JURNAL STATISTIKA DAN APLIKASINYA

e-ISSN 2620 - 8369
VOLUME 10 (ISSUE 1) JUNE 2026: 01-16
https://doi.org/10.21009/JSA.10101

MODELING THE AGGREGATE LOSS DISTRIBUTION IN MOTOR THIRD-PARTY LIABILITY INSURANCE USING MONTE CARLO SIMULATION

Tajmahal Ghaza Antoni[1], Nur Rahmadani Ahmad[2], Viery Salsaputra Triana[3], Gemala Azzahra Ocan[4], Ruhiyat[5*]

[1,2,3,4,5]*Division of Economic, Financial, and Actuarial Mathematics, School of Data Science, Mathematics, and Informatics, IPB University, Jalan Meranti, Kampus IPB Dramaga, Bogor, 16680, Indonesia*

Corresponding author's e-mail: * ***ruhiyat-mat@ipb.ac.id***

ABSTRACT

Article History:

Received: January 27, 2026
Revised: June 12, 2026
Accepted: June 12, 2026
Published: June 30, 2026
Available online.

Keywords:

Burr XII Severity Model;
Claim Frequency Modeling;
Heavy-Tailed Insurance Data;
Value at Risk Estimation;
Z123M Negative Binomial.

This study models the aggregate loss distribution for motor third-party liability insurance using Monte Carlo simulation. Aggregate loss estimation is essential because it depends on claim frequency and severity, which often exhibit overdispersion and heavy tails, making analytical solutions intractable and motivating simulation-based approaches for accurate tail-risk assessment. The objective of this study is to identify appropriate distributions for frequency and severity using French Motor Third-Party Liability (MTPL) insurance data and to construct the aggregate loss distribution through Monte Carlo simulation. The modeling procedure involves distribution selection, goodness-of-fit assessment using Chi-Square and Kolmogorov-Smirnov tests, graphical comparison, and model evaluation using the Akaike Information Criterion (AIC). The selected distributions are then combined to generate simulated aggregate losses, from which Value at Risk (VaR) and Tail Value at Risk (TVaR) are computed. The results show that the Zero-One-Two-Three Modified Negative Binomial (Z123M-NB) distribution provides the best fit for claim frequency, while the Burr XII distribution effectively represents claim severity. Monte Carlo simulation with 10 million iterations produces stable estimates of the aggregate loss mean and variance, and the estimated VaR at the 95%, 97.5%, and 99% confidence levels are 105.85, 1,506.61, and 3,629.14, with corresponding TVaR values of 4,122.93, 7,418.70, and 15,075.21, indicating substantial tail heaviness. The study is limited by the sensitivity of variance estimation under extreme severity values and the assumption of a continuous severity model. The novelty of this study lies in integrating the Z123M-NB frequency model with Burr XII severity within a Monte Carlo framework for real MTPL data, offering enhanced flexibility in modeling extreme aggregate losses.

How to cite this article:
Antoni, T. G., Ahmad, N. R., Triana, V. S., Ocan, G. A., & Ruhiyat, "MODELING THE AGGREGATE LOSS DISTRIBUTION IN MOTOR THIRD-PARTY LIABILITY INSURANCE USING MONTE CARLO SIMULATION", Journal Statistika dan Aplikasinya, vol. 10, iss. 1, pp. 01 - 16, June 2026

Journal homepage: https://journal.unj.ac.id/unj/index.php/statistika
Journal e-mail: jsa@unj.ac.id
Research Article · Open Access

1. INTRODUCTION

The insurance industry is one of the key sectors in the modern financial system that provides protection against financial uncertainty risks that may cause economic losses [1]. One of the main branches of this industry is motor third-party liability (MTPL) insurance, which provides compensation for physical damage or bodily injury caused by traffic accidents [2]. In many countries, including Europe, this product is mandatory for all motor vehicle owners, aiming to maintain economic stability for individuals as well as insurance companies [3].

In actuarial practice, one of the main focuses is estimating aggregate losses, which represent the total amount of loss incurred within a specific period. The magnitude of aggregate losses is influenced by two main components: claim frequency (the number of claims) and claim severity (the monetary value of claims paid) [4]. Mathematically, the aggregate loss variable can be modeled as a random sum of independent individual claims [1]. Therefore, the aggregate loss distribution generally does not admit a simple analytical form, requiring a simulation-based approach to obtain accurate estimates of tail risk [3].

One widely used method to estimate aggregate loss distribution is the Monte Carlo simulation. This approach allows actuaries to simulate various possible loss outcomes based on the probabilistic distributions of claim frequency and severity [4]. The main advantage of this method lies in its flexibility for use with complex distributions lacking analytical solutions and its ability to estimate extreme loss probabilities, which are critical for insurance risk management [3]. [4] demonstrated that the Monte Carlo approach is highly effective in computing tail probabilities for aggregate distributions under the natural exponential families (NEF) assumption, particularly when claim distributions exhibit overdispersion and heavy-tailed characteristics.

The data used in this study come from the freMTPL2freq and freMTPL2sev datasets, which are publicly available through the freMTPL2 repository on Hugging Face. These datasets contain information on third-party liability insurance claims for motor vehicles in France and have been widely used in actuarial research to investigate claim frequency and severity patterns [1]. The freMTPL2freq dataset exhibits typical characteristics of motor insurance claim data, including a large proportion of policyholders with zero claims and overdispersion in claim frequencies, where the variance exceeds the mean. These characteristics indicate that standard frequency distributions may not adequately capture the observed claim frequency behavior. Therefore, a flexible discrete distribution, namely the Zero-One-Two-Three Modified Negative Binomial (Z123M-NB) distribution, is considered appropriate because it allows additional probability adjustments for low claim frequencies while maintaining the overdispersion property of the negative binomial framework. In this study, claim frequency and severity are modeled separately before being combined into the aggregate loss simulation using the Monte Carlo method, where severity is represented by the Burr Type XII (Burr XII) distribution and frequency by the Z123M-NB distribution [4].

In modern actuarial contexts, modeling claim frequency and severity separately and then combining them into a single aggregate loss distribution framework is known as the collective risk model [1]. This model is essential for calculating Value at Risk (VaR) and Tail Value at Risk (TVaR), which are key measures in determining minimum capital requirements and technical reserves for insurance companies [4]. As data complexity and accuracy requirements increase, Monte Carlo simulation has become increasingly relevant for producing more realistic results than deterministic methods.

Previous studies have applied Monte Carlo simulation and stochastic modeling methods in insurance risk modeling. [1] developed a stochastic model to assess price distortion in motor insurance. [2] emphasized the importance of real datasets in improving risk model accuracy. [4] used the Monte Carlo approach to estimate aggregate loss distributions. The Monte Carlo method has proven effective for modeling complex and realistic aggregate losses.

