Modeling Aggregate Losses for Third Party Liability Insurance Using the Panjer Recursive Method

Amellia Putri Adelliawan(1), Aceng Komarudin Mutaqin(2)

[1,2]Department Statistics, Faculty of Mathematics and Natural Science, Universitas Islam Bandung

Jln. Ranggagading No.8, Tamansari, Kecamatan Bandung Wetan, Kota Bandung, Jawa Barat
e-mail: amelliaadelliawan@gmail.com(1), aceng.k.mutaqin@gmail.com(2)

ABSTRAK

Peningkatan jumlah kendaraan penumpang di wilayah padat penduduk seperti DKI Jakarta, Jawa Barat, dan Banten meningkatkan risiko kerugian finansial akibat kecelakaan, pencurian, dan kerusakan kendaraan. Penelitian ini bertujuan untuk memodelkan distribusi kerugian agregat pada asuransi tanggung jawab pihak ketiga (*Third Party Liability*, TPL) menggunakan metode rekursif Panjer. Data yang digunakan merupakan data riil PT. XYZ tahun underwriting 2018 untuk kendaraan angkutan penumpang dengan nilai pertanggungan antara Rp125 juta hingga Rp200 juta. Frekuensi klaim dimodelkan menggunakan distribusi Poisson, sedangkan besar klaim dimodelkan menggunakan distribusi Pareto Tipe II. Estimasi parameter dilakukan dengan metode *Maximum Likelihood Estimation* (MLE), dan kesesuaian model diuji menggunakan uji Chi-kuadrat dan Kolmogorov–Smirnov. Hasil penelitian menunjukkan bahwa distribusi kerugian agregat dapat dibentuk dengan baik menggunakan metode rekursif Panjer. Komponen diskrit terjadi pada kerugian agregat nol dengan peluang sebesar 0,992664, sedangkan komponen kontinu mencakup kerugian positif hingga Rp23.400.000. Tingginya peluang kerugian agregat nol menunjukkan bahwa klaim TPL pada portofolio yang diamati bersifat sangat jarang, yang memiliki implikasi penting dalam penentuan premi dan pengelolaan risiko pada asuransi kendaraan bermotor.
Kata kunci: Kerugian Agregat; Panjer Rekursif; Distribusi Poisson; Distribusi Pareto Tipe II; Asuransi Kendaraan Bermotor

ABSTRACT

The increasing number of passenger vehicles in densely populated areas such as Jakarta, West Java, and Banten has increased the risk of financial losses due to accidents, theft, and vehicle damage. This study aims to model the aggregate loss distribution in Third Party Liability (TPL) insurance using the Panjer recursive method. The data used are real observations from PT. XYZ for passenger vehicles insured between IDR 125 million and IDR 200 million in the 2018 underwriting year. Claim frequency is modeled using a Poisson distribution, while claim severity follows a Pareto Type II distribution. Model parameters are estimated using the Maximum Likelihood Estimation (MLE) method, and goodness-of-fit is evaluated using the Chi-square and Kolmogorov–Smirnov tests. The results show that the aggregate loss distribution can be effectively constructed using the Panjer recursive method. A dominant discrete probability mass occurs at zero aggregate loss with a probability of 0.992664, while the continuous component covers positive losses up to IDR 23,400,000. This result indicates that TPL claims in the observed portfolio are extremely sparse, which has important implications for premium pricing and risk management in motor vehicle insurance.
Keywords: *Aggregate Loss; Panjer Recursion; Poisson Distribution; Pareto Type II Distribution; Motor Vehicle Insurance*

INTRODUCTION

The increasing number of passenger motor vehicles in Indonesia, particularly in densely populated regions such as DKI Jakarta, West Java, and Banten, has raised the risk of financial losses due to accidents, theft, and natural disasters. Given that insured vehicle values typically range from IDR 125 million to IDR 200 million, motor vehicle insurance plays a vital role in providing financial protection against such risks.

Motor vehicle insurance products generally include Total Loss Only (TLO) and Comprehensive coverage, which can be extended with additional protections such as Third Party Liability (TPL). TPL insurance protects the insured against legal liabilities arising from bodily injury or property damage suffered by third parties. Understanding the aggregate loss distribution in TPL insurance is essential for insurers to assess risk exposure, determine appropriate premium levels, and ensure financial solvency.

