International Journal of Mathematical, Engineering and Management Sciences
Vol. 10, No. 2, 486-505, 2025
https://doi.org/10.33889/IJMEMS.2025.10.2.024

Ram Arti Publishers

Ruin Probability Analysis in Automobile Insurance Using Gamma-Cubic Transmuted Exponential Distribution for Claim Severity

Jiramet Hengcharoensuk
Department of Statistics,
Faculty of Science, Silpakorn University, Nakhon Pathom, 73000, Thailand.
E-mail: Hengcharoensuk_j@silpakorn.edu

Adisak Moumeesri
Department of Statistics,
Faculty of Science, Silpakorn University, Nakhon Pathom, 73000, Thailand.
Corresponding author: moumeesri_a@silpakorn.edu

(Received on September 10, 2024; Revised on November 17, 2024 & December 16, 2024; Accepted on December 24, 2024)

Abstract
Insurance companies must confront and manage financial risks to ensure their survival. This article analyzes and manages risks by estimating ruin probability using the newly developed Gamma-Cubic Transmuted Exponential (GCTE) claim severity distribution. The GCTE distribution is obtained by mixing distribution and Cubic Rank Transmuted map techniques. Its parameters are estimated using the maximum likelihood estimation (MLE) method. The performance of GCTE distribution is compared with well-known loss distributions such as Gamma, Weibull, Lognormal, Inverse Gaussian, mixture Lognormal, and the Cubic Transmuted Exponential (CTE) distribution to identify the best-fitting distribution for claim severity data. To estimate ruin probability, a sample data simulation technique known as acceptance-rejection is employed in the surplus process. Additionally, the exponential function is used to predict the minimum capital reserve necessary to safeguard the company under acceptable risk levels. These methodologies aid decision-making and promote financial stability within non-life insurance companies.

Keywords- Cubic rank transmutations, Loss distribution, Claim severity, Ruin probability, Capital reserve.

1. Introduction

Insurance is an important business in the economies of many countries around the world. It is a financial arrangement in which the insured pays premiums to an insurance company to receive protection against risks that may occur in the event of a catastrophic event or damage, as specified in the policy. The insurance company provides compensation according to the specified conditions. There are many different types of insurance. Life insurance, for example, focuses on providing protection for the life of the insured or others who are financially affected in the event of death. Non-life insurance covers damages to property or businesses, protecting against losses that occur from accidents or natural disasters. Casualty insurance can be further divided into various types, such as motor insurance and miscellaneous casualty insurance. Insurance is a crucial tool for mitigating or distributing risk for policyholders. However, the insurance business itself also faces risks in investment and operations.

Risks are a constant challenge for insurance companies, as well as for investment in other industries. Each industry has its own approach to managing investment risks, depending on the specific factors involved. For insurance companies, investment risks are crucial for the survival of the insurance industry. Various factors contribute to these risks, including financial instability, legal uncertainties, and exposure to natural disasters. These risks can impact the probability of ruin and the capital reserves of insurance companies. On the other hand, financial risk is equally significance to the companies in making their decisions for the following year. This article will propose a loss distribution to estimate ruin probability, aiding in the

analysis and management of risks, thereby enhancing the overall stability and resilience of insurance operations.

Filip Lundberg, a Swedish actuary, laid the groundwork for modern risk theory in 1903 with his model for non-life insurance. He focused on the arrival and amount of claims, which are critical to an insurance company's capital. In the classical risk process, ruin probability occurs at claim times T_n such that $0 \leq T_1 \leq T_2 \leq \cdots$. The discrete-time risk process is defined by $U_n = u + cT_n - \sum_{k=1}^{n} X_k$, with $U_0 = u$, $n = 1,2,3, \ldots$, with u as initial capital and c as the premium rate. Chan and Zhang (2006) simplified this by assuming $T_n = n$. Sattayatham et al. (2013) provided a closed form for the probability of ruin and studied the minimum capital reserve, while Klongdee et al. (2013) further investigated the risk model for motor insurance. Liu et al. (2018) studied the ruin probability over a discrete finite time period using a heavy-tailed probability distribution. Santana and Rincón (2020) use the Pollaczek–Khinchine formula to present a general approach for calculating the ultimate ruin probability in a Gerber–Dickson risk model, considering the claim severity under a probability distribution. Cruz et al. (2024) applied complex sequence theory to calculate ruin probabilities in the risk process over a discrete finite time period. It is evident that these studies use random variables in probability theory to describe the distribution of claim severity.

A probability distribution is a theoretical framework used to describe the nature of data. Heavy-tailed distributions are particularly suitable for data characterized by a significant right tail. In insurance, heavy-tailed distributions are crucial because they can model catastrophic events and large claims more accurately than lighter-tailed distributions. Commonly used heavy-tailed distributions include the Gamma distribution, Inverse Gaussian distribution, Weibull distribution, and Lognormal distribution. Over time, new probability distributions have been developed to better fit insurance claim data and enhance model efficiency. Insurance premiums are often calculated using probability distributions. Frangos and Vrontos (2001), Mert and Saykan (2005), Ni et al. (2014), Moumeesri et al. (2020), Pongsart et al. (2021), Moumeesri and Pongsart (2022) and Ieosanurak et al. (2023) utilized techniques such as mixing distributions to create distributions for calculating insurance premiums based on the probability distribution of the number of claims and the severity of claims. Rahman et al. (2019) proposed a new cubic transmuted family of distributions. Additionally, Adetunji and Sabri (2023) developed the Cubic Transmuted Exponential distribution (CTED) using the Cubic Rank Transmuted map technique to improve the fit of the exponential distribution to data. We continue to create a new distribution that better matches the data, aiming to improve the accuracy and efficiency of its model.

Currently, the data obtained for analysis do not always conform to the assumptions of traditional probability distributions. Over the past century, many models have been developed to address these issues, with a focus on expanding the scope of interest. This has included modifying base distributions (such as Poisson, Exponential, Weibull, Lindley, etc.) into more general forms, known as generalized probability distributions.

The Generalized Poisson distribution, introduced by Satterthwaite (1942), was modified from the Poisson distribution by adding a shape parameter to make the model more flexible. The Generalized Weibull distribution was presented by Mudholkar et al. (1996). The Generalized Exponential distribution, introduced by Gupta and Kundu (1999), is another notable example. Shaw and Buckley (2009) developed a new family of distributions known as the "Quadratic Transmuted Family of Distributions." The transmutation approach has been recognized as a better fit for the data, leading to the development and widespread use of the Quadratic Transmuted family of distributions, as shown in the equation:

$$F(x) = (1 + \lambda)G(x) - \lambda G^2(x), \ \lambda \in [-1,1].$$

where, $G(x)$ is the cumulative distribution function (CDF) of the specified distribution, and when $\lambda = 0$, the equation simplifies to the base distribution. A decade later, Granzotto et al. (2017) introduced a new family of transmuted distributions called the "Cubic Transmuted Family of Distributions," which enhanced the flexibility of transmuted distributions for applications in risk models, such as those used in financial companies or insurance. The most recent development in this field was made by Rahman et al. (2019), who created a new Cubic Transmuted family of distributions, which has been utilized in this work.

In this article, we introduce a new heavy-tailed distribution, the Gamma-Cubic Transmuted Exponential (GCTE) distribution, derived from a mixture of the Gamma distribution and the Cubic Transmuted Exponential (CTE) distribution. Additionally, we compare its efficiency with other loss distributions to identify a claim severity distribution that accurately reflects the characteristics of insurance claim data. We estimate the ruin probability based on the proposed distribution and identify the minimum capital reserve necessary within acceptable risk thresholds at various levels for establishing a company. This new loss distribution serves as a valuable tool in the insurance industry for assessing investment risks and formulating investment strategies that align with the company's risk tolerance.

2. Generating Distributions for Claim Severity

The primary objective of this study is to identify the distribution of claim severity that best represents the characteristics of dataset, thereby enhancing the efficiency of the risk process model. Initial analysis indicated that traditional heavy-tailed distributions did not adequately capture the heavy-tail behavior typical of insurance risk data. This led to the development of a new heavy-tailed distribution specifically tailored to better reflect real-world insurance data. Two key techniques were employed: cubic rank transmutation and a mixing distribution approach.