Based on the background, this study aims to construct an aggregate loss distribution model using the Monte Carlo simulation by utilizing insurance claim data from the freMTPL2freq and freMTPL2sev datasets. Through this simulation, it is expected to obtain a more accurate representation of total loss distribution and provide a foundation for insurance companies in developing appropriate risk management strategies and premium pricing [1]. This study also aims to contribute to actuarial literature in applying simulation methods to real-world non-life insurance data, particularly MTPL insurance [2].

2. METHODS

Material and Data

The data used to model the distribution of aggregate loss were obtained from the French MTPL dataset, accessible through the freMTPL2 repository on Hugging Face. Risk features were collected for 678,013 motor third-party liability policies, observed mostly over one year, across different regions in France, prior to 2016 [5]. The dataset provides both the frequency of claims (number of claims per policyholder) and the severity of claims (amount of loss per claim), making it suitable for modeling the aggregate loss distribution in insurance applications.

Research Method

The study followed the statistical modeling procedure proposed by Kaishev and Krachunov for fitting probability distributions to insurance claim data. First, candidate distribution families were selected for both claim frequency and claim severity. Second, the parameters of the candidate distributions were estimated using the maximum likelihood estimation (MLE) method. Third, the goodness-of-fit of each fitted distribution was evaluated using the Chi-Square test for claim frequency and the Kolmogorov-Smirnov test for claim severity [6]. The candidate models that satisfied the goodness-of-fit criteria were subsequently compared using the Akaike Information Criterion (AIC) to identify the most suitable distributions. The AIC is calculated using the following formula:

$$\text{AIC} = 2k - 2\ln\left(\hat{L}\right)$$

where k denotes the number of model parameters and $\hat{L}$ denotes the maximum likelihood of the estimated model. The model with the smallest AIC value is selected as the most suitable for representing both claim frequency and claim severity.

After selecting the best-fitting claim frequency and claim severity models, the aggregate loss distribution was generated using Monte Carlo simulation. Monte Carlo simulation is a numerical method that uses random numbers to represent random events within a process [7]. In each iteration, the number of claims and the corresponding claim severities were randomly generated from the selected distributions, and their sum represented one realization of aggregate loss. Let N denotes the claim frequency and $X_1, X_2, \dots, X_N$ denote independent and identically distributed random variables representing individual claim severities. The independence assumption implies that the amount of one claim does not influence the amount of another claim. The aggregate severity for N Simulated claims are then defined as $S = X_1 + X_2 + \cdots + X_N$. Furthermore, within the collective risk model framework, the claim frequency (N) is assumed to be independent of the individual claim severities (X_i), meaning that the number of claims occurring during a period does not affect the distribution of the claim amounts. Under these assumptions, the sample mean estimator is expressed as $\bar{X}_N = \frac{S}{N}$. If the simulation process is repeated m times, producing independent realizations $\bar{X}_{N_1}, \bar{X}_{N_2}, \dots, \bar{X}_{N_m}$, the overall sample mean estimator is given by $\bar{X} = \frac{1}{m}\sum_{i=1}^{m} \bar{X}_{N_i}$. This estimator converges to the expected value ($E[\bar{X}]$) [8]. Repeating the simulation multiple times produced an empirical approximation of the aggregate loss distribution [9].

Prior to performing the full simulation, an appropriate number of iterations n was determined through a convergence analysis to ensure numerical stability. The simulation was repeated for increasing values of n while monitoring the stabilization of key outputs, including the expected aggregate loss, Value at Risk (VaR), and Tail Value at Risk (TVaR), as well as the relative error of the estimated expected aggregate loss and the variance of aggregate loss. The optimal value of n was selected when these quantities exhibited sufficiently small relative errors and successive increases in

n produced negligible changes, indicating convergence and reliable simulation accuracy. Finally, the insurer's risk exposure was assessed using VaR and TVaR derived from the simulated aggregate loss distribution [9]. All statistical analyses and simulations were performed using R software.

3. RESULTS

Distribution Models for Claim Frequency

Descriptive Statistics

The claim frequency data used in this study were obtained from the French MTPL insurance dataset, specifically the freMTPL2freq dataset. The dataset contains policy-level information with a total of 678,013 motor insurance policies. This dataset provides the number of claims reported for each policy within the observation period. The following section presents the statistical description of the claim frequency data derived from the dataset, as summarized in Table 1.

Table 1. Descriptive statistics of claim frequency in French MTPL insurance

Minimum	Maximum	Mean	Variance	Skewness	Kurtosis
0	16	0.0533	0.0579	5.5996	79.8413

The positive skewness value indicates that the claim frequency distribution is strongly right-skewed, characterized by a long tail extending toward higher claim frequencies. This pattern reflects the predominance of policyholders who did not submit any claims during the observation period, with only a relatively small proportion reporting one or more claims. Furthermore, the kurtosis value, which is substantially greater than three, signifies a leptokurtic distribution with a pronounced peak and heavy tails. Such characteristics are typical in motor insurance datasets, where observations are heavily concentrated at zero claims while a limited number of higher claim occurrences contribute to the tail heaviness. The asymmetrical shape observed in the histogram, as shown in Figure 1, further supports the conclusion that the distribution deviates markedly from normality.

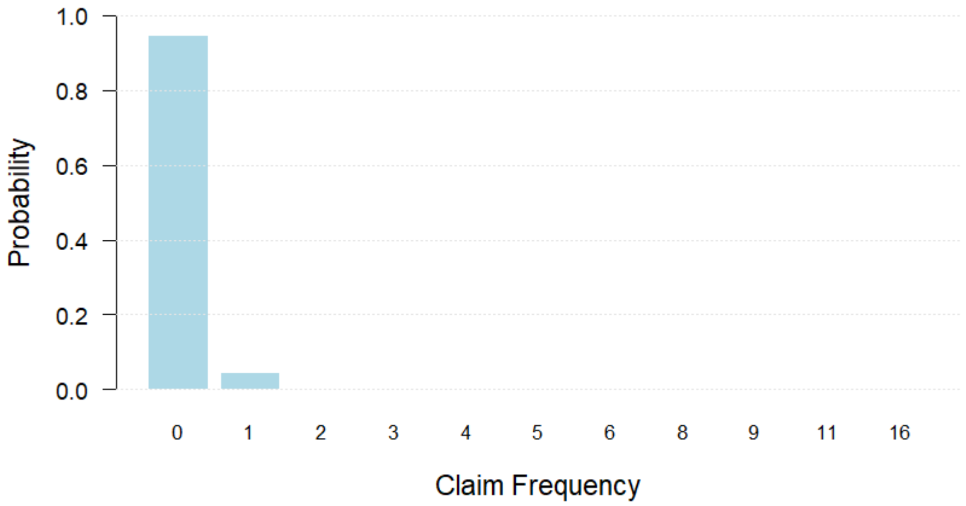

Figure 1. Histogram of claim frequency

Probability Mass Function and Parameter Estimation

Based on the statistical description of the dataset, the negative binomial distribution was selected because the variance exceeded the mean, indicating overdispersion in the claim frequencies. For comparison, the Poisson distribution was also considered since it represents the standard model for frequency data with equal mean and variance. Parameter estimation for both distributions was performed using the MLE method, and the corresponding probability mass function (PMF) for the negative binomial and Poisson models was evaluated using the estimated parameters. The complete parameter estimates are presented in Table 2.