Several studies have examined aggregate loss modeling using different numerical approaches. Kartini and Mutaqin [1] applied the Fast Fourier Transform method, while Sa'diah and Mutaqin [2] used numerical inversion of characteristic functions. Aryanti and Mutaqin [3] employed Laplace transforms, and Utami and Mutaqin [4] demonstrated the efficiency of the Panjer recursive method in modeling aggregate losses. International studies, such as those by Shevchenko [5] and Klugman et al. [6], emphasized the importance of aggregate loss distributions in actuarial risk management.

Unlike previous studies that mainly focus on general motor vehicle insurance or rely on transform-based approaches, this study specifically examines Third Party Liability (TPL) insurance using real Indonesian portfolio data. The contribution of this paper lies in the application of the Panjer recursive method to a mixed discrete–continuous aggregate loss distribution, supported by empirical modeling of claim frequency and severity using Poisson and Pareto Type II distributions. The results provide practical actuarial insights relevant to pricing and risk management of TPL insurance portfolios.

METHOD

This study uses secondary data obtained from PT. XYZ for the 2018 underwriting year. The dataset consists of 20,707 policyholders with 155 observed TPL claims after data validation. The data include claim frequency and claim severity for passenger vehicles insured under comprehensive motor vehicle insurance with TPL extensions in DKI Jakarta, West Java, and Banten, with insured values ranging from IDR 125 million to IDR 200 million.

Motor Vehicle Insurance

Motor vehicle insurance is an agreement between the insurer and the insured to provide compensation for losses due to risks such as collision, theft, and other perils. The two main types are *Total Loss Only (TLO)* and *Comprehensive*. Additionally, *Third Party Liability (TPL)* provides coverage for losses suffered by third parties due to accidents caused by the insured vehicle.

Poisson Distribution

Claim frequency is assumed to follow a Poisson distribution with parameter $\lambda > 0$, with probability mass function:

$$P(k|\lambda) = \frac{e^{-\lambda}\lambda^k}{k!}, for\ k = 0,1,2,\dots \quad (1)$$

The parameter λ is estimated using the Maximum Likelihood Estimation (MLE) method[7][8].

Pareto Type II Distribution

Claim severity is modeled using the Pareto Type II distribution with shape parameter α and scale parameter θ. Its probability density function and cumulative distribution function are given by

$$f(x) = \frac{\alpha\,\theta^{\alpha}}{(x+\theta)^{\alpha+1}}, x > 0, \alpha, \theta > 0. \quad (2)$$

$$F(x) = 1 - (\frac{\theta}{x+\theta})^{\alpha}. \quad (3)$$

The expectation and variance depend on the parameters α and θ, where $E(X) = \frac{\theta}{\alpha-1}$ for $\alpha > 1$ and $Var(X) = \frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}$for $\alpha > 2$. Parameter estimation is performed using MLE. Since the likelihood equations cannot be solved analytically, the Newton–Raphson iteration method is applied. The iteration is stopped when the absolute change in parameter estimates is less than 10^{-6}or when the maximum number of iterations is reached [9].

Goodness-of-Fit Test

The Chi-square test is used to evaluate the goodness-of-fit of the Poisson distribution for claim frequency, while the Kolmogorov–Smirnov test is applied to assess the suitability of the Pareto Type II distribution for claim severity data [10].

Chi-Square Test

Used for discrete data, the Chi-square statistic is defined as:

$$\chi^2 = \sum_{k=1}^{m} \frac{(n_k - np_k)^2}{np_k} \quad (4)$$

where n_k is the observed frequency, $np_k$is the expected frequency, and $m$is the number of categories. The statistic follows a Chi-square distribution with $(m - r - 1)$ degrees of freedom. Reject H_0 if $\chi^2 \geq \chi^2_{(m-r-1,1-\alpha)}$. If any expected frequency is less than 5, categories should be combined [11][12].

