus D. Schmidt. An extension of panjer's recursion. *ASTIN Bulletin: The Journal of the IAA*, 32(2):283–297, 2002.

[54] Ole Hesselager. A recursive procedure for calculation of some compound distributions. *ASTIN Bulletin: The Journal of the IAA*, 24(1):19–32, 1994.

[55] Alexandre Hocquard, Nicolas Papageorgiou, and Bruno Rémillard. The payoff distribution model: an application to dynamic portfolio insurance. *Quantitative Finance*, 15(2):299–312, 2015.

[56] Louisa Hofmann. Approximation methods for the total claim amount in collective risk modeling. Manuscript, 2022.

[57] Robert V. Hogg and Stuart A. Klugman. *Loss Distributions*. Wiley, 2009.

[58] E. Hovinen. A method to compute convolution. In *Congress of Actuaries*, 1964.

[59] Weenakorn Ieosanurak and Adisak Moumeesri. Estimating ruin probability in an insurance risk model using the wang–ph transform through claim simulation. *Emerging Science Journal*, 9:188–209, 2025.

[60] Pavla Jindrová and Viera Pacáková. Modelling of extreme losses in natural disasters. *International Journal of Mathematical Models and Methods in Applied Sciences*, 10, 2016.

[61] Norman L. Johnson, Samuel Kotz, and Narayanaswamy Balakrishnan. *Continuous Univariate Distributions, Volume 2*, volume 2. John Wiley & Sons, 1995.

[62] Norman L. Johnson, Adrienne W. Kemp, and Samuel Kotz. *Univariate Discrete Distributions*, volume 444. John Wiley & Sons, 2005.

[63] Paul M. Kahn. An introduction to collective risk theory and its application to stop-loss reinsurance. *Transactions of Society of Actuaries*, 14(40):400–425, 1962.

[64] Leo Katz. Unified treatment of a broad class of discrete probability distributions. *Classical and Contagious Discrete Distributions*, 1:175–182, 1965.

[65] Stuart A. Klugman, Harry H. Panjer, and Gordon E. Willmot. *Loss Models: From Data to Decisions*, volume 715. John Wiley & Sons, 2012.

[66] Mohamed Kouadria and Halim Zeghdoudi. The truncated new-xlindley distribution with applications. *Journal of Computational Analysis and Applications*, 34 (3), 2025.

[67] Markus Kreer, Ayşe Kızılersü, Anthony W. Thomas, and Alfredo D. Egídio dos Reis. Goodness-of-fit tests and applications for left-truncated weibull distributions to non-life insurance. *European Actuarial Journal*, 5:139–163, 2015.

[68] K KRISHNAKUMARI and Dais George. A new left truncated distribution for modeling failure time data: Estimation, robustness study and application. *Statistics, Optimization & Information Computing*, 13(6):2264–2277, 2025.

[69] J. F. Lawless. Truncated distributions. *Wiley StatsRef: Statistics Reference Online*, 2014.

[70] Hao Liu and Ruodu Wang. Collective risk models with dependence uncertainty. *ASTIN Bulletin: The Journal of the IAA*, 47(2):361–389, 2017.

[71] Filip Lundberg. Some supplementary researches on the collective risk theory. *Scandinavian Actuarial Journal*, 1932(3):137–158, 1932.

[72] Filip Lundberg. *On the Numerical Application of the Collective Risk Theory.* Almqvist & Wiksell, 1934.

[73] Andrew Luong and Louis G. Doray. General quadratic distance methods for discrete distributions definable recursively. *Insurance: Mathematics and Economics*, 30(2):255–267, 2002.

[74] Andrew Luong and José Garrido. Minimum quadratic distance estimation for a parametric family of discrete distributions defined recursively. *Australian Journal of Statistics*, 35(1):59–67, 1993.

[75] Ján Mačutek and Gabriel Altmann. Discrete and continuous modelling in quantitative linguistics. *Journal of Quantitative Linguistics*, 14(1):81–94, 2007.

[76] Debiao Meng, Yipeng Guo, Yihe Xu, Shiyuan Yang, Yongqiang Guo, Lidong Pan, and Xinkai Guo. Saddlepoint approximation method in reliability analysis: A review. *Computer Modeling in Engineering & Sciences (CMES)*, 139(3), 2024.

[77] Mohammed A. Meraou, Noriah M. Al-Kandari, Mohammad Z. Raqab, and Debasis Kundu. Analysis of skewed data using compound poisson–exponential distribution with applications to insurance claims. *Journal of Statistical Computation and Simulation*, 92(5):928–956, 2022.

[78] Glenn G. Meyers. An analysis of retrospective rating. *PCAS*, 67:110, 1980.

[79] Shaw Mong. Estimating aggregate loss probability and increased limit factor. *Pricing Property and Casualty Insurance Products, Casualty Actuarial Society*, page 358, 1980.

[80] Mohanan Monisha, Radhakumari Maya, Muhammed Rasheed Irshad, Christophe Chesneau, and Damodaran Santhamani Shibu. A new generalization of the zero-truncated negative binomial distribution by a lagrange expansion with associated regression model and applications. *International Journal of Data Science and Analytics*, pages 1–15, 2023.