Table 2. PMF and the estimated parameter values of the negative binomial and Poisson

Distribution	PMF	Parameter estimates
Negative binomial	$f_{NB}(k) = P(X = k) = \binom{k+r-1}{k}\left(\frac{r}{r+\mu}\right)^{r}\left(\frac{\mu}{r+\mu}\right)^{k},$ $k = 0,1,2,\ldots,\ r > 0,\ \mu > 0$	$\hat{r} = 0.77833$ $\hat{\mu} = 0.05323$
Poisson	$f_{Pois}(k) = P(X = k) = \frac{\lambda^{k} e^{-\lambda}}{k!}, \quad k = 0,1,2,\ldots,\ \lambda > 0$	$\hat{\lambda} = 0.05325$

Furthermore, these modified structures provide an enhanced capability to model datasets characterized by substantial probability mass at lower claim frequencies, which is a common feature in motor insurance portfolios. By explicitly adjusting the probabilities at the specified frequency levels, the Zero-One-Two-Three Modified negative binomial (Z123M-NB) and the Zero-One-Two-Three Modified Poisson (Z123M-Poisson) offer greater flexibility in capturing departures from the standard parametric forms. The PMF for each modified model is presented in Table 3, outlining the full density specifications used in the subsequent analysis.

Table 3. PMF the estimated parameter values of the Z123M-NB and Z123M-Poisson

Distribution	PMF	Parameter estimates
Z123M-NB	$P(X = k) = p_0$ for $k = 0$; p_1 for $k = 1$; p_2 for $k = 2$; p_3 for $k = 3$; $(1 - p_0 - p_1 - p_2 - p_3)\, f_{NB}(k)$ for $k \geq 4$.	$\hat{p}_0 = 0.94977$ $\hat{p}_1 = 0.04746$ $\hat{p}_2 = 0.00263$ $\hat{p}_3 = 0.00012$ $\hat{r} = 9.39728$ $\hat{\mu} = 6.87478$
Z123M-Poisson	$P(X = k) = p_0$ for $k = 0$; p_1 for $k = 1$; p_2 for $k = 2$; p_3 for $k = 3$; $(1 - p_0 - p_1 - p_2 - p_3)\, f_{Pois}(k)$ for $k \geq 4$.	$\hat{p}_0 = 0.94977$ $\hat{p}_1 = 0.04746$ $\hat{p}_2 = 0.00263$ $\hat{p}_3 = 0.00012$ $\hat{\lambda} = 6.875$

Graphical Model Analysis

Before presenting the numerical evaluation of the fitted frequency models, a graphical comparison is provided to visually examine how each candidate distribution represents the empirical pattern of claim frequencies. This visual assessment is useful for identifying the overall shape, concentration at lower frequencies, and tail behavior captured by each model prior to analyzing the statistical goodness-of-fit measures. Figure 2 displays the histogram of the observed claim frequencies alongside the fitted PMF of the four selected models.

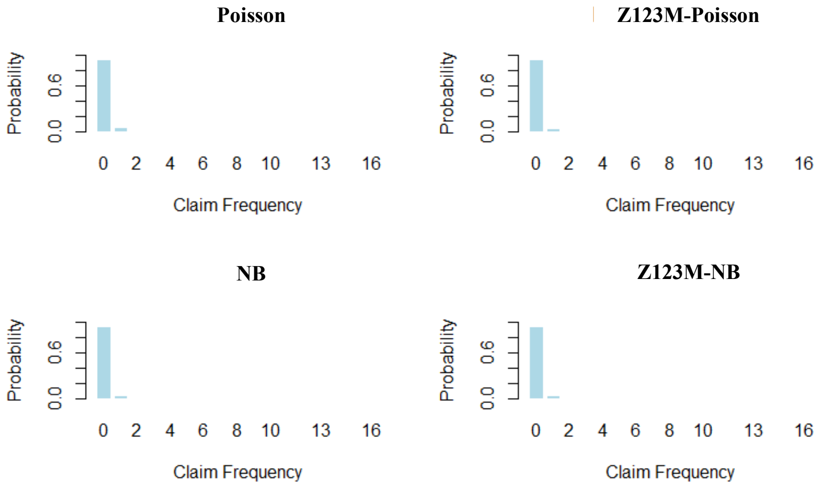

Figure 2. **Histogram of claim frequency with fitted PMF from negative binomial, Z123M-NB, Poisson, and Z123M-Poisson**

Figure 2 presents the graphical comparison between the empirical claim frequency histogram and the fitted PMF of the four candidate models: the negative binomial distribution, the Poisson distribution, the Z123M-NB, and the Z123M-Poisson. Each subplot illustrates how closely the theoretical probability structure of the corresponding model aligns with the observed pattern of claim frequencies. These graphical displays provide an initial visual assessment of how well each distribution captures the characteristics of the claim frequency data before further statistical evaluation is performed.

Chi-Square Test

The claim frequency data were evaluated using four distributions: negative binomial (NB), Poisson, Z123M-NB, and Z123M-Poisson. The chi-square test was conducted using the following hypotheses:

H_0: The claim frequency data follow the specified distribution.
H_1: The claim frequency data do not follow the specified distribution.

The expected claim frequencies obtained from each distribution are summarized in Table 4.