Kolmogorov–Smirnov Test

Used for continuous data, the K–S statistic is given by:

$$D = \max_{1\leq i\leq n} | F_n(x_i) - F^*(x_i) | \quad (5)$$

where $F_n(x_i)$is the empirical CDF and $F^*(x_i)$is the theoretical CDF. Accept H_0if $D < D_{critical}$; otherwise, reject H_0. [9][11]

Aggregate Loss Model

Aggregate loss is defined as the sum of all claim amounts within a given period. The Panjer recursive method is used to compute the aggregate loss distribution efficiently for frequency distributions belonging to the (a, b, 0) class. Continuous claim severity data are discretized using the central difference approximation before applying the recursion[13].

Distribution of Aggregate Loss

The aggregate loss distribution is constructed from the claim frequency (N) and claim severity (X) distributions. Its cumulative distribution involves repeated convolution of the claim severity distribution for different values of N[14][15].

Panjer Recursive Method

The Panjer recursive method efficiently computes the aggregate loss distribution without direct convolution. It applies to Panjer (a,b,0) class distributions, including Poisson, Binomial, and Negative Binomial. For continuous data, claim severities are discretized using the *central difference approximation* method.

Panjer Recursive Theorem

If the frequency distribution satisfies $p_n = (a + b/n)p_{n-1}$, the Panjer recursion applies. For the Poisson distribution, $a = 0$and $b = \lambda$.

Panjer Recursive Algorithm

Here is the algorithm of the Panjer recursion:

1. Initialization: Compute $f_X(0)$and $h_S(0)$based on the given formulas, and set $H_S(0) = h_S(0)$.
2. For $l = 1,2, \ldots$:
   a. Compute $f_X(l)$. If the claim severity distribution is continuous, then $f_X(l)$can be discretized using the theory described in Subsection 2.8.3.
   b. Compute
   $$h_S(l) = \frac{1}{1 - af_X(0)} \sum_{j=1}^{l} (a + \frac{bj}{l})\, f_X(j)\, h_S(l-j) \tag{6}$$
   c. Compute
   $$H_S(l) = H_S(l-1) + h_S(l) \tag{7}$$
   d. The procedure is temporarily stopped when $H_S(l)$exceeds the desired quantile level α, for example $\alpha = 0.999$. Increase l(i.e., set $l = l + 1$) and return to Step 2.

Discretization in the Panjer Recursion

Discretization of continuous claim severity is performed using central, forward, or backward difference approximations. These discretized values are then used in the Panjer recursion. The approach used to discretize the claim severity distribution is the central difference approximation, given by:

$$f_X(0) = F\left(\frac{\delta}{2}\right) \quad (8)$$

$$f_X(l) = F(l\delta + \frac{\delta}{2}) - F(l\delta - \frac{\delta}{2}), l = 1,2, \dots \quad (9)$$

RESULT AND DISCUSSION

Poisson Distribution Goodness-of-Fit Test

The Poisson goodness-of-fit test for the frequency of motor vehicle insurance claims at PT. XYZ, underwriting year 2018, was performed using the Chi-square test.

Hypotheses:

H_0: The frequency of TPL claims for PT. XYZ's motor vehicle insurance in underwriting year 2018 follows a Poisson distribution.

H_1: The frequency of TPL claims for PT. XYZ's motor vehicle insurance in underwriting year 2018 does not follow a Poisson distribution.

The estimated Poisson parameter is $\hat{\lambda} = 0.0075$.

The estimated probabilities for each claim frequency are:

$$P(K = 0) = e^{-0.0075} \times \frac{0.0075^0}{0!} = 0.99254$$

$$P(K = 1) = e^{-0.0075} \times \frac{0.0075^1}{1!} = 0.00743$$

$$P(K \geq 2) = 1 - [P(K = 0) + P(K = 1)] = 0.00003$$

Expected frequencies were calculated as:

$$np_0 = 20{,}707 \times 0.9925 = 20{,}552.58$$

$$np_1 = 20{,}707 \times 0.0074 = 153.84$$

$$np_2 = 20{,}707 \times 0.00003 = 0.58$$

The results for the estimated probabilities and expected values are presented in Table 1.