2.1 Cubic Rank Transmuted Exponential Distribution

Let Y be a random variable following an exponential distribution with a parameter θ. Its cumulative distribution function (CDF) is expressed as

$$G(y) = 1 - e^{-\theta y};\ \theta > 0, y > 0 \tag{1}$$

The CDF of the cubic rank transmutation map (Rahman et al., 2019) is given by

$$F(y) = (1-p)G(y) + 3p\big(G(y)\big)^2 - 2p\big(G(y)\big)^3,\ p \in [-1,1] \tag{2}$$

The Cubic Transmuted Exponential (CTE) distribution is obtained by substituting Equation (1) into Equation (2). Therefore, the CDF and probability density function (PDF) of the CTE distribution are displayed as Equation (3) and (4), respectively.

$$F(y) = 1 - e^{-\theta y} + pe^{-\theta y} + 2pe^{-3\theta y} - 3pe^{-2\theta y} \tag{3}$$

$$f(y) = \theta e^{-\theta y}\big(1 - p + 6pe^{-\theta y} - 6pe^{-2\theta y}\big) \tag{4}$$

Figure 1 shows the PDF shape of the CTE distribution for different parameter values. Despite the varying parameters, the distribution remains right-skewed.

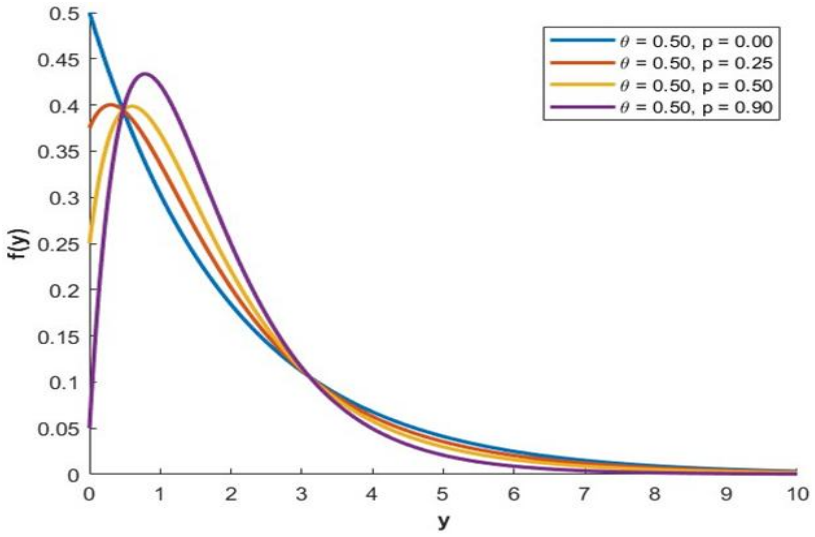

Figure 1. Shape of the PDF for the CTE distribution.

When Y follows the CTE distribution, the moment about the origin is given as follows:

$$\mu'_r = E(Y^r) = \int_0^\infty y^r f(y)dy = \left(1 - p + \frac{3p}{2^r} - \frac{2p}{3^r}\right)\frac{r!}{\theta^r} \quad (5)$$

If Y follows the CTE distribution, the moment generating function (MGF) is given as follows:

$$M_y(t) = E(e^{ty}) = \int_0^\infty e^{ty} f(y)dy = \frac{\theta}{\theta - t} - \frac{p^\theta}{\theta - t} + \frac{6p\theta}{2\theta - t} - \frac{6p\theta}{3\theta - t} \quad (6)$$

As shown in Equation (5), the mean and variance of the CTE distribution are given in Equations (7) and (8), respectively.

$$E(Y) = \frac{\left(1 - p + \frac{3p}{2} - \frac{2p}{3}\right)}{\theta} \quad (7)$$

$$Var(Y) = \frac{(36 - 22p - p^2)}{36\theta^2} \quad (8)$$

2.2 Mixed Gamma with Cubic Transmuted Exponential Distribution

Proportions 1. Let's suppose that X is distributed according to a Gamma distribution with parameters α and β, represented as $X \sim \text{Gam}(\alpha, \beta)$, where the parameter β follows the CTE distribution in Equation (4). Then, the PDF of the Gamma-Cubic Transmuted exponential (GCTE) distribution is obtained as:

$$f_X(x) = \theta\alpha x^{\alpha-1}\left(\frac{1-p}{(\theta+x)^{\alpha+1}} + \frac{6p}{(2\theta+x)^{\alpha+1}} - \frac{6p}{(3\theta+x)^{\alpha+1}}\right) \quad (9)$$

where, $x > 0, \theta > 0, \alpha > 0$, and $p \in [-1,1]$.

Proof:

$$f_X(x) = \int_0^\infty f(x;\alpha,\beta) f(\beta)\, d\beta$$

$$= \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \cdot \theta e^{-\theta\beta}\left(1 - p + 6pe^{-\theta\beta} - 6pe^{-2\theta\beta}\right) d\beta$$

$$= \frac{\theta x^{\alpha-1}}{\Gamma(\alpha)}\left(\int_0^\infty \beta^\alpha e^{-\beta(\theta+x)}d\beta - p\int_0^\infty \beta^\alpha e^{-\beta(\theta+x)}d\beta + 6p\int_0^\infty \beta^\alpha e^{-\beta(2\theta+x)}d\beta - 6p\int_0^\infty \beta^\alpha e^{-\beta(3\theta+x)}d\beta\right)$$

$$= \frac{\theta x^{\alpha-1}}{(\alpha-1)!}\left(\frac{\alpha!}{(\theta+x)^{\alpha+1}} - \frac{p\alpha!}{(\theta+x)^{\alpha+1}} + \frac{6p\alpha!}{(2\theta+x)^{\alpha+1}} - \frac{6p\alpha!}{(3\theta+x)^{\alpha+1}}\right)$$

$$= \theta\alpha x^{\alpha-1}\left(\frac{1-p}{(\theta+x)^{\alpha+1}} + \frac{6p}{(2\theta+x)^{\alpha+1}} - \frac{6p}{(3\theta+x)^{\alpha+1}}\right)$$

The PDF of the GCTE distribution exhibits right-skewness for various parameter values, as illustrated in **Figure 2**.

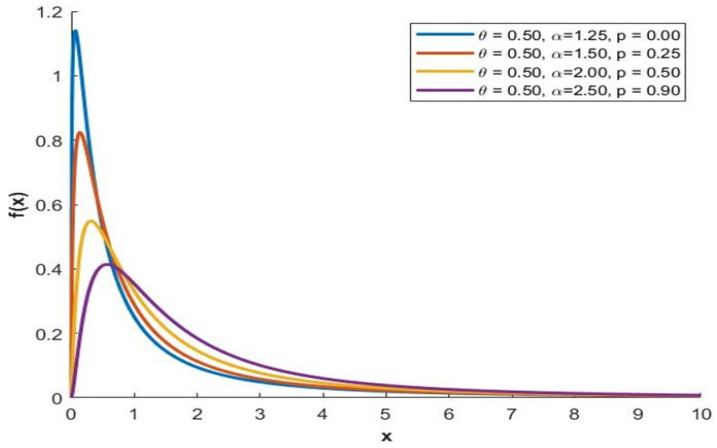

Figure 2. Shape of the PDF for the GCTE distribution.

2.2.1 Parameter Estimation

In this article, the parameters are estimated through the maximum likelihood estimation (MLE) method. Let $x_1, x_2, \ldots, x_n$ be a random sample of size n that is distributed from the GCTE distribution with parameters α, θ and p. The likelihood function is

$$L = \prod_{i=1}^{n}\left[\theta\alpha {x_i}^{\alpha-1}\left(\frac{1-p}{(\theta+x_i)^{\alpha+1}} + \frac{6p}{(2\theta+x_i)^{\alpha+1}} - \frac{6p}{(3\theta+x_i)^{\alpha+1}}\right)\right].$$

The log-likelihood function is

$$\ln(L) = \sum_{i=1}^{n}\ln\left[\theta\alpha {x_i}^{\alpha-1}\left(\frac{1-p}{(\theta+x_i)^{\alpha+1}} + \frac{6p}{(2\theta+x_i)^{\alpha+1}} - \frac{6p}{(3\theta+x_i)^{\alpha+1}}\right)\right].$$

Since direct solutions are not feasible, numerical methods are employed. In this study, we use the 'optim' commands in the R programming language for ease of estimation.