Table 4. Expectations of the claim frequency from the NB distribution, Poisson distribution, Z123M-NB distribution, and Z123M-Poisson distribution

k	Claim frequency	NB	Poisson	Z123M-NB	Z123M-Poisson
0	643,953	643,987	642,855	643,953	643,953
1	32,178	32,085	34,230	32,178	32,178
2	1,784	1,826	911	1,784	1,784
3	82	108	16	82	82
4	7	7	0	2	2
5	2	0	0	2	2
6	1	0	0	2	3
7	0	0	0	2	3
8	1	0	0	2	2
9	1	0	0	2	2
10	0	0	0	1	1
11	3	0	0	1	1

k	**Claim frequency**	**NB**	**Poisson**	**Z123M-NB**	**Z123M-Poisson**
12	0	0	0	1	0
13	0	0	0	0	0
14	0	0	0	0	0
15	0	0	0	0	0
16	1	0	1	0	0
Total	678,013	678,013	678,013	678,013	678,013
	Chi-square	19.30349	2,372.981	0.28676	0.19919
	Degrees of freedom	2	3	2	3
	p-value	0	6.43131×10^{-5}	0.90520	0.86642

Based on the results in Table 4, the Poisson distribution produced a chi-square value of 2,372.981 with a p-value of 6.43×10^{-5}, indicating a substantial deviation from the observed frequencies. Therefore, H_0 is rejected, and the Poisson distribution is not suitable for the claim data. The NB distribution performs better than Poisson but still shows noticeable discrepancies, with a chi-square value of 19.30349 and a p-value of 0. Since the p-value < 0.05, H_0 is rejected, meaning the NB distribution is also not appropriate.

In contrast, the two modified distributions, Z123M-NB and Z123M-Poisson, exhibit excellent fit. They yield very small chi-square values (0.28676 and 0.19919, respectively) and high p-values (0.90520 and 0.86642, respectively). Since the p-values are far above 0.05, the null hypothesis is not rejected for either model, indicating that both distributions adequately represent the observed claim frequency data. Therefore, both models were retained as candidate frequency distributions for further evaluation. The final model selection was subsequently based on the Akaike Information Criterion (AIC) and the theoretical suitability of the distribution for overdispersed claim frequency data.

Model Selection via Akaike Information Criterion (AIC)

After confirming that the Z123M-NB and Z123M-Poisson models passed the chi-square goodness-of-fit test, the next step was to compare their overall statistical performance using the Akaike Information Criterion (AIC). Lower AIC values indicate a better balance between model fit and model complexity. The results of the AIC calculations for both models are presented in Table 5.

Table 5. Log-Likelihood and AIC Values of the Z123M-NB and Z123M-Poisson Models

	Z123M-NB	**Z123M-Poisson**
Log-likelihood	-142,810.7	-142,813.7
AIC	285,625.4	285,629.3

Based on the AIC values, the Z123M-NB model yields a lower value than the Z123M-Poisson model, indicating a better overall fit to the observed claim frequency data. Since both models successfully passed the chi-square goodness-of-fit test, the final model selection was based on a combination of statistical performance and theoretical suitability. Given that the claim frequency data exhibit overdispersion, as indicated by a variance exceeding the mean, the negative binomial framework is more appropriate than the Poisson framework [10]. Therefore, the Z123M-NB model was selected as the frequency model for the subsequent aggregate loss simulation because it provides the best overall fit according to the AIC criterion while preserving the ability to accommodate overdispersion in the claim frequencies.

Distribution Models for Claim Severity

Descriptive Statistics

The data used to model claim severity consists of claim amounts in the French MTPL severity dataset. The descriptive statistics of this dataset are presented in Table 6.

Table 6. Descriptive statistics of claim amount in French MTPL insurance

Mean	Variance	Skewness	Kurtosis
4,678.241	2,835,934,146	61.15125	4,422.063

By removing several uppermost values so that outliers do not dominate the visualization, the histogram of the lower 95% of the data is obtained, as shown in Figure 3. The distribution is concentrated around small to medium claim amounts, indicating that most claims are of low-cost [11]. The right tail of the distribution is noticeably long (positively skewed), suggesting the presence of high-cost claims, although they occur infrequently [12]. The high kurtosis value further indicates the existence of significant outliers and a pronounced peak in the distribution.

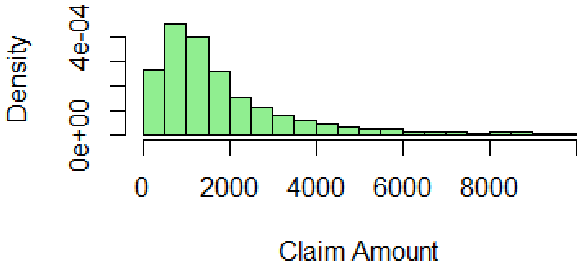

Figure 3. Histogram of claim amount (95% Data)

The histogram plotted on a logarithmic scale, presented in Figure 4, reveals a distribution that is considerably more symmetric than the original scale. This transformation reduces the influence of extreme values and shortens the right tail, causing the mean and median to lie closer to the center of the data. Visually, the distribution of the log-transformed data appears approximately normal due to its symmetry. This provides evidence that the original claim severity data may follow a lognormal distribution, consistent with the common assumption that the logarithm of cost-type data is normally distributed or close to normal [13]. Therefore, this study will evaluate the lognormal model for claim severity. In addition, Burr XII and log-logistic models will also be tested because both distributions are flexible heavy-tailed models commonly used in actuarial severity modeling and are capable of capturing extreme losses.

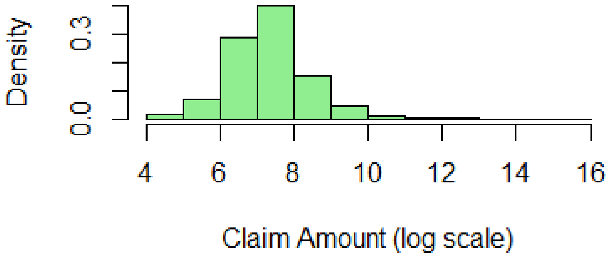

Figure 4. Histogram of claim amount on a logarithmic scale

Probability Density Function and Parameter Estimation

Parameter estimation for the lognormal, Burr XII, and log-logistic models in the claim severity analysis was carried out using the MLE method. The probability density function (PDF) of the distributions, sourced from the book [14], along with the corresponding estimated parameter values, are presented in Table 7.

Table 7. PDF of lognormal distribution, Burr XII distribution, log-logistic distribution, and the estimated value of the parameters

Distribution	PDF	Parameter estimates
Lognormal	$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right); x > 0$	$\hat{\mu} = 7.33077, \hat{\sigma}^2 = 1.10216$
Burr XII	$f(x) = \frac{\alpha\gamma\left(\frac{x}{\theta}\right)^\gamma}{x\left[1 + \left(\frac{x}{\theta}\right)^\gamma\right]^{\alpha+1}}; x, \alpha, \gamma, \theta > 0$	$\hat{\alpha} = 0.63324, \hat{\gamma} = 2.07123,$ $\hat{\theta} = 1{,}017.08214$
Log-logistic	$f(x) = \frac{\gamma\left(\frac{x}{\theta}\right)^\gamma}{x\left[1 + \left(\frac{x}{\theta}\right)^\gamma\right]^2}; x, \gamma, \theta > 0$	$\hat{\gamma} = 1.70713, \hat{\theta} = 1{,}457.89989$

Graphical Model Analysis

The plots for the three competing models based on the lower 95% of the claim severity data in the original scale are presented in Figure 5. Visual inspection of the curves indicates that the Burr XII model provides the closest fit to the empirical distribution compared with the other two models.