Table 1 Estimated Probability Value and Expected Number of Claims

Frekuensi Klaim (k) (1)	**Banyaknya Tertanggung** (2)	**Peluang Terjadinya Klaim (p_k)** (3)	**Nilai Harapan Terjadinya Klaim** (4)	$\frac{(n_k - np_k)^2}{np_k}$ (5)
0	20.553	0,99254	20.552,57867	0,00001
1	153	0,00743	153,84410	0,00463
2	1	0,00003	0,57723	0,30964
Jumlah	**20.707**	**1**	**20.707**	**0,31428**

The Chi-square statistic was 0.31428, while the critical value at α = 0.05 and 1 degree of freedom is 3.84145. Since 0.31428 < 3.84145, $H_0$is accepted, indicating that the frequency of claims follows a Poisson distribution.

The estimated Poisson parameter is very small, indicating that TPL claims occur infrequently per policyholder. This finding reflects the low-frequency nature of TPL claims in the observed insurance portfolio.

Estimation of Pareto II Distribution Parameters

Parameter estimation of the Pareto II distribution was performed using the *Maximum Likelihood Estimation* method via the Newton–Raphson iteration, assisted by RStudio.
The initial estimate of parameter θ was obtained using the *method of moments*:

$$m = \frac{1}{n}\sum_{i=1}^{n} x_i = 4.384.108{,}213$$

$$t = \frac{1}{n}\sum_{i=1}^{n} {x_i}^2 = 66.266.280.700.560{,}50$$

$$\theta = \frac{mt}{t - 2m^2} = 10{,}440{,}741.32$$

After iteration (as shown in the table below), the converged estimate was $\hat{\theta} = 10{,}162{,}823.25$.

Table 2 Newton–Raphson Iteration Results

Nilai Taksiran θ	Nilai Δ	Jumlah Iterasi
10.162.823,2407	176,7593	1
10.162.823,2466	0,0059	2
10.162.823,2466	0,0000	3
10.162.823,2466	0,0000	4

Using this $\hat{\theta}$, the estimated parameter $\hat{\alpha}$was calculated by MLE:

$$\hat{\alpha} = 3.3534$$

The estimated parameters are $\hat{\alpha} = 3.3534$ and $\hat{\theta} = 10{,}162{,}823.25$.

Pareto II Goodness-of-Fit Test

Parameter estimation using the Newton–Raphson method yields α = 3.3534 and θ = 10,162,823.25. The Kolmogorov–Smirnov test confirms that the Pareto Type II distribution provides an adequate fit to the claim severity data.

Hypotheses:

H_0: The TPL claim severity data follows the Pareto II distribution.
H_1: The TPL claim severity data does not follow the Pareto II distribution.
The theoretical cumulative distribution function (CDF) of the Pareto II distribution is:

$$F^*(x) = 1 - (\frac{\theta}{x + \theta})^{\alpha}$$

For example, for $x = 110{,}000$:

$$F^*(110{,}000) = 1 - (\frac{10{,}162{,}823.25}{110{,}000 + 10{,}162{,}823.25})^{3.3534} = 0.0355$$

The empirical CDF for the same observation is:

$$F_n(110,000) = \frac{2}{155} = 0.0129$$

The maximum absolute difference between theoretical and empirical CDFs is:

$$D = \max\ |\ F_n(x_i) - F^*(x_i)\ | = 0.0687$$

The critical value for $n = 155$and $\alpha = 0.05$is:

$$D_{crit} = \frac{1.36}{\sqrt{155}} = 0.1092$$

Since $D = 0.0687 < 0.1092$, H_0 is accepted the claim severity data follows the Pareto II distribution. The fitted Pareto Type II distribution indicates moderately heavy-tailed behavior, suggesting that although large TPL claims are rare, they may still have a significant financial impact when they occur.

Aggregate Loss Distribution Using Panjer Recursion

The Panjer recursive method was applied to determine the aggregate loss distribution of TPL policyholders.

From previous sections:

- Claim frequency follows Poisson (λ = 0.0075)
- Claim severity follows Pareto II (α = 3.3534, θ = 10,162,823.25)

Let the discretization step be δ = 100,000.

The smallest claim is Rp 100,000 and the largest is Rp 47,700,000 → approximately 480 intervals.