3. Numerical Application

The dataset consists of actual motor insurance claims provided by a non-life insurance company in Thailand for the year 2009 (Sattayatham and Talangtam, 2012). The data covers a voluntary plan for Type 2+ vehicles and focuses on the claim severity variable (in Thai baht). The average claim severity is 17,662 baht, with a standard deviation of 41,331.63 baht. Claim severity ranges from a minimum of 159 baht to a maximum of 899,879 baht based on a total of 1,296 claims. To illustrate the nature of the additional information derived

from the above statistics (such as the distribution of the data, skewness, etc.), the total claim severity data of 1,296 observations is presented using a histogram, as shown in **Figure 3**.

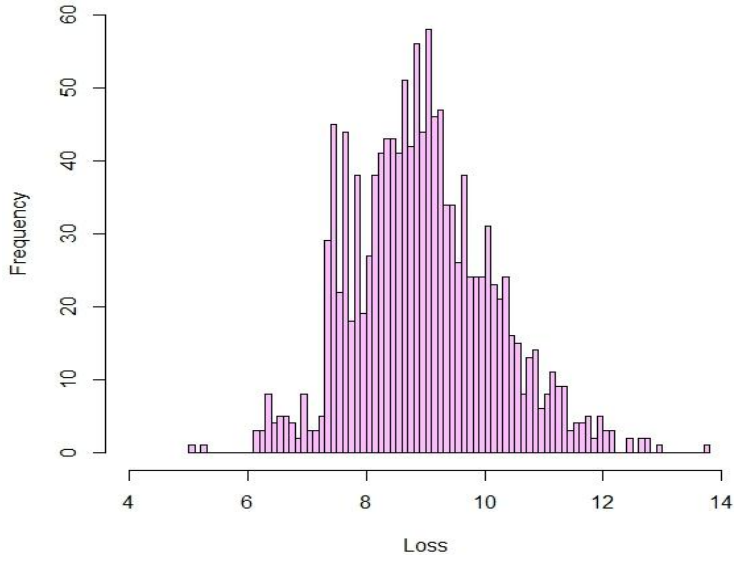

Figure 3. Histogram shows the distribution of total claim values for 1,296 observations (log scale).

3.1 Fitting Models

The goodness of fit test is a statistical method used to assess how well a model fits the data. This test can be performed in several ways. In this article, our focus is on the Kolmogorov-Smirnov (KS) test and Anderson-Daring (AD), which compares the distance between the empirical cumulative distribution and the theoretical cumulative distribution. Model selection criteria are determined using the Akaike Information Criterion (AIC).

Testing the hypothesis of the KS test and AD test:

H_0: The data follows the specified distribution.

H_1: The data does not follow the specified distribution.

Table 1. PDF and CDF of the claim severity distributions.

CTE	PDF: $f(y) = \theta e^{-\theta y}\left(1 - p + 6pe^{-\theta y} - 6pe^{-2\theta y}\right)$ CDF: $F(y) = 1 - e^{-\theta y} + pe^{-\theta y} + 2pe^{-3\theta y} - 3pe^{-2\theta y}$
Gamma	PDF: $f(y) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y}$ CDF: $F(y) = \frac{1}{\Gamma(\alpha)} \gamma(\alpha, \beta y)$
Weibull	PDF: $f(y) = \frac{\lambda y^{\lambda-1}}{\delta^{\lambda}} e^{-\left(\frac{y}{\delta}\right)^{\lambda}}$ CDF: $F(y) = 1 - e^{-\left(\frac{y}{\delta}\right)^{\lambda}}$
Lognormal	PDF: $f(y) = \frac{1}{y\,\sigma\sqrt{2\pi}} e^{-\frac{(\ln y - \mu)^2}{2\sigma^2}}$ CDF: $F(y) = \Phi\left(\frac{\ln(y) - \mu}{\sigma}\right)$
Inverse Gaussian	PDF: $f(y) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left[-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right],\ \mu > 0, \lambda > 0, x > 0$ CDF: $F(y) = \Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right) + \exp\left(\frac{2\lambda}{\mu}\right) \Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} + 1\right)\right)$
GCTE	PDF: $f(y) = \theta\alpha x^{\alpha-1}\left(\frac{1-p}{(\theta+x)^{\alpha+1}} + \frac{6p}{(2\theta+x)^{\alpha+1}} - \frac{6p}{(3\theta+x)^{\alpha+1}}\right)$ CDF: Not in closed form

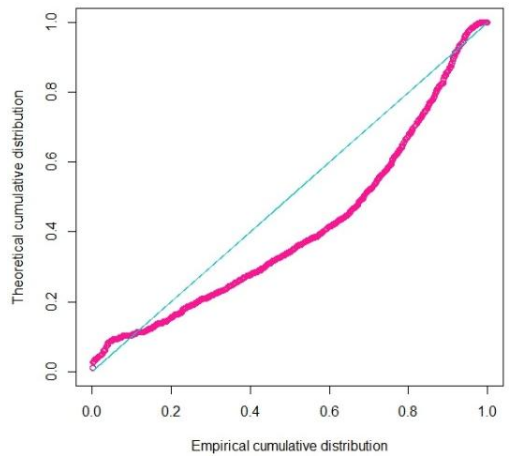

Figure 4. P-P plot of CTE.

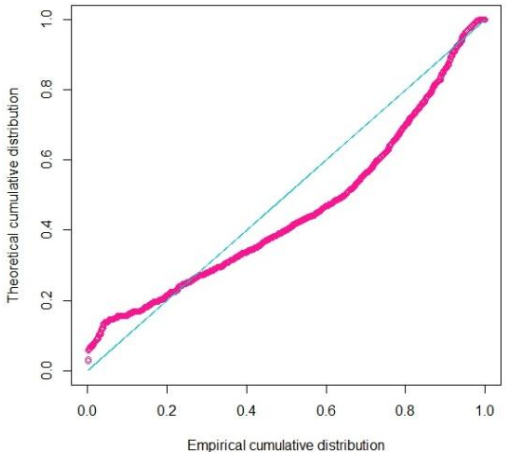

Figure 5. P-P plot of Gamma.

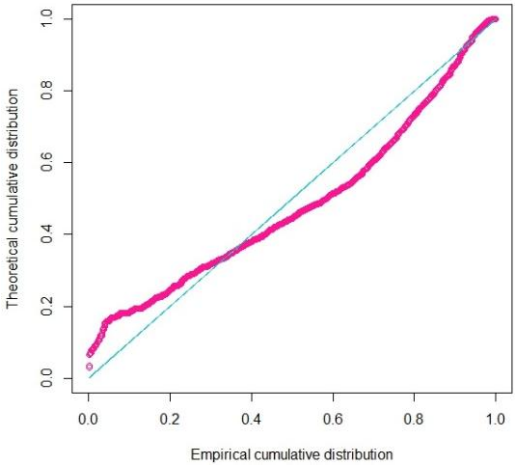

Figure 6. P-P plot of Weibull.

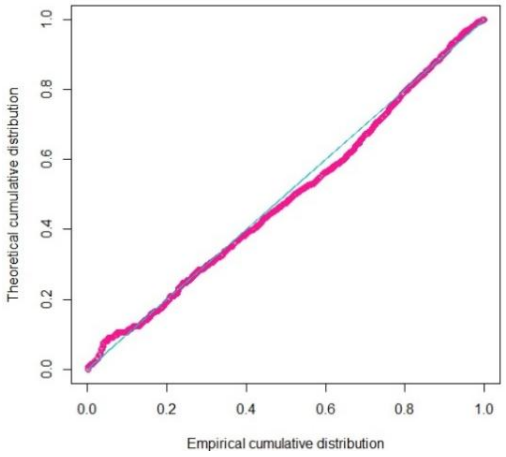

Figure 7. P-P plot of Lognormal.

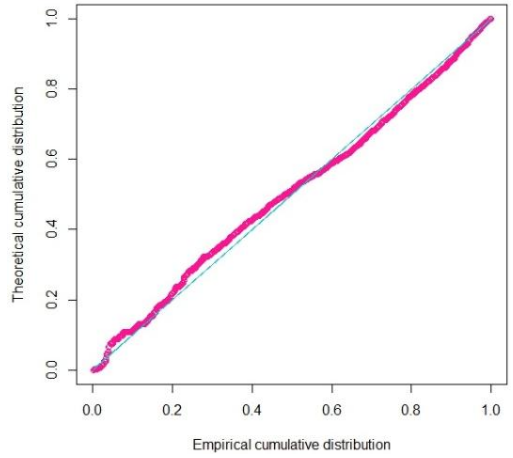

Figure 8. P-P plot of Inverse Gaussian.