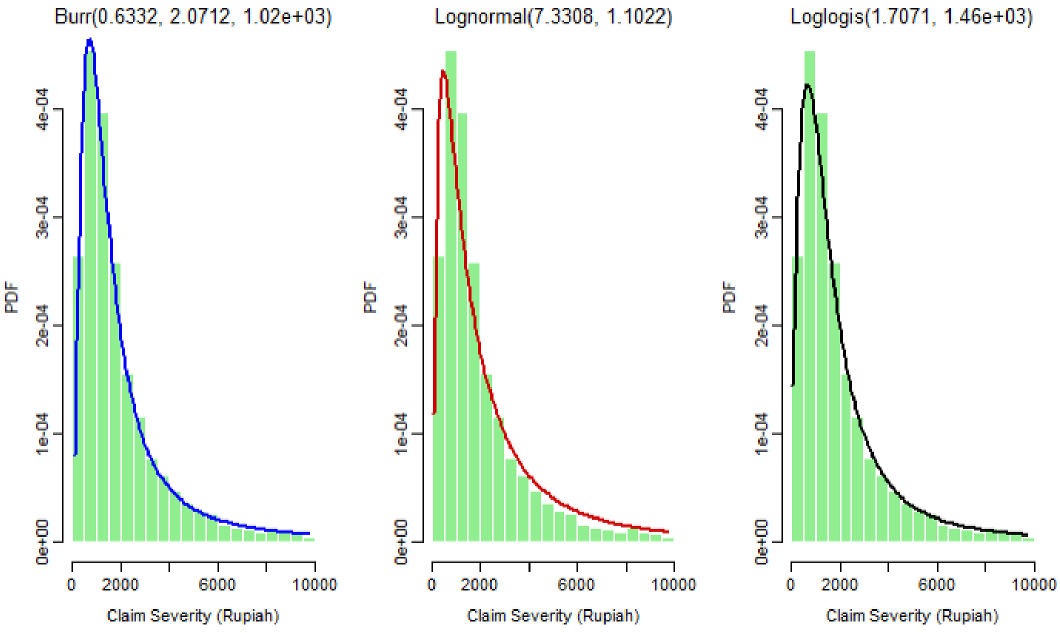

Figure 5. Plots of the lognormal, Burr XII, and log-logistic distribution for the lower 95% of claim severity data

The plots for the three models based on the full dataset on a logarithmic scale are shown in Figure 6. The graphical analysis suggests that the log-logistic model most accurately captures the behavior of the data relative to the Burr XII and lognormal models.

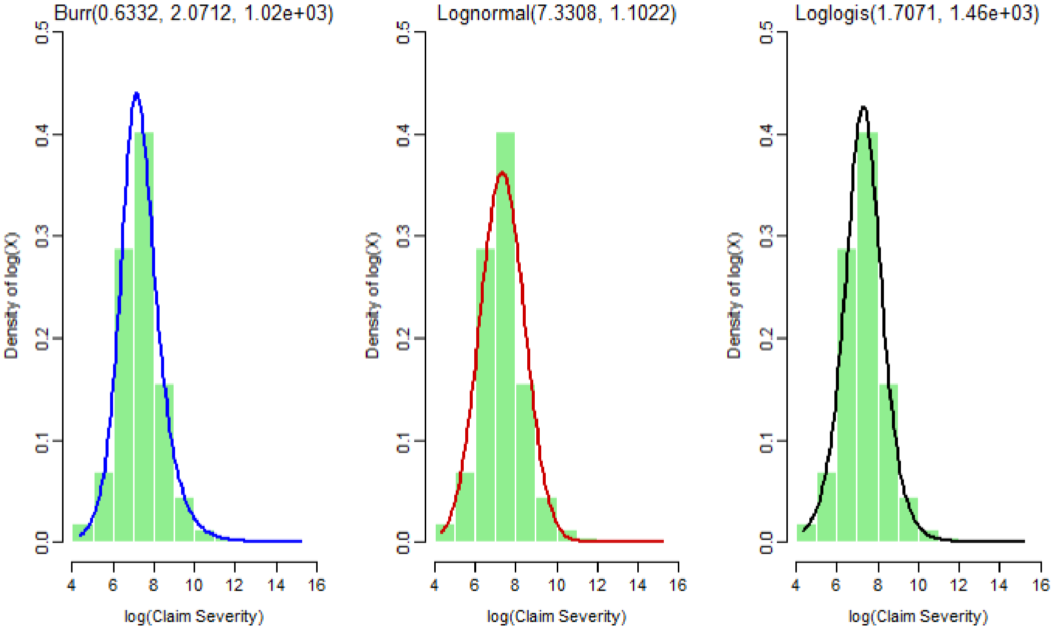

Figure 6. Plots of the lognormal, Burr XII, and log-logistic distributions for the full claim severity data on a logarithmic scale

Kolmogorov-Smirnov Test

To determine which model best represents the data, the Kolmogorov-Smirnov goodness-of-fit test was conducted. The hypotheses for this test are formulated as follows:

H_0: The claim severity data in the French MTPL severity dataset follow the specified distribution.
H_1: The claim severity data in the French MTPL severity dataset do not follow the specified distribution.

Table 8. Kolmogorov-Smirnov test p-values and discrepancy measures for French MTPL severity data

	Lognormal	Burr XII	Log-logistic
Discrepancy value	0.05771	0.01507	0.02469
p-value	2.2×10^{-16}	0.05479	0.00013

The test results are presented in Table 8. As shown, the Burr XII model is the only candidate distribution with a p-value greater than 0.05, namely 0.05479. Accordingly, the decision is to fail to reject H_0. This indicates that the claim severity data in the French MTPL severity dataset are adequately modeled by the Burr XII distribution.