For initialization:

$$f_X(0) = F(\frac{\delta}{2}) = F(50,000) = 1 - (\frac{10,162,823.25}{50,000 + 10,162,823.25})^{3.3534} = 0.0163$$

$$h_S(0) = e^{\lambda(f_X(0)-1)} = e^{0.0075(0.0163-1)} = 0.9927$$

$$H_S(0) = h_S(0) = 0.9927$$

For $l = 1$:

$$f_X(1) = F(150,000) - F(50,000) = 0.0316$$

$$h_S(1) = (0.0075)(0.0316)(0.9927) = 0.0001$$

$$H_S(1) = 0.9927 + 0.0001 = 0.9928$$

Continuing this process recursively yields the aggregate loss density and CDF up to $H_S(l) = 1$(or nearly 1) at around Rp 23,400,000.

Tabel 3 Aggregate Loss Density Function Value and Cumulative Distribution Function Value

l	$s = l\delta$	$f_x(l)$	$h_s(l)$	$H_s(l)$
0	0	0,0163	0,9927	0,992664
1	100.000	0,0316	0,0001	0,992785
2	200.000	0,0303	0,0002	0,993021
3	300.000	0,0291	0,0002	0,993247
4	400.000	0,0279	0,0002	0,993463
5	500.000	0,0268	0,0002	0,993671

6	600.000	0,0257	0,0002	0,993870
⋮	⋮	⋮	⋮	⋮
229	22.900.000	0,0002	0,0000	0,999992
230	23.000.000	0,0002	0,0000	0,999993
231	23.100.000	0,0002	0,0000	0,999995
232	23.200.000	0,0002	0,0000	0,999996
233	23.300.000	0,0002	0,0000	0,999997
234	23.400.000	0,0002	0,0000	0,999999

The probability of zero aggregate loss is 0.992664, while the continuous component covers aggregate losses up to IDR 23,400,000, where the cumulative distribution function approaches one. The high probability mass at zero aggregate loss reflects the sparsity of TPL claims, whereas the continuous component captures the financial consequences of non-zero claims. These results are important for determining appropriate premiums and managing risk exposure in TPL insurance portfolios.

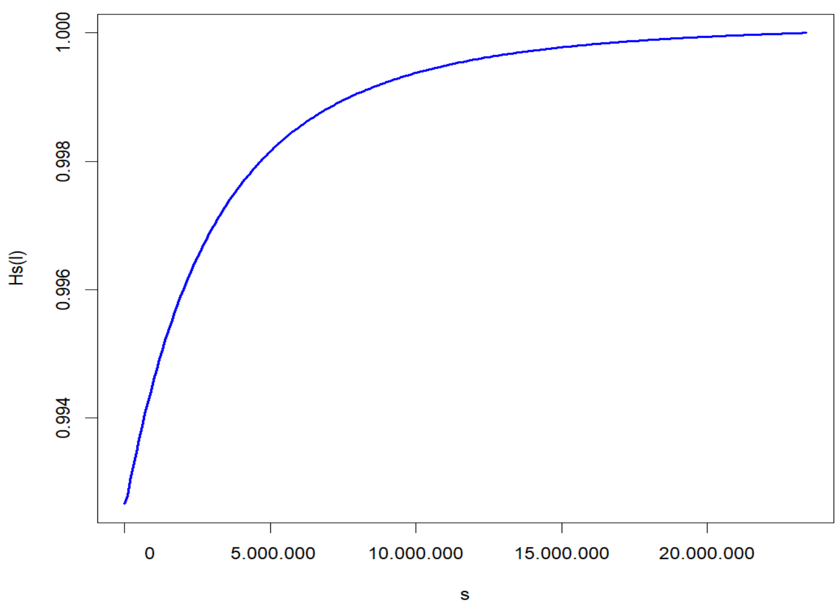

Figure 1. The cumulative distribution function curve of aggregate losses using the Panjer recursive method.

Figure 1 shows the cumulative distribution function of aggregate losses obtained using the Panjer recursive method. The curve increases sharply at zero aggregate loss and approaches one at approximately IDR 23,400,000, indicating that almost all probability mass is concentrated below this value. This behavior reflects the low frequency of TPL claims in the observed portfolio and the limited contribution of large aggregate losses.