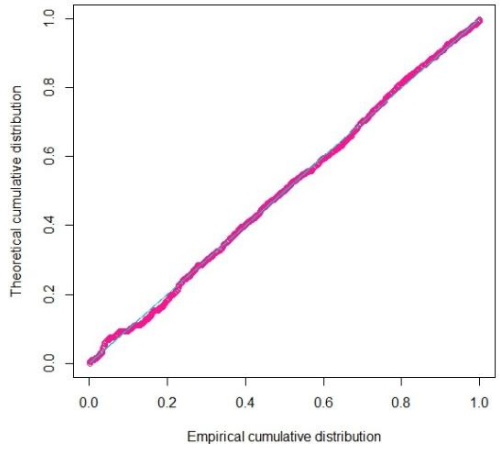

Figure 9. P-P plot of GCTE.

In this study, we analyze the claim severity distributions (or loss distributions) whose PDFs and CDFs are shown in **Table 1**. Their parameters are estimated using the MLE method, and these estimates are presented in **Table 2**. By assessing the goodness of fit using the KS statistic and AD statistic. We found that the values obtained from both test statistics are consistent and indicate results in the same direction.The mixture Lognormal (Sattayatham and Talangtam, 2012). and GCTE distribution had a p-value of greater 0.10, indicating that the null hypothesis could not be rejected for these distributions. For other loss distributions, the null hypothesis was rejected. Therefore, it can be concluded that the data follows the mixture Lognormal distribution and GCTE distribution at a significant level of 0.05

We also considered model selection based on the AIC criterion. Observing that the AIC values for the GCTE distribution are the lowest compared to those of the other distributions, we conclude that the mixture Lognormal and GCTE distributions provide a better fit for this dataset than the CTE, Gamma, Weibull, Lognormal, and Inverse Gaussian distributions. The statistical values obtained from the GCTE distribution are similar to and align with those of the mixture Lognormal distribution, which is formed by combining 40 components of the Lognormal distribution. However, as it does not have a closed form, it is complicated to use. Therefore, we choose the GCTE distribution, which can describe the characteristics of the data similarly to the mixture Lognormal distribution, to model claim severity and analyze ruin probability.

Table 2. Comparing the statistical values of claim severity distributions.

Distributions	Estimated parameters	KS statistic	p-value	AD statistic	p-value	AIC
CTE	$\hat{\theta} = 5.5869\times10^{-5}$ $\hat{p} = -0.1031$	0.1983	< 0.01	82.3545	< 0.01	27941.38
Gamma	$\hat{\alpha} = 0.7385$ $\hat{\beta} = 4.181\times10^{-5}$	0.1474	< 0.01	48.0170	< 0.01	27855.46
Weibull	$\hat{\lambda} = 0.7751$ $\hat{k} = 14420.85$	0.1157	< 0.01	33.0205	< 0.01	27723.14
Lognormal	$\hat{\mu} = 8.9669$ $\hat{\sigma} = 1.1805$	0.0471	< 0.01	3.3667	0.0244	27354.16
Inverse Gaussian	$\hat{\mu} = 17669.4180$ $\hat{\lambda} = 5494.4150$	0.0440	0.0166	3.8224	0.0110	27386.14
Mixture Lognormal (40 components)	-	**0.0217**	**> 0.20**	**0.7989**	**>0.10**	-
GCTE	$\hat{\alpha} = 2.9130$ $\hat{\theta} = 1956.7662$ $\hat{p} = 0.5957$	**0.0239**	**> 0.20**	**0.9667**	**>0.10**	**27335.42**

The probability-probability (P-P) plots of the CTE, Gamma, Weibull, Lognormal, Inverse Gaussian, and GCTE distributions, shown in **Figures 4** to **9**, respectively. The P-P plots illustrate the relationship between the empirical cumulative distribution line and the theoretical cumulative distribution line, represented as a 45-degree line. When these two lines closely overlap, it indicates that the empirical cumulative distribution function aligns with the theoretical cumulative distribution function. This alignment suggests that the sample distribution follows the specified distribution, consistent with the KS statistic value, which reflects the maximum distance between the empirical and theoretical cumulative distribution functions. These results align with the findings presented in **Table 2**.

4. Discrete-Time Risk Process

The discrete-time risk process is a special case of the surplus process $\{U_n\ ;\ n \geq 0\}$ when time is not continuous $\{U_n;\ n = 0,1,2,\dots\}$. Defined by

$$U_n = u + cT_n - \sum_{k=1}^{n} X_k\,,\ n \in \mathbb{N};$$

where, $U_0 = u > 0$ is the initial capital reserve, X_k represents the claim severity random variables at time T_n such that $0 \le T_1 \le T_2 \le \cdots$, and $c > 0$ is the constant premium rate.

Chan and Zhang (2006) presented the discrete-time risk process under the assumption $T_n = n$, is a model used in the insurance industry to study changes in a company's capital reserves over specific time periods, such as months, quarters, or years. The discrete-time risk model provides an approach specifically designed for finite-time ruin probabilities in a discrete setting. This framework is necessary for our study because it aligns well with the periodic nature of our dataset, where claims and premiums are observed at discrete intervals. Unlike continuous models, this discrete-time approach allows for straightforward recursive and explicit formula derivations, enabling us to assess ruin probabilities within finite time horizons accurately. This process considers initial capital reserve, premium rate, claim severity, and the time intervals being analyzed. In this article, we examine the discrete-time risk process is defined as follows:

$$U_n = u + cn - \sum_{k=1}^{n} X_k = U_{n-1} + c - X_n,\ n \in \mathbb{N};$$

where, $U_0 = u > 0$ is the initial capital reserve, X_n represents the claim severity random variables at time n, and $c > 0$ is the daily premium rate, defined as follows:

$$c = (1 + \vartheta)E(X),$$

where, ϑ is the safety loading and $E(X)$ is the expected value of the claim severity over n days.

The main purpose of modeling the discrete-time risk process (surplus process) is to determine whether a portfolio will survive over time ($U_n > 0, n = 0,1,2,...$). Therefore, researchers are interested in modeling it based on a distribution of claim severity, effectively described by the GCTE distribution, within a given time period and safety loading. The finite-time ruin probability at time $k = 1,2,..,n$ is expressed as follows:

$$\phi_n(u) = P(U_k(u) < 0 \text{ for some } k = 0,1,2,\ldots,n | U_0 = u).$$

4.1 Simulation

4.1.1 Finite-Time Ruin Probability

We analyze a dataset of actual motor insurance claims provided by a non-life insurance company in Thailand for the year 2009. **Table 2** indicates that the GCTE distribution is the best fit for describing the severity of these claims. Consequently, we assume that X_n follows the GCTE distribution with parameters α = 2.9130, θ = 1956.7662, and p = 0.5957 for all n. The acceptance-rejection technique, a traditional sampling method, is used to generate samples from distributions that are challenging or impossible to simulate using inverse transformation. The basic idea of acceptance-rejection sampling method is to begin by defining an auxiliary distribution $g(x)$ (which should be a distribution with not complicated, known function, such as a normal or uniform distribution). Next, define constant values c such that $cg(x)$ covers the target distribution $f(x)$. Once the auxiliary distribution satisfies $cg(x) \ge f(x)$. Then samples are drawn x from $g(x)$ and u from a uniform distribution over [0,1]. If $u \le \frac{f(x)}{cg(x)}$ the sample x is accepted; otherwise, it is rejected. This process is repeated until the desired number of accepted samples is obtained as described by Gamerman and Lopes (2006). We show the capital reserve u to be 100000, 250000, 400000, 550000, 700000 and 800000 Thai bath (THB), and consider safety loading values, ϑ, of 0.10, 0.15, 0.20, 0.25, and 0.30. Additionally, we examine time periods of n = 2 years, 3 years, 4 years, and 5 years. The simulation results are obtained using 10000 paths for the surplus process.

Table 3. The approximation of the finite-time ruin probability of the surplus process from simulation with n = 2 years

ϑ	u = 100000	u = 250000	u = 400000	u = 550000	u = 700000	u = 800000
0.10	0.5420	0.2635	0.1090	0.0386	0.0127	0.0053
0.15	0.4347	0.1746	0.0598	0.0182	0.0050	0.0016
0.20	0.3569	0.1173	0.0339	0.0093	0.0015	0.0005
0.25	0.3014	0.0852	0.0214	0.0052	0.0008	0.0002
0.30	0.2557	0.0640	0.0147	0.0028	0.0004	0.0001

Table 4. The approximation of the finite-time ruin probability of the surplus process from simulation with n = 3 years.