[81] Saralees Nadarajah. Some truncated distributions. *Acta Applicandae Mathematicae*, 106(1):105–123, 2009.

[82] Sintya Putri Nugrahainy and Azizah Azizah. Estimasi premi bruto asuransi kendaraan bermotor menggunakan metode panjer recursion dan fast fourier transform. *Jambura Journal of Mathematics*, 7(2):129–137, 2025.

[83] V. Pacáková and J. Gogola. Pareto distribution in insurance and reinsurance. In *Financial Management of Firms and Financial Institutions, 9th Int. Scientific Conf., VŠB Ostrava*, pages 298–306, 2013.

[84] V. Packová and D. Brebera. Loss distributions in insurance risk management. *Recent Advances on Economics and Business Administration*, pages 17–22, 2015.

[85] Ro J. Pak. Estimating loss severity distribution convolution approach. *Journal of Mathematics and Statistics*, 10(3):247–254, 2014.

[86] H. H. Panjer and G. E. Willmot. Recursions for compound distributions. *ASTIN Bulletin: The Journal of the IAA*, 13(1):1–12, 1982.

[87] Harry H. Panjer. The aggregate claims distribution and stop-loss reinsurance. *Proceedings of the Casualty Actuarial Society*, 32, 1980.

[88] Harry H. Panjer. Recursive evaluation of a family of compound distributions. *ASTIN Bulletin: The Journal of the IAA*, 12(1):22–26, 1981.

[89] Harry H. Panjer and B. W. Lutek. Practical aspects of stop-loss calculations. *Insurance: Mathematics and Economics*, 2(3):159–177, 1983.

[90] Harry H. Panjer and Gordon E. Willmot. *Insurance Risk Models*. Society of Actuaries, 1992.

[91] Dinis D. Pestana and Sílvio F. Velosa. Extensions of katz–panjer families of discrete distributions. *REVSTAT—Statistical Journal*, 2(2), 2004.

[92] C. Philipson. Some results concerning compound poisson distributions. *Skandinavisk Aktuarietidskrift*, 38:57–72, 1955.

[93] Carl Philipson. Some convolution theorems for compound poisson processes. *Scandinavian Actuarial Journal*, 1967(3–4):203–210, 1967.

[94] Carl Philipson. A review of the collective theory of risk: Part i. comments on the development of the theory. *Scandinavian Actuarial Journal*, 1968(1-2):45–68, 1968.

[95] Chudamani Poudyal and Vytaras Brazauskas. Robust estimation of loss models for truncated and censored severity data. arXiv:2202.13000, 2022.

[96] Narahari U. Prabhu. On the ruin problem of collective risk theory. *The Annals of Mathematical Statistics*, 32(3):757–764, 1961.

[97] Hosenur Rahman Prodhani and Rama Shanker. The truncated sujatha distribution with properties and applications in engineering and medical sciences. *Journal of Statistical Theory and Practice*, 19(4):85, 2025.

[98] S. Ramasubramanian. Poisson process and insurance: An introduction. Technical report, Indian Statistical Institute, New Delhi, 2005. Lecture notes.

[99] Mohammad Z. Raqab, Debasis Kundu, and Fahimah A. Al-Awadhi. Compound zero-truncated poisson normal distribution and its applications. *Communications in Statistics—Theory and Methods*, 50(13):3030–3050, 2021.

[100] Nancy Reid. Saddlepoint methods and statistical inference. *Statistical science*, pages 213–227, 1988.

[101] Tomasz Rolski, Hanspeter Schmidli, Volker Schmidt, and Jozef L. Teugels. *Stochastic Processes for Insurance and Finance*. John Wiley & Sons, 2009.

[102] T. Saxon. On the probability of ruin in the collective risk theory for insurance enterprises with only negative risk sums. *Skandinavisk Aktuarietidskrift*, 1948.

[103] Klaus J. Schröter. On a family of counting distributions and recursions for related compound distributions. *Scandinavian Actuarial Journal*, 1990(2–3):161–175, 1990.

[104] Hilary L. Seal. Stochastic theory of a risk business. Unpublished, 1969.

[105] C.-O. Segerdahl. *On Homogeneous Random Processes and Collective Risk Theory*. PhD thesis, Stockholm University, 1939.

[106] J. M. Shapiro. Sums of independent truncated random variables. *The Annals of Mathematical Statistics*, pages 754–761, 1957.

[107] Peng Shi and Zifeng Zhao. Regression for copula-linked compound distributions with applications in modeling aggregate insurance claims. *The Annals of Applied Statistics*, 14(1):357–380, 2020.

[108] Sanjay Kumar Singh, Umesh Singh, and Vikas Kumar Sharma. The truncated lindley distribution: Inference and application. *Journal of Statistics Applications & Probability*, 3(2):219, 2014.

[109] Surang Sitho, Sunthree Denthet, and Hira Nadeem. Zero truncated negative binomial weighted weibull distribution and its application. *Lobachevskii Journal of Mathematics*, 42(13):3241–3252, 2021.