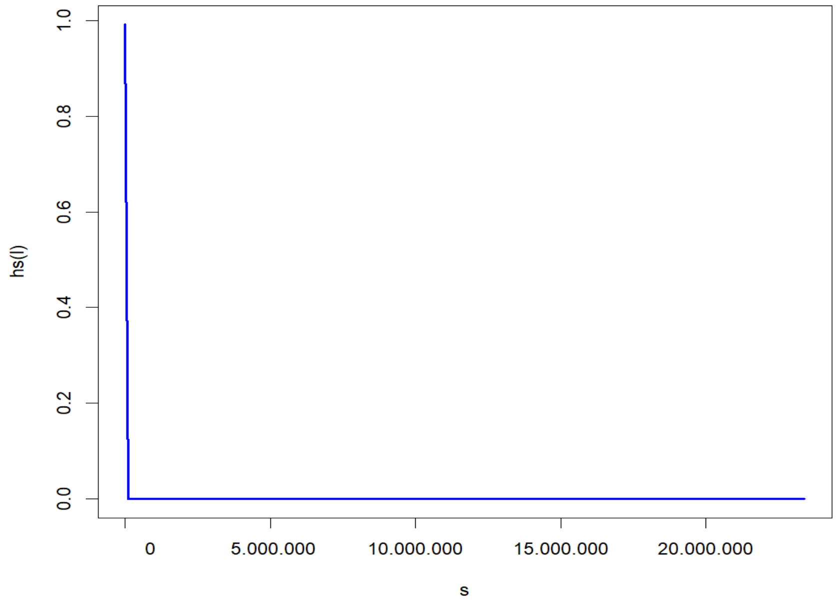

Figure 2. The density function curve of aggregate losses using the Panjer recursive method.

Figure 2 presents the aggregate loss density function, which is dominated by a high probability mass at zero loss. This indicates that most policyholders do not generate any TPL claims during the observation period. The presence of a long but thin tail reflects the possibility of non-zero aggregate losses, although such events occur with relatively low probability.

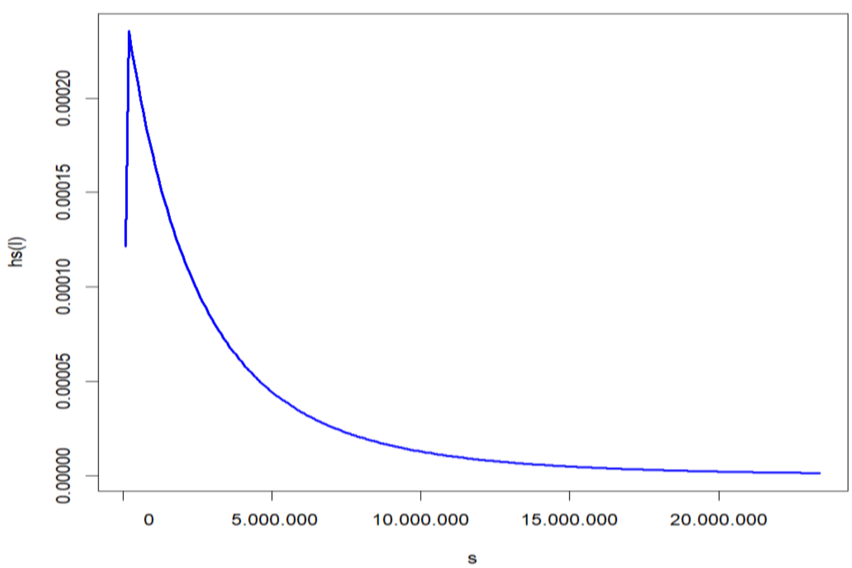

Figure 3. The density function curve of aggregate losses using the Panjer recursive method for non-zero aggregate losses.

Figure 3 illustrates the density function of non-zero aggregate losses. The right-skewed shape of the distribution indicates that smaller aggregate losses are more likely, while larger losses occur less frequently. This pattern highlights the importance of modeling claim severity accurately, as rare but larger claims may still have a significant financial impact.

CONCLUSION

This study applies the Panjer recursive method to model the aggregate loss distribution of Third-Party Liability insurance for passenger vehicles insured in Indonesia. Claim frequency is modeled using a Poisson distribution, while claim severity follows a Pareto Type II distribution. The results show that the aggregate loss distribution has a mixed structure consisting of a dominant discrete component at zero loss and a continuous component for positive losses.