ϑ	u = 100000	u = 250000	u = 400000	u = 550000	u = 700000	u = 800000
0.10	0.5467	0.2822	0.1323	0.0562	0.0218	0.0116
0.15	0.4349	0.1870	0.0720	0.0271	0.0091	0.0044
0.20	0.3589	0.1273	0.0439	0.0146	0.0040	0.0020
0.25	0.3004	0.0926	0.0291	0.0087	0.0020	0.0007
0.30	0.2531	0.0706	0.0202	0.0051	0.0011	0.0003

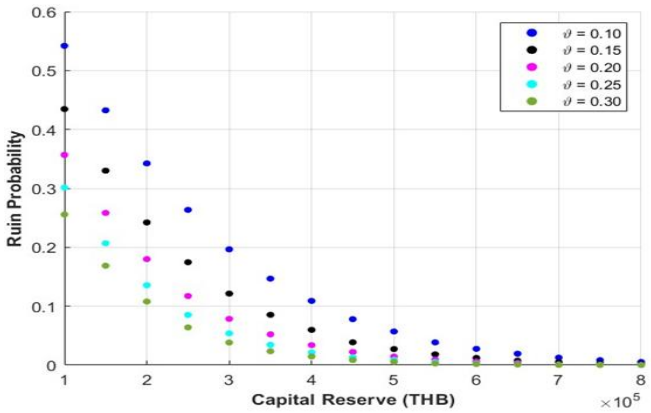

Figure 10. The relation between the finite-time ruin probabilities and capital reserves for n = 2 years.

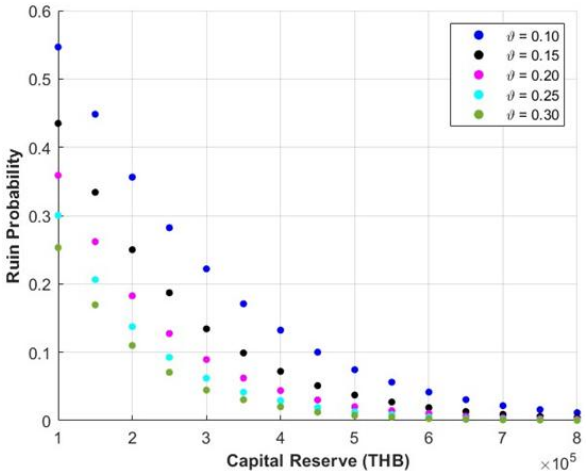

Figure 11. The relation between the finite-time ruin probabilities and capital reserves for n = 3 years.

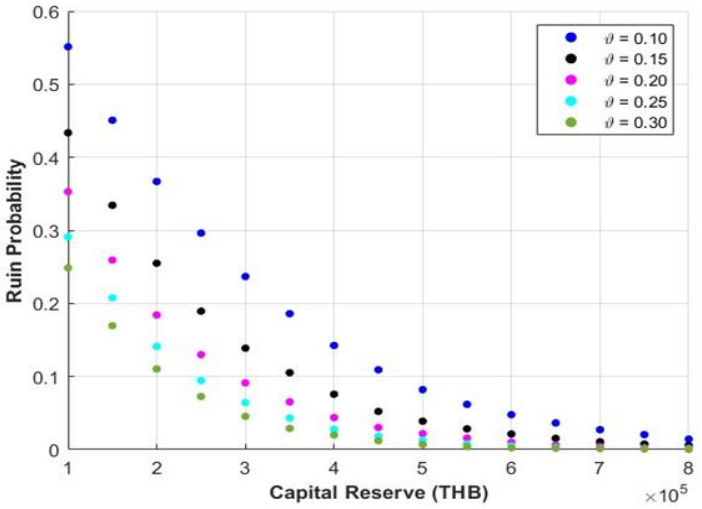

Figure 12. The relation between the finite-time ruin probabilities and capital reserves for n = 4 years.

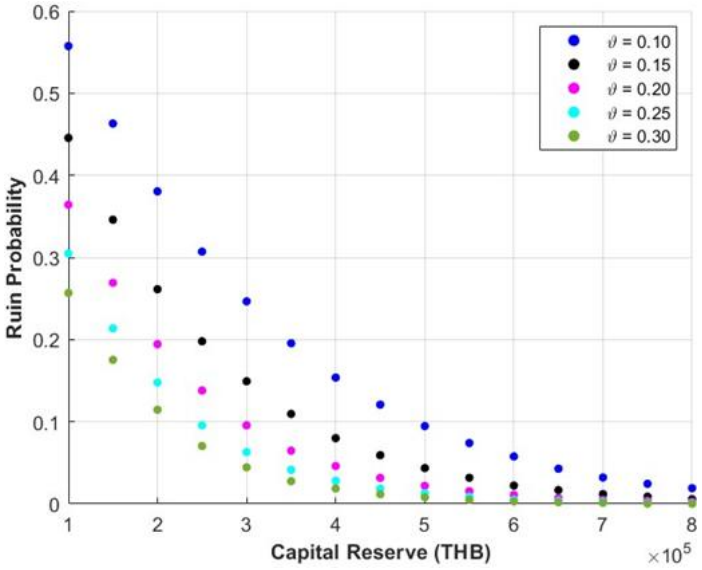

Figure 13. The relation between the finite-time ruin probabilities and capital reserves for n = 5 years.

Table 5. The approximation of the finite-time ruin probability of the surplus process from simulation with n = 4 years.

ϑ	u = 100000	u = 250000	u = 400000	u = 550000	u = 700000	u = 800000
0.10	0.5513	0.2963	0.1425	0.0619	0.0272	0.0143
0.15	0.4335	0.1894	0.0757	0.0283	0.0107	0.0057
0.20	0.3529	0.1299	0.0438	0.0157	0.0042	0.0018
0.25	0.2911	0.0945	0.0277	0.0079	0.0018	0.0010
0.30	0.2487	0.0726	0.0200	0.0041	0.0013	0.0002

Table 6. The approximation of the finite-time ruin probability of the surplus process from simulation with n = 5 years.

ϑ	u = 100000	u = 250000	u = 400000	u = 550000	u = 700000	u = 800000
0.10	0.5574	0.3071	0.1537	0.0741	0.0321	0.0191
0.15	0.4455	0.1980	0.0800	0.0318	0.0119	0.0057
0.20	0.3641	0.1380	0.0461	0.0155	0.0049	0.0027
0.25	0.3048	0.0956	0.0282	0.0086	0.0029	0.0008
0.30	0.2568	0.0705	0.0188	0.0049	0.0012	0.0001

The results of the approximation of the finite-time ruin probability corresponding to the capital reserve u and safety loading ϑ are shown in **Tables 3** to **6**. These tables cover scenarios where u = 100000, 250000, 400000, 550000, 700000 and 800000 THB and ϑ = 0.10, 0.15, 0.20, 0.25 and 0.30, for time periods of n = 2 years, 3 years, 4 years, and 5 years, respectively, Consistent with **Figures 10** to **13**, which show the relationship between safety loading, capital reserve, and their effect on the ruin probability, it can be observed that this relationship follows a curvilinear shape resembling an exponential function. For any given safety loading ϑ, increasing the capital reserve u leads to a decrease in the probability of ruin. Similarly, for any given capital reserve u, increasing the safety loading ϑ consistently reduces the probability of ruin. The impact of capital reserve and safety loading on ruin probability is interconnected, with higher values of both leading to lower ruin probabilities. By analyzing these results, insurance companies can better understand how to structure their financial strategies to reduce the risk of ruin over finite periods.

The criteria for setting capital reserves and safety loadings are largely determined by the discretion of each insurance company, as there is no standardized formula. Capital reserve requirements vary based on multiple company-specific factors. Safety loadings, on the other hand, are typically set within the range of 0.2%–0.3%. Higher safety loading increases premiums, potentially impacting the affordability for policyholders, while a lower safety loading can elevate the company's risk of insolvency. Striking a balance in safety loading is therefore crucial to maintain financial stability while remaining competitive in the market.