[110] Walter L. Smith. A note on truncation and sufficient statistics. *The Annals of Mathematical Statistics*, pages 247–252, 1957.

[111] W. L. Stevens. The truncated normal distribution. *Annals of Applied Biology*, 24(4):815–852, 1937.

[112] E. Stroh. Actuarial note: The distribution function of collective risk theory as linear compounds. *Actuarial Research Clearing House*, 1978.

[113] Bjørn Sundt. On some extensions of panjer's class of counting distributions. *ASTIN Bulletin: The Journal of the IAA*, 22(1):61–80, 1992.

[114] Bjørn Sundt and William S. Jewell. Further results on the recursive evaluation of compound distributions. *ASTIN Bulletin: The Journal of the IAA*, 12(1):27–39, 1981.

[115] Bjørn Sundt and Raluca Vernic. *Recursions for Convolutions and Compound Distributions with Insurance Applications*. Springer, 2009.

[116] Gábor Szűcs. Panjerove rekurzie v neživotnom poistení. Diplomová práca, Fakulta matematiky, fyziky a informatiky, Univerzita Komenského v Bratislave, Bratislava, 2011.

[117] Gábor Szűcs. Metódy odhadu parametrov rekurentnỳch tried diskrétnych rozdelení: Parameter estimation methods for recurrent classes of discrete probability distributions. In *Forum Statistikum Slovacum*, page 154, 2013.

[118] Gábor Szűcs. Schröterova trieda rozdelení. *Forum Statisticum Slovacum*, IX(7): 227–232, 2013. ISSN 1336-7420.

[119] Yanbo Tang and Nancy Reid. Laplace and saddlepoint approximations in high dimensions. *Bernoulli*, 31(3):1759–1788, 2025.

[120] H. R. Thompson. Truncated normal distributions. *Nature*, 165(4194):444–445, 1950.

[121] James Tilley. Discussion of harry panjer, "the aggregate claim distribution and stop-loss reinsurance". *Transactions of the Society of Actuaries*, 32:538, 1980.

[122] Gary Venter. Transformed beta and gamma distributions and aggregate losses. In *Proceedings of the Casualty Actuarial Society*, volume 70, pages 289–308, 1983.

[123] Raluca Vernic. Sundt's classes of distributions. *Wiley StatsRef: Statistics Reference Online*, 2014.

[124] Shaun Wang and Monica Sobrero. Further results on hesselager's recursive procedure for calculation of some compound distributions. *ASTIN Bulletin: The Journal of the IAA*, 24(2):161–166, 1994.

[125] Suojin Wang. Saddlepoint approximations for bivariate distributions. *Journal of Applied Probability*, 27(3):586–597, 1990.

[126] Mohammad Kafeel Wani and Peer Bilal Ahmad. One-inflated zero-truncated poisson distribution: Statistical properties and real life applications. *Annals of Data Science*, 12(2):639–666, 2025.

[127] Deborah Warren. Statistical modeling of insurance data with heavy-tailed distributions: Computation of actuarial functions through equilibrium distribution simulation. Preprint, 2025.

[128] Gordon E. Willmot and Harry H. Panjer. Difference equation approaches in evaluation of compound distributions. *Insurance: Mathematics and Economics*, 6(1):43–56, 1987.

[129] Gordon E. Willmot and Bjørn Sundt. On evaluation of the delaporte distribution and related distributions. *Scandinavian Actuarial Journal*, 1989(2):101–113, 1989.

[130] Gejza Wimmer and Gabriel Altmann. *Thesaurus of Univariate Discrete Probability Distributions*. Stamm, 1999.

[131] Esther Yartey. *The (a, b, r) Class of Discrete Distributions with Applications*. PhD thesis, Laurentian University of Sudbury, 2020.

[132] Haitham Yousof, Mahmoud Afshari, Morad Alizadeh, Vahid Ranjbar, R. Minkah, Mohamed S. Hamed, and Moustafa Salem. A novel insurance claims (revenues) xgamma extension: Distributional risk analysis with financial port-var analysis. *Pakistan Journal of Statistics and Operation Research*, pages 83–117, 2025.

[133] Pengcheng Zhang, Enrique Calderin, Shuanming Li, and Xueyuan Wu. On the type i multivariate zero-truncated hurdle model with applications in health insurance. *Insurance: Mathematics and Economics*, 90:35–45, 2020.

[134] Pengcheng Zhang, David Pitt, and Xueyuan Wu. A comparative analysis of several multivariate zero-inflated and zero-modified models with applications in insurance. *Communications in Statistics—Theory and Methods*, 54(7):2130–2157, 2025.

[135] Qian Zhao, Sahadeb Upretee, and Daoping Yu. Model uncertainty and selection of risk models for left-truncated and right-censored loss data. *Risks*, 11(11):188, 2023.

[136] Alain F. Zuur, Elena N. Ieno, Neil J. Walker, Anatoly A. Saveliev, and Graham M. Smith. Zero-truncated and zero-inflated models for count data. In *Mixed Effects Models and Extensions in Ecology with R*, pages 261–293. Springer, 2009.