The high probability of zero aggregate loss indicates that TPL claims are extremely sparse in the observed portfolio, while the continuous component highlights the potential impact of non-zero claims. These findings demonstrate that the Panjer recursive method provides an effective and practical tool for aggregate loss modeling and supports actuarial decision-making related to pricing and risk management in TPL insurance.

REFERENCE

[1] N. Kartini and A. K. Mutaqin, “Penentuan Distribusi Kerugian Agregat Tertanggung Asuransi Kendaraan Bermotor di Indonesia Menggunakan Metode Fast Fourier Transform,” *J. Ilm. Teori dan Apl. Stat.*, vol. 2, no. 1, pp. 45–54, 2018.

[2] K. Sa’diah and A. K. Mutaqin, “Penentuan Distribusi Kerugian Agregat Melalui Invers Numerik dari Fungsi Karakteristik Menggunakan Aturan Trapezoidal Quadrature,” *J. Ilm. Teori dan Apl. Stat.*, vol. 4, no. 2, pp. 67–75, 2020.

[3] P. Aryanti and A. K. Mutaqin, “Penerapan Solusi Analitik Distribusi Kerugian Agregat Melalui Transformasi Laplace pada Data Klaim Asuransi Kendaraan Bermotor di Indonesia,” *J. Ilm. Teori dan Apl. Stat.*, vol. 7, no. 1, pp. 21–30, 2023.

[4] D. A. Utami and A. K. Mutaqin, “Penentuan Distribusi Kerugian Agregat Tertanggung Asuransi Kendaraan Bermotor di Indonesia Menggunakan Metode Rekursif Panjer,” *J. Ilm. Teori dan Apl. Stat.*, vol. 1, no. 2, pp. 55–63, 2017.

[5] P. Shevchenko, “Calculation of Aggregate Loss Distributions,” *J. Oper. Risk*, vol. 5, no. 2, pp. 3–40, 2010, doi: 10.21314/jop.2010.077.

[6] S. A. Klugman, H. H. Panjer, and G. E. Willmot, *Loss Models: From Data to Decisions*. Hoboken, NJ, USA: Wiley, 2012.

[7] Rani Kurnia Putri, Muhammad Athoillah, and Restu Ria Wantika, “Identifikasi Faktor Pembentuk Minat Belajar Matematika Pada Peserta Didik,” Seminar Nasional Hasil Riset dan Pengabdian, vol. 7, no. 1, pp. 70–82, Dec. 2025, [Online]. Available: https://snhrp.unipasby.ac.id/prosiding/index.php/snhrp/article/view/1186

[8] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, Applied logistic regression. John Wiley & Sons, 2013.

[9] J. Kenouza and A. Belguerna, “Optimal Solutions of Parameters of the Modified Weibull Distribution Using Least Squares and Maximum Likelihood Methods,” *Turk. J. Comput. Math. Educ.*, vol. 13, no. 2, pp. 202–210, 2022.

[10] R. K. Putri, M. Athoillah, and A. Haqiqiyah, "Analisis Faktor Yang Mempengaruhi Ketepatan Kelulusan Mahasiswa Dengan Algoritma Regresi Linear," Jurnal Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika, vol. 5, no. 2, pp. 671–680, 2024.

[11] M. Sarstedt, E. Mooi, M. Sarstedt, and E. Mooi, "Regression analysis," A concise guide to market research: The process, data, and methods using IBM SPSS Statistics, pp. 209–256, 2019.

[12] M. L. McHugh, "The Chi-square Test of Independence," *Biochem. Med.*, vol. 23, no. 2, pp. 143–149, 2013, doi: 10.11613/BM.2013.018.

[13] S. Nugroho, *Statistika Nonparametrika*. Yogyakarta, Indonesia: Gadjah Mada Univ. Press, 2008.

[14] E. M. P. Hermanto, M. Athoillah, W. N. Hamidah, and D. P. Putra, "Pelatihan Penggunaan Software R Untuk Menguji Perbandingan Berganda Dan Asumsi Residual Pada Rancangan Percobaan," J-ABDI: Jurnal Pengabdian kepada Masyarakat, vol. 1, no. 4, pp. 449–458, 2021.

[15] Nuryadi, T. D. Astuti, E. S. Utami, and M. Budianta, *Dasar-Dasar Statistik Penelitian*. Yogyakarta, Indonesia: SIBUKU MEDIA, 2017.