4.1.2 Empirical Function of Ruin Probability

Based on **Figures 10** to **13**, a relationship is observed between ruin probability and capital reserve, and we consider an exponential function given by

$$\phi_n(u) = \gamma e^{\delta u}.$$

We will transform the data to estimate parameters of the exponential equation, utilizing the relationship between ϕ and u as follows:

$$\phi = \gamma e^{\delta u}.$$

Taking the natural log of both sides yields:

$$\ln\phi = \ln\gamma + \ln e^{\delta u} = \ln\gamma + \delta u.$$

Let $z = \ln\phi, a_0 = \ln\gamma$ and $a_1 = \delta$. Then, we have

$$z = a_0 + a_1 u.$$

When considering linear regression and estimating parameters using the least squares method, the format is as follows:

$$a_1 = \frac{n\sum_i^n u_i z_i - \sum_i^n u_i \sum_i^n z_i}{n\sum_i^n u_i^2 - \left(\sum_i^n u_i\right)^2},$$

$$a_0 = \bar{z} - a_1\bar{u}.$$

Once a_0 and a_1 are found, the original constants of the model $\phi = \gamma e^{\delta u}$ are $\delta = a_1$ and $\gamma = e^{a_0}$, i.e.:

$$\delta = \frac{n\sum_i^n u_i z_i - \sum_i^n u_i \sum_i^n z_i}{n\sum_i^n u_i^2 - \left(\sum_i^n u_i\right)^2},$$

and

$$\gamma = e^{\bar{z} - a_1\bar{u}}.$$

where, u_i is the capital reserve such that u = 100000, 150000, 200000, …, 800000 THB, and ϕ_i is the ruin probability from simulation corresponding to capital reserve u_i, i=1,2,3, …,15, shown in **Figures 10** to**13** for 2 years, 3 years, 4 years, and 5 years, respectively. The results of the R-squared values and parameters of the exponential equation are provided in **Tables 7** to **10** for 2 years, 3 years, 4 years, and 5 years, respectively. The relationship between ruin probabilities and capital reserves is depicted in **Figures 14** to **17**, with exponential equation lines for n = 2 to 5 years, respectively.

Table 7. Parameters and R-squared values of the exponential equation for the simulation with 2 years.

ϑ	Parameters		R^2
	γ	δ	
0.10	1.3259185	-0.0000066	0.9369
0.15	1.2276884	-0.0000079	0.9316
0.20	1.2294597	-0.0000093	0.8879
0.25	1.1032493	-0.0000102	0.9058
0.30	1.0304146	-0.0000110	0.8948

Table 8. Parameters and R-squared values of the exponential equation for the simulation with 3 years.

ϑ	Parameters		R^2
	γ	δ	
0.10	1.1107227	-0.0000055	0.9748
0.15	0.9505879	-0.0000066	0.9854
0.20	0.8301759	-0.0000075	0.9924
0.25	0.7862434	-0.0000085	0.9855
0.30	0.7299839	-0.0000093	0.9814

Table 9. Parameters and R-squared values of the exponential equation for the simulation with 4 years.

ϑ	Parameters		R^2
	γ	δ	
0.10	1.0652859	-0.0000052	0.9789
0.15	0.8873703	-0.0000063	0.9925
0.20	0.8567560	-0.0000075	0.9823
0.25	0.7629440	-0.0000084	0.9854
0.30	0.7672282	-0.0000095	0.9665

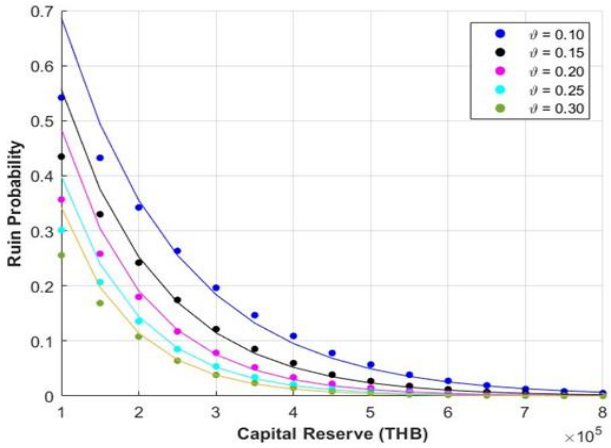

Figure 14. Exponential equation lines for n = 2 years.

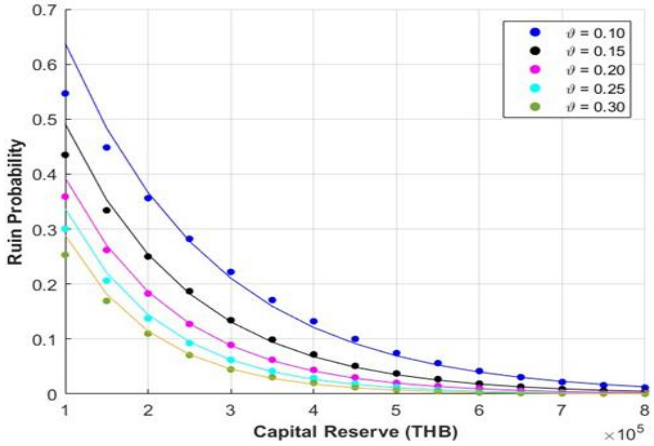

Figure 15. Exponential equation lines for n = 3 years.

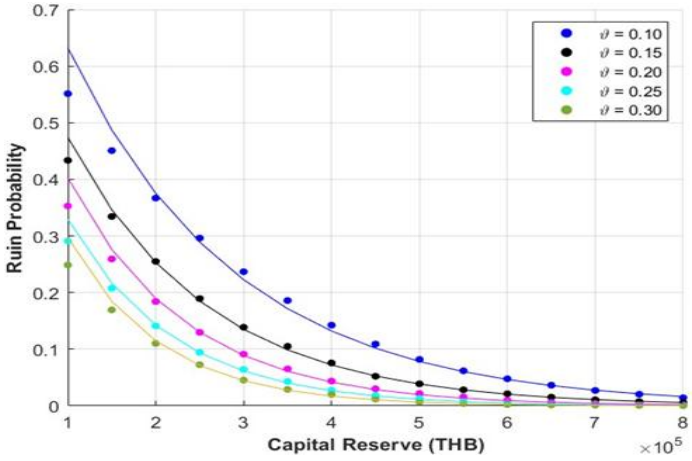

Figure 16. Exponential equation lines for n = 4 years.

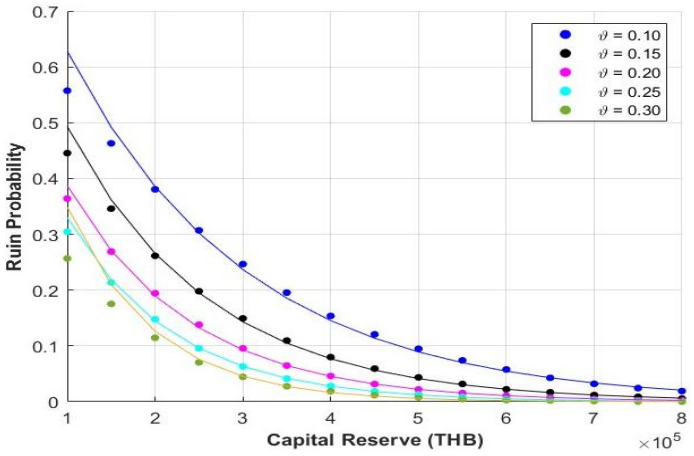

Figure 17. Exponential equation lines for n = 5 years.

Table 10. Parameters and R-squared values of the exponential equation for the simulation with 5 years.

ϑ	Parameters		R^2
	γ	δ	
0.10	1.0228005	-0.0000049	0.9848
0.15	0.9154813	-0.0000062	0.9899
0.20	0.7912944	-0.0000072	0.9965
0.25	0.7532703	-0.0000083	0.9942
0.30	0.9652014	-0.0000102	0.8793

From **Figures 14** to **17**, it is shown that the data characteristics are well-predicted by an exponential function, aligning with the results in **Tables 7** to **10**. The high R^2 indicates that this exponential model, applied to capital reserve and safety loading variables, effectively explains the variation in ruin probability. Consequently, we use this model to estimate the minimum required initial capital reserve based on the risk level the company can accept.