Aggregate Loss Distribution

Calibration of Simulation Size Based on Empirical Mean and Variance

Because the claim severity distribution exhibits infinite theoretical variance, the variance of the aggregate loss cannot be reliably estimated using standard Monte Carlo methods alone, making a bootstrap approach necessary to obtain a stable empirical estimate. Figure 7 shows that the relative error for the estimated mean $E(S)$ quickly stabilizes near zero, indicating that its estimate converges even as the number of simulations increases. The relative error of the variance $Var(S)$fluctuates more substantially and only reaches convergence at around 10 million simulations. Beyond this point, the increasing variance estimate is not a sign of divergence, but rather a consequence of the Monte Carlo property that larger simulation sizes produce estimates that move closer to the true empirical value. In this case, the claim severity distribution has an infinite theoretical variance, so as n becomes large, the Monte Carlo estimator naturally drifts upward toward this unbounded value.

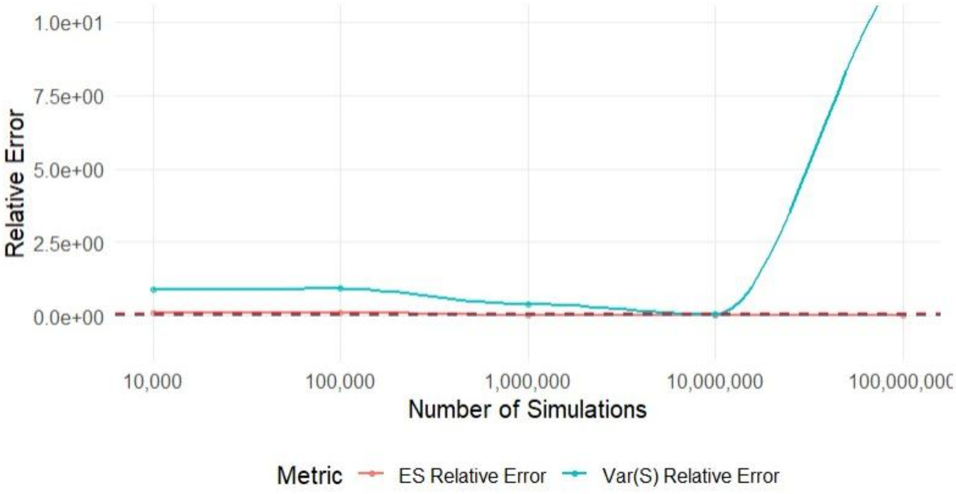

Figure 7. Convergence of error Monte Carlo estimates for aggregate loss

Based on Table 9, the number of simulations that yields the smallest relative error, which is 10 million simulations, is selected as the appropriate value of n.

Table 9. Aggregate loss estimates and relative error at different simulation levels

n-simulations	$E(S)$ simulation	$Var(S)$ simulation	$E(S)$ relative error	$Var(S)$ relative error
10^4	187.30331	18,020,930.46	0.10777	0.88175
10^5	187.47195	13,207,660.23	0.10697	0.91333
10^6	206.10761	92,577,836.29	0.01819	0.39251
10^7	**206.16643**	**159,121,167.6**	**0.01791**	**0.04414**
10^8	212.35608	1,950,183,716	0.01157	11.79697

Determining the Aggregate Loss Distribution Model with Monte Carlo Simulation

The aggregate loss distribution is obtained by combining the claim frequency and claim severity distributions. In this study, the frequency of claims is modeled using the Z123M-NB distribution, while the claim severity is modeled using the Burr XII distribution. The severity values, assumed to follow a Burr XII distribution with parameter estimates of $\hat{\alpha} = 0.63324$, $\hat{\gamma} = 2.07123, \hat{\theta} = 1017.08214$, were then simulated for each generated claim frequency.

Let N be a random variable representing the claim frequency within a given period. Let X be a random variable representing the claim severity per claim, assumed to be independent of N. The expectation and variance of X are obtained by applying the following formulas:

$$E(X) = \theta\, \alpha\, B\left(1 + \frac{1}{\gamma}, \alpha - \frac{1}{\gamma}\right)$$

$$Var(X) = \theta^2 \alpha\, B\left(1 + \frac{2}{\gamma}, \alpha - \frac{2}{\gamma}\right) - \left[\theta\, \alpha\, B\left(1 + \frac{1}{\gamma}, \alpha - \frac{1}{\gamma}\right)\right]^2.$$

The random variable S, representing the aggregate loss, is defined as

$$S = X_1 + X_2 + \cdots + X_N.$$

Theoretically, the expectation value and variance of S are given by the following expressions.

$$E(S) = E(N)\,E(X)$$
$$Var(S) = E(N)Var(X) + Var(N)\,[E(X)]^2.$$

In this study, a Monte Carlo simulation is employed to obtain empirical estimates of the expectation and variance of S. The estimated values of the expectation and variance of N, X, and S are presented in Table 10. For the Burr XII distribution, the existence of moments depends on the relationship between its shape parameters. In particular, the second moment exists only when $\alpha > \frac{2}{\gamma}$. Based on the parameter estimates obtained in this study, $\hat{\alpha} = 0.63324$ and $\hat{\gamma} = 2.07123$, yielding $\hat{\alpha} < \frac{2}{\hat{\gamma}}$. Consequently, the theoretical second moment, and therefore the theoretical variance, diverges and is not finite. This result reflects the extremely heavy-tailed nature of the fitted severity distribution, where rare but exceptionally large claim amounts have a substantial influence on higher-order moments. Because the theoretical variance cannot be computed directly, a non-parametric bootstrap procedure with 10,000 replications was employed to obtain a stable empirical estimate of the claim severity variance. The bootstrap approach repeatedly resamples the observed claim data and estimates the variance from the empirical distribution, thereby providing a practical measure of variability without relying on the existence of a finite theoretical variance under the fitted Burr XII model.

Table 10. The estimated values of the expectation and variance of N, X, and S

N		X		S	
$E(N)$	$Var(N)$	$E(X)$	$Var(X)$	$E(S)$	$Var(S)$
0.05327	0.05787	3,941.015	2,844,064,408	206.16643	159,121,167.6

Errors on the Aggregate Loss Distribution

The error calculation begins by obtaining the expected value and variance of the aggregate loss using Monte Carlo simulation, both analytically and through simulated estimation. The absolute error is then computed as the difference between the analytical and simulated values in their original units, while the relative error represents the magnitude of this difference relative to the analytical value [15]. The results of both error measures are presented in Table 11.

Table 11. Absolute and Relative Errors of the Estimated Mean and Variance of Aggregate Loss

	$E(S)$	$Var(S)$
Simulation results	206.16643	159,121,167.6
Analytical results	209.92696	152,394,224.1
Absolute error	3.76053	6,726,943.5
Relative error	0.01791	0.04414

The higher discrepancy observed in the variance is theoretically reasonable, as variance is more sensitive to fluctuations caused by extreme simulated values. Overall, the relative error remaining below 10% indicates that the Monte Carlo simulation procedure is sufficiently efficient. Therefore, these results are considered reliable for subsequent risk measures such as Value at Risk (VaR) and Tail Value at Risk (TVaR).