4.1.3 Minimum Capital Reserve

We define the maximum acceptable risk to be represented by the acceptable risk α; thus, the ruin probability should be less than α, namely,

$$\phi(u) \leq \alpha.$$

The capital reserve has to satisfy the inequality,

$$\gamma e^{\delta u} \leq \alpha.$$

Therefore,

$$u \geq \frac{1}{\delta} \ln\left(\frac{\alpha}{\gamma}\right).$$

In situations where the portfolio faces no risk, a sufficiently high premium rate is paid. The capital reserve, denoted by u, may be negative. In such instances, no capital reserve is necessary. Consequently, the minimum capital reserve (MCR) is determined as follows:

$$\text{MCR} = \max\left\{0, \frac{\ln\alpha - \ln\gamma}{\delta}\right\}.$$

Tables 11 to **14** present the minimum capital reserve under varying acceptable risks over periods of 2, 3, 4, and 5 years, respectively. **Figures 18** to **21** depict the relationship between safety loading and minimum capital reserve for ruin probabilities or acceptable risks over 2, 3, 4, and 5 years, respectively. As safety loading ϑ increases, the MCR generally decreases for all levels of risk α. As α decreases, the MCR increases for all levels of ϑ, indicating that a lower acceptable risk requires a higher capital reserve to mitigate ruin probability. Higher ϑ values allow for lower MCR, suggesting that a higher safety loading can reduce the required capital reserve. These results provide insights into structuring optimal capital reserve strategies based on the acceptable risk level and safety loading, helping insurance companies manage their financial risks effectively.

As the time horizon n increases, the MCR generally increases for all levels of risk α and safety loading ϑ. The time horizon has an impact on the required capital reserve, with longer time horizons requiring higher reserves to mitigate insolvency risk. The results in **Tables 11** to **14** show that as the safety loading increases, the MCR decreases. Insurance companies need to consider the time horizon when determining their capital reserve strategies to ensure they are adequately protected against ruin over longer periods. These results highlight the importance of considering the time horizon in determining optimal capital reserve strategies, providing valuable insights for insurance companies to manage their financial risks effectively over different time frames.

Table 11. Minimum capital reserve in the finite-time surplus process and safety loading in relation to acceptable risks in 2 years.

$\alpha = 0.01$		$\alpha = 0.05$		$\alpha = 0.10$	
ϑ	MCR	ϑ	MCR	ϑ	MCR
0.10	742,300.47	0.10	497,852.11	0.10	392,573.93
0.15	608,527.96	0.15	404,925.83	0.15	317,239.17
0.20	516,473.55	0.20	343,722.89	0.20	269,323.23
0.25	462,233.17	0.25	304,064.43	0.25	235,944.86
0.30	420,791.88	0.30	274,682.02	0.30	211,755.93

Table 12. Minimum capital reserve in the finite-time surplus process and safety loading in relation to acceptable risks in 3 years.

$\alpha = 0.01$		$\alpha = 0.05$		$\alpha = 0.10$	
ϑ	MCR	ϑ	MCR	ϑ	MCR
0.10	849,340.50	0.10	559,126.44	0.10	434,138.04
0.15	689,808.86	0.15	446,048.70	0.15	341,066.92
0.20	590,428.19	0.20	375,391.70	0.20	282,780.53
0.25	514,998.47	0.25	325,097.30	0.25	243,311.32
0.30	462,532.51	0.30	289,026.34	0.30	214,301.30

Table 13. Minimum capital reserve in the finite-time surplus process and safety loading in relation to acceptable risks in 4 years.

$\alpha = 0.01$		$\alpha = 0.05$		$\alpha = 0.10$	
ϑ	MCR	ϑ	MCR	ϑ	MCR
0.10	894,737.14	0.10	586,276.05	0.10	453,429.10
0.15	714,987.66	0.15	458,453.76	0.15	347,970.63
0.20	589,540.48	0.20	376,347.74	0.20	284,530.63
0.25	516,151.40	0.25	324,504.26	0.25	241,966.33
0.30	456,772.09	0.30	287,391.31	0.30	214,442.98

Table 14. Minimum capital reserve in the finite-time surplus process and safety loading in relation to acceptable risks in 5 years.

α=0.01		α=0.05		α=0.10	
ϑ	MCR	ϑ	MCR	ϑ	MCR
0.10	948,608.62	0.10	618,699.19	0.10	476,614.93
0.15	729,868.95	0.15	469,803.88	0.15	357,799.95
0.20	611,213.20	0.20	386,163.88	0.20	289,240.42
0.25	523,725.24	0.25	328,691.77	0.25	244,695.43
0.30	449,111.36	0.30	290,937.15	0.30	222,815.23

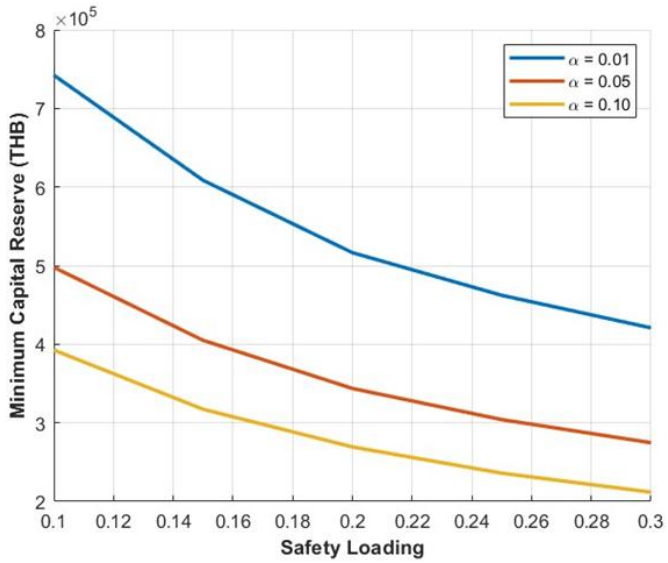

Figure 18. The link between MCR and safety loading under acceptable risks over a period of 2 years.

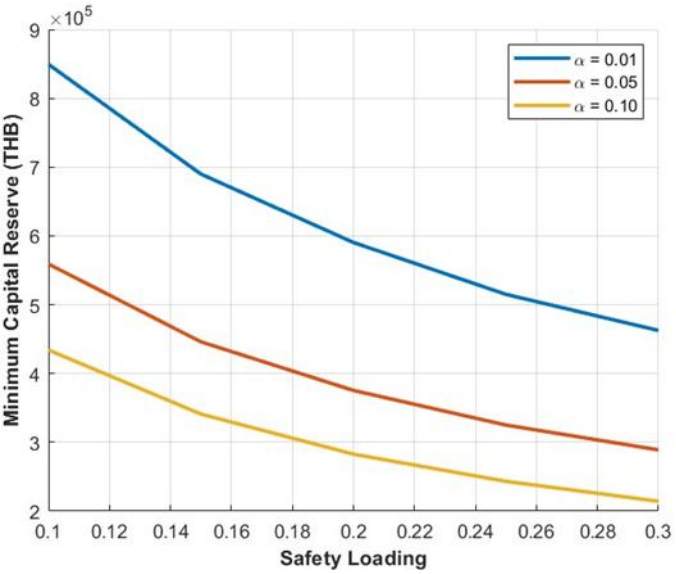

Figure 19. The link between MCR and safety loading under acceptable risks over a period of 3 years.

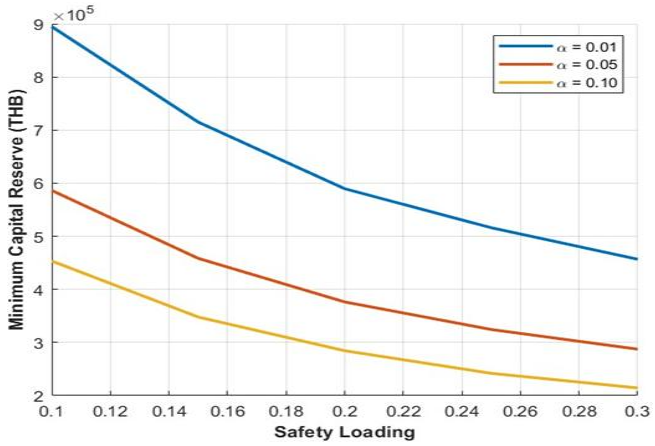

Figure 20. The link between MCR and safety loading under acceptable risks over a period of 4 years.