Descriptive Statistics of the Aggregate Loss Distribution

The descriptive statistics of the aggregate loss are presented in Table 12.

Table 12. Descriptive statistics of the aggregate loss

Mean	Variance	Skewness	Kurtosis
206.1664	1.59×10^8	767.57178	751,028.7348

By removing several uppermost values so that outliers do not dominate the visualization, the histogram of the lower 99% data in Figure 8 shows that most of the aggregate losses are clustered very close to zero, while only a few reach very large values, creating a long right tail. The high kurtosis value confirms that the data contain sharp peaks and a number of extreme outliers. This means that the aggregate loss can fluctuate a lot, mostly because of rare but very large claims [16].

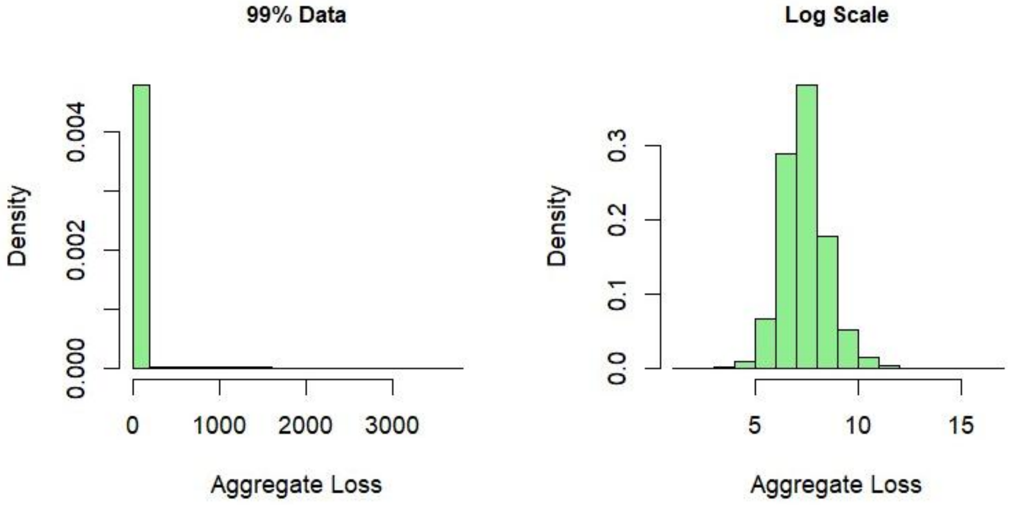

Figure 8. Histogram of aggregate loss for the lower 99% data and the logarithmic scale

Value at Risk (VaR) and Tail Value at Risk (TVaR) for the Aggregate Loss Claim Distribution with Monte Carlo Simulation

Value at Risk (VaR) is a risk measurement that represents the maximum potential loss that may occur within a given period at a specified confidence level. VaR provides an estimate of the loss that will not be exceeded with a probability equal to the chosen confidence level. However, VaR does not convey how severe the losses may be when this threshold is exceeded. In other words, it does not provide information about the magnitude of tail losses beyond the VaR boundary. Therefore, Tail Value at Risk (TVaR) is computed to estimate the average loss that occurs when the VaR level is surpassed. The VaR and TVaR values for the 95%, 97.5%, and 99% confidence levels are presented in Table 13.

Table 13. Estimated values of VaR and TVaR for the 95%, 97.5%, and 99% confidence levels

Level of confidence	VaR	TVaR
95%	105.85	4,122.93
97.5%	1,506.61	7,418.70
99%	3,629.14	15,075.21

At the 95% confidence level, the VaR is 105.85 while the TVaR is 4,122.93, indicating that losses beyond the VaR threshold are much larger on average. As the confidence level increases to 97.5% and 99%, both VaR and TVaR rise substantially, reflecting higher potential losses under more extreme scenarios. TVaR is consistently higher than VaR at all levels, highlighting the severity of tail losses in the aggregate loss distribution.

4. DISCUSSION

The findings of this study indicate that the Z123M-NB distribution provides the best fit for modeling claim frequency in French MTPL insurance data, while the Burr XII distribution is most suitable for claim severity. These results reflect the underlying characteristics of the dataset, where claim frequencies are dominated by zero-claim observations, which create strong overdispersion, while severity values exhibit extremely heavy-tailed behavior due to a small number of exceptionally large claims. The Z123M-NB model captures the unusually high probability mass at claim frequencies of 0-3, which standard negative binomial or Poisson models fail to accommodate. Likewise, the Burr XII distribution effectively represents extreme claim amounts due to its flexible shape and heavy-tail structure, a property well-established in actuarial risk modeling. This aligns with the theoretical findings of [17], who demonstrated that Burr XII is appropriate for modeling insurance losses and ruin-related quantities due to its capability to capture extreme variability. The large discrepancies observed between VaR and TVaR in this study further highlight the dominance of rare but severe loss events, confirming that tail behavior plays a critical role in aggregate risk modeling.

When compared with previous empirical work, the conclusions of this study are consistent with [18], who reported overdispersed claim frequencies and heavy-tailed severity patterns in motor insurance data, emphasizing the need for flexible frequency-severity modeling frameworks. Similarly, [19] found that motor insurance claims often display strong zero-inflation and substantial tail risks, supporting the appropriateness of using modified discrete distributions for frequency and heavy-tailed models for severity. In addition, the suitability of Burr XII for capturing large losses is reinforced by [20], who showed that Burr-type models outperform standard light-tailed alternatives when modeling extreme claims in insurance tariff applications. Together, these studies confirm that the modeling choices in this research are theoretically justified and empirically aligned with current practices in non-life insurance analytics, particularly when dealing with datasets characterized by heavy tails and rare catastrophic claims.

The implications of these findings are substantial for both actuarial science and industry practice. Theoretically, the study demonstrates that modified frequency distributions such as Z123M-NB are essential when claim data exhibit extreme zero-inflation, and that Burr XII's heavy-tail properties make it a robust model for severe motor insurance losses. Methodologically, the need for large-scale Monte Carlo simulation is evident when dealing with severity distributions whose theoretical variance is infinite, ensuring stable and reliable aggregate loss estimation. From a practical standpoint, the results highlight the importance for insurers and regulators to incorporate tail-sensitive metrics such as TVaR rather than relying solely on VaR, given the potential financial impact of rare but substantial claims. At a broader socio-economic level, more accurate modeling of aggregate motor insurance risk supports better pricing strategies, capital requirement calculations, and public policy decisions, ultimately strengthening consumer protection and the financial resilience of the insurance sector.