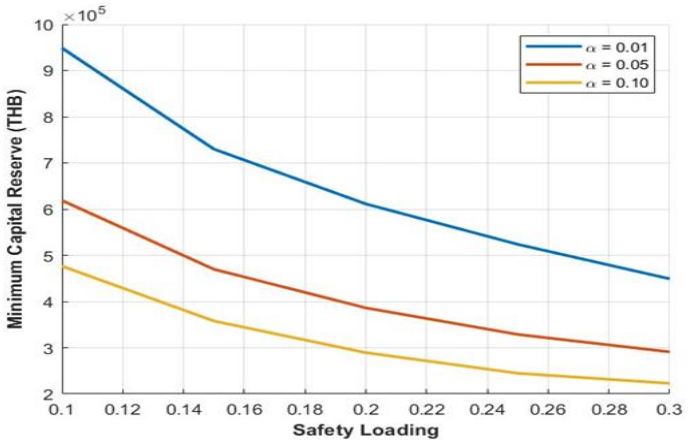

Figure 21. The link between MCR and safety loading under acceptable risks over a period of 5 years.

5. Conclusion and Discussion

The findings of this study indicate that the new distribution (GCTE) was developed using two techniques. The first is the Cubic Transmuted Rank family method, which modifies the distribution's characteristics to increase density in key ranges, enhancing its ability to capture data asymmetry, particularly for loss data with heavy tails. The second is a mixing distributions approach, grounded in Bayesian concepts, where parameters are treated as random variables. The distributions obtained from this mixing approach yield variable failure parameters, unlike traditional distributions that assume constant failure parameters. The results will help the new distributions better describe the heavy-tailed characteristic of insurance. This aligns with previous research by Sattayatham and Talangtam (2012), which found that mixtures of Lognormal distributions across different components provided a better representation of insurance data than traditional distributions alone. When evaluated using the KS and AD tests, the performance of this newly developed distribution is comparable to that of previous studies. Additionally, the distribution presented here has the advantage of a closed form, making it more accessible for interpretation and application. However, in cases where the data has a small right-tailed distribution, this distribution's advantage in describing data characteristics is less pronounced.

While creating a new distribution to better represent the data can indeed improve model fit, it also typically results in a more complex or non-closed form, potentially complicating its practical use. Therefore, selecting a distribution for analysis requires careful consideration of the trade-off between increased accuracy and added complexity in application, which should be weighed based on the analyst's goals and priorities.

The results indicate that the GCTE distribution is more suitable for describing insurance loss data (heavy-tailed data) than traditional loss distributions, such as Gamma, Weibull, Lognormal, and Inverse Gaussian, which are commonly used in insurance, finance and risk-related fields. The GCTE distribution is constructed using the previously mentioned techniques to enhance its flexibility in accurately reflecting the characteristics of insurance data.

In the application section of this study, the GCTE distribution is utilized to create a loss model based on claim severity, which is integral to the insurance risk assessment process. The GCTE distribution aids in evaluating the survival probability of an insurance company within a finite time frame and represents the risk as a probability of loss. Employing this loss distribution enhances the accuracy and reliability of loss probability calculations, providing valuable insights for the insurance industry to develop effective risk management strategies. Additionally, this study's risk model is limited to a finite time frame; however, future research could extend this model to incorporate an infinite time horizon for broader risk assessment.

Conflict of Interest
The authors confirm that there is no conflict of interest to declare for this publication.

AI Disclosure
The author(s) declare that no assistance is taken from generative AI to write this article.

Acknowledgments
The authors sincerely thank the editor and reviewers for their insightful comments and valuable suggestions, which have significantly improved the quality of our manuscript.

References

Adetunji, A.A., & Sabri, S.R.M. (2023). On zero-inflated mixed Poisson transmuted exponential distribution: properties and applications to observation with excess zeros. *Maejo International Journal of Science and Technology*, *17*(01), 68-80.

Chan, W.S., & Zhang, L. (2006). Direct derivation of finite-time ruin probabilities in the discrete risk model with exponential or geometric claims. *North American Actuarial Journal*, *10*(4), 269-279.

Cruz, E., Rincón, L., & Santana, D.J. (2024). Ruin probabilities as recurrence sequences in a discrete-time risk process. *Methodology and Computing in Applied Probability*, *26*(3), 1-16. https://doi.org/10.1007/s11009-024-10102-0.

Frangos, N.E., & Vrontos, S.D. (2001). Design of optimal bonus-malus systems with a frequency and a severity component on an individual basis in automobile insurance. *ASTIN Bulletin: The Journal of the IAA*, *31*(1), 1-22.

Gamerman, D., & Lopes, H.F. (2006). *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. Chapman and Hall/CRC, New York. https://doi.org/10.1201/9781482296426.

Granzotto, D.C.T., Louzada, F., & Balakrishnan, N. (2017). Cubic rank transmuted distributions: inferential issues and applications. *Journal of Statistical Computation and Simulation*, *87*(14), 2760-2778.

Gupta, R.D., & Kundu, D. (1999). Theory & methods: generalized exponential distributions. *Australian & New Zealand Journal of Statistics, 41*(2), 173-188. https://doi.org/10.1111/1467-842x.00072.

Ieosanurak, W., Khomkham, B., & Moumeesri, A. (2023). Claim modeling and insurance premium pricing under a bonus–malus system in motor insurance. *International Journal of Applied Mathematics and Computer Science, 33*(4), 637-650. https://doi.org/10.34768/amcs-2023-0045.

Klongdee, W., Sattayatham, P., & Boonta, S. (2013). On approximating the ruin probability and the minimum initial capital of the finite-time risk process by separated claim technique of motor insurance. *Far East Journal of Mathematical Sciences, 2013*(IV), 597-604.

Liu, R., Wang, D., & Guo, F. (2018). The finite-time ruin probability of a discrete-time risk model with GARCH discounted factors and dependent risks. *Communications in Statistics-Theory and Methods, 47*(17), 4170-4186.

Mert, M., & Saykan, Y. (2005). On a bonus-malus system where the claim frequency distribution is geometric and the claim severity distribution is Pareto. *Hacettepe Journal of Mathematics and Statistics, 34*(1), 75-81.

Moumeesri, A., & Pongsart, T. (2022). Bonus-malus premiums based on claim frequency and the size of claims. *Risks, 10*(9), 181. https://doi.org/10.3390/risks10090181.

Moumeesri, A., Klongdee, W., & Pongsart, T. (2020). Bayesian bonus-malus premium with Poisson-Lindley distributed claim frequency and Lognormal-Gamma distributed claim severity in automobile insurance. *WSEAS Transactions on Mathematics, 19*(46), 443-451. https://doi.org/10.37394/23206.2020.19.46.

Mudholkar, G.S., Srivastava, D.K., & Kollia, G.D. (1996). A generalization of the Weibull distribution with application to the analysis of survival data. *Journal of the American Statistical Association, 91*(436), 1575-1583.

Ni, W., Constantinescu, C., & Pantelous, A.A. (2014). Bonus–Malus systems with Weibull distributed claim severities. *Annals of Actuarial Science, 8*(2), 217-233. https://doi.org/10.1017/S1748499514000062.

Pongsart, T., Moumeesri, A., Mayureesawan, T., & Phaphan, W. (2021). Computing Bayesian bonus-malus premium distinguishing among different multiple types of claims. *Lobachevskii Journal of Mathematics, 42*(13), 3208-3217. https://doi.org/10.1134/s1995080222010164.

Rahman, M.M., Al-Zahrani, B., Shahbaz, S.H., & Shahbaz, M.Q. (2019). Cubic transmuted uniform distribution: An alternative to beta and kumaraswamy distributions. *European Journal of Pure and Applied Mathematics, 12*(3), 1106-1121. https://doi.org/10.29020/nybg.ejpam.v12i3.3410.

Santana, D.J., & Rincón, L. (2020). Approximations of the ruin probability in a discrete time risk model. *Modern Stochastics: Theory and Applications*, *7*(3), 221-243. https://doi.org/10.15559/20-vmsta158.

Sattayatham, P., & Talangtam, T. (2012). Fitting of finite mixture distributions to motor insurance claims. *Journal of Mathematics and Statistics*, *8*(1), 49-56.

Sattayatham, P., Sangaroon, K., & Klongdee, W. (2013). Ruin probability-based initial capital of the discrete-time surplus process. *Variance, 7*(1), 74-81.

Satterthwaite, F.E. (1942). Generalized Poisson distribution. *The Annals of Mathematical Statistics*, *13*(4), 410-417.

Shaw, W.T., & Buckley, I.R. (2009). The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *Computational Finance.* arXiv preprint arXiv:0901.0434.