5. CONCLUSION

In conclusion, the claim frequency data are best modeled by the Z123M-NB distribution with parameters $\hat{p}_0 = 0.94977$, $\hat{p}_1 = 0.04746$, $\hat{p}_2 = 0.00263$, $\hat{p}_3 = 0.00012$, $\hat{r} = 9.39728$, and $\hat{\mu} = 6.87478$. The claim severity data are adequately represented by the Burr XII distribution with parameters $\hat{\alpha} = 0.63324$, $\hat{\gamma} = 2.07123$, and $\hat{\theta} = 1{,}017.08214$. Combining the Z123M-NB frequency model and the Burr XII severity model through Monte Carlo simulation with 10 million iterations yields a stable estimate of the aggregate loss distribution. The estimated VaR at the 95%, 97.5%, and 99% confidence levels are 105.85, 1,506.61, and 3,629.14, respectively. The corresponding TVaR estimates are 4,122.93, 7,418.70, and 15,075.21, respectively.

6. ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to all individuals and institutions who provided support throughout the completion of this research. Special appreciation is extended to the School of Data Science, Mathematics, and Informatics, IPB University, for providing financial support as well as facilitating access to computational resources and research infrastructure. The authors also acknowledge the availability of the freMTPL2 repository on Hugging Face, which significantly contributed to the feasibility of this study.

7. REFERENCES

[1] P. Fersini, S. Forte, G. Melisi, and G. Olivieri, "A stochastic model to evaluate pricing distortions in indemnity insurance methods for MTPL insurance," *Decis. Econ. Financ.*, vol. 42, no. 1, pp. 103–133, 2019, doi: 10.1007/s10203-019-00240-3.

[2] J. Segura-Gisbert, J. Lledó, and J. M. Pavía, "Dataset of an actual motor vehicle insurance portfolio," *Eur. Actuar. J.*, vol. 15, no. 1, pp. 241–253, 2025, doi: 10.1007/s13385-024-00398-0.

[3] S. Chen, Z. Wang, and M. Kelly, "Aggregate loss model with Poisson-Tweedie frequency," *Big Data Inf. Anal.*, vol. 6, no. 0, pp. 56–73, 2021, doi: 10.3934/bdia.2021005.

[4] S. K. Bar-Lev and A. Ridder, "Monte Carlo Methods for Insurance Risk Computation," *Int. J. Stat. Probab.*, vol. 8, no. 3, p. 54, 2019, doi: 10.5539/ijsp.v8n3p54.

[5] A. Noll, R. Salzmann, and M. V. Wuthrich, "Case Study: French Motor Third-Party Liability Claims," *SSRN Electron. J.*, pp. 1–41, 2018, doi: 10.2139/ssrn. 3164764.

[6] C. O. Omari, S. G. Nyambura, and J. M. W. Mwangi, "Modeling the Frequency and Severity of Auto Insurance Claims Using Statistical Distributions," *J. Math. Financ.*, vol. 08, no. 01, pp. 137–160, 2018, doi: 10.4236/jmf. 2018.81012.

[7] K. N. Nasution, "Prediksi Penjualan Barang pada Koperasi PT. Perkebunan Silindak dengan Menggunakan Metode Monte Carlo," *J. Ris. Komput.*, vol. 3, no. 6, pp. 65–69, 2016.

[8] Lusiana, S. Martha, and S. W. Rizki, "Simulasi pergerakan harga saham menggunakan pendekatan metode Monte Carlo," *Bul. Ilm. Math.Stat. dan Ter.*, vol. 07, no. 2, pp. 119–126, 2018.

[9] Z. G. Ignatov, V. K. Kaishev, and R. S. Krachunov, "An improved finite-time ruin probability formula and its Mathematica implementation," *Insur. Math. Econ.*, vol. 29, no. 3, pp. 375–386, 2001, doi: 10.1016/S0167-6687(01)00078-6.

[10] M. Denuit, X. Maréchal, S. Pitrebois, and J.-F. Walhin, *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Systems*. Chichester: John Wiley & Sons, 2007.

[11] U. Mukhaiyar, A. Dianpermatasari, A. Dzakiya, S. B. Widyani, and H. K. Syam, "The Value at Risk Analysis using Heavy-Tailed Distribution on the Insurance Claims Data," *JTAM (Jurnal Teor. dan Apl. Mat.)*, vol. 8, no. 4, p. 1233, 2024, doi: 10.31764/jtam.v8i4.25053.

[12] M. Alemany, R., Bolance, C., & Guillen, "Accounting for severity of risk when pricing insurance products," 2014.

[13] R. Diwakar, "An evaluation of normal versus lognormal distribution in data description and empirical analysis," *Pract. Assessment, Res. Eval.*, vol. 22, no. 13, pp. 1–15, 2017.

[14] S. A. Klugman, H. H. Panjer, and G. E. Willmot, *Loss Models from Data to Decisions*, 5th ed. Hoboken: John Wiley & Sons Inc., 2019. doi: 10.4018/979-8-3693-3820-9.ch010.

[15] M. Jacquemin, S. Putot, and F. Védrine, "A Reduced Product of Absolute and Relative Error Bounds for Floating-Point Analysis," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11002 LNCS, pp. 223–242, 2018, doi: 10.1007/978-3-319-99725-4_15.

[16] K. Dutta and J. Perry, "A Tale of Tails : An Empirical Analysis of Loss Distribution Models for Estimating Operational Risk Capital," vol. 2007, no. April, 2007.

[17] J. Das and D. C. Nath, "Burr Distribution as an Actuarial Risk Model and the Computation of Some of Its Actuarial Quantities Related to the Probability of Ruin," *J. Math. Financ.*, vol.

06, no. 01, pp. 213–231, 2016, doi: 10.4236/jmf.2016.61019.

[18] K. Antonio and R. Plat, “Micro-level stochastic loss reserving for general insurance,” *Scand. Actuar. J.*, vol. 2014, no. 7, pp. 649–669, Oct. 2014, doi: 10.1080/03461238.2012.755938.

[19] P. Shi and E. A. Valdez, “Multivariate negative binomial models for insurance claim counts,” *Insur. Math. Econ.*, vol. 55, pp. 18–29, Mar. 2014, doi: 10.1016/j.insmatheco.2013.11.011.

[20] C. Laudagé, S. Desmettre, and J. Wenzel, “Severity modeling of extreme insurance claims for tariffication,” *Insur. Math. Econ.*, vol. 88, pp. 77–92, Sep. 2019, doi: 10.1016/j.insmatheco.2019.06.002.