

Bonus-Malus Premium for Third Party Liability Insurance with Poisson-Lindley Distribution Claim Frequency and Exponential-Inverse Gamma Distribution Claim Severity

Bilqis Nur Rizkia ⁽¹⁾, Aceng Komarudin Mutaqin ⁽²⁾

^{1,2}Department of Statistics, Faculty of Mathematics and Natural Science, Universitas Islam

Bandung

Jln. Ranggagading No.8, Tamansari, Kecamatan Bandung Wetan, Kota Bandung, Jawa Barat

E-mail: bilqisnr76@gmail.com⁽¹⁾, aceng.k.mutaqin@gmail.com⁽²⁾

ABSTRACT

Asuransi adalah bentuk gotong royong yang memberikan perlindungan terhadap risiko tak terduga. Dengan asuransi, seseorang merasa lebih tenang menghadapi kemungkinan buruk, baik bagi dirinya maupun hartanya. Seiring meningkatnya jumlah kendaraan bermotor di Indonesia, risiko kecelakaan juga naik, sehingga asuransi kendaraan bermotor, khususnya asuransi tanggung jawab pihak ketiga (*Third Party Liability* (TPL), menjadi penting. Untuk meningkatkan keadilan, perusahaan asuransi menerapkan sistem premi berdasarkan riwayat klaim, salah satunya sistem bonus malus. Dalam penelitian ini dibahas perhitungan premi pada sistem bonus malus untuk asuransi TPL, dengan asumsi frekuensi klaim berdistribusi Poisson-Lindley dan besar klaim berdistribusi eksponensial-invers gamma. Data yang digunakan adalah data sekunder dari PT. XYZ pada tahun underwriting 2019 untuk pemegang polis kategori dua. Hasil analisis menunjukkan bahwa distribusi tersebut sesuai dengan data, dan sistem bonus malus yang optimal memberikan premi murni awal sebesar Rp22.970 untuk pemegang polis baru. Premi tahun berikutnya disesuaikan berdasarkan klaim: meningkat jika ada klaim, dan menurun jika tidak ada klaim.

Kata kunci: Bonus-Malus; Asuransi *Third Party Liability*; Distribusi Poisson-Lindley dan Eksponensial-Invers Gamma

ABSTRACT

Insurance is a form of mutual cooperation that provides protection against unforeseen risks. With insurance, individuals can feel more secure about potential future losses, whether related to themselves or their property. As the number of motor vehicles in Indonesia increases, so does the risk of traffic accidents, making motor vehicle insurance particularly Third Party Liability (TPL) insurance increasingly important. To enhance fairness, insurance companies implement premium systems based on claim history, one of which is the bonus-malus system. This study discusses premium calculation in a bonus-malus system for TPL insurance, assuming that claim frequency follows a Poisson-Lindley distribution and claim severity follows an exponential-inverse gamma distribution. The data used are secondary data obtained from PT. XYZ for the 2019 underwriting year, focusing on policyholders in category two. The analysis results indicate that the selected distributions fit the data well. The optimal bonus-malus system determines that the initial pure premium to be paid by new policyholders is Rp22,970. Premiums in subsequent years are adjusted based on claim activity: increasing if a claim is made and decreasing if no claim occurs.

Keywords: Bonus-Malus; Third Party Liability Insurance; Poisson-Lindley and Exponential-Inverse Gamma Distributions

INTRODUCTION

Indonesia, with the number of motor vehicles increasing every year, faces major challenges in managing transportation and highway safety. Data from the Badan Pusat Statistik (BPS) noted that from 2018 to 2023, the number of motorcycles increased from 106 million to 132 million units, up 26 million units in six years. Cars also increased from 14 million to 18 million units [1]. This growth not only causes congestion, but also increases the potential for financial losses due to accidents. In this case, there are still many road users, including third parties who are harmed in accidents, do not have adequate financial protection.

Motor vehicle insurance is an important solution. There are two main types of coverage in motor vehicle insurance in Indonesia, namely Comprehensive and Total Loss Only (TLO), where the insured must choose one of them. In addition, additional protections such as Third Party Liability (TPL) can be added as needed. TPL Insurance provides protection against claims for compensation from third parties who suffer losses due to accidents caused by policyholders. As a measure to improve protection for all road users, the Indonesian government has announced a plan to implement mandatory TPL insurance for all motor vehicles starting January 2025 [2].

One of the important things in motor vehicle insurance is to determine the amount of premium that must be paid by the policyholder [3]. One method to determine the amount of premium is the bonus-malus system. This system provides incentives in the form of premium reductions for drivers who do not file claims (bonus) and imposes higher premiums for those who often file claims (malus). The system is designed to encourage safer and more responsible driving behavior. The simplest bonus-malus system is based on the frequency of claims filed by policyholders. The system is considered unfair because insured who submits a claim with a small value is subject to the same premium as a claim with a large value. The solution to improve the fairness of the bonus-malus system is to consider two components, namely the frequency of claims and the severity of claims.

Various previous studies have examined the bonus-malus system in motor vehicle insurance using a different distribution approach. Eygenio and Qoyyimi [4] designed a generalized bonus-malus system with the frequency of Negative Binomial distributed claims and the magnitude of Pareto distributed claims on motor vehicle insurance in the 2014 motor vehicle insurance policy data of the insurance company Jasindo. Adisti and Mutaqin [5] applied pure premium calculation to the bonus-malus system where the frequency of Negative Binomial distributed claims and the amount of Weibull distributed claims on category 8 motor vehicle insurance data in region 3 in Indonesia. Sevina and Purwadi [6] applied pure premium calculation to the bonus-malus system where the frequency of Geometric distributed claims and the amount of Weibull distributed claims on the secondary data of motor vehicle insurance in Indonesia. In this study, the application of premium calculation in the bonus-malus system will be carried out on motor vehicle insurance in Indonesia for Third Party Liability (TPL) coverage, with the assumption that the claim frequency follows a mixture Poisson-Lindley distribution and the claim severity follows a mixture Exponential-Inverse Gamma distribution. The Poisson-Lindley distribution is selected because it offers greater flexibility than the standard Poisson distribution in modeling overdispersed count data, a common feature in insurance claim frequencies. It also provides a better fit when the variance of the data exceeds the mean, which is often observed in real insurance datasets.

METHOD

The data used in this study is secondary data taken from the recording of PT. XYZ in the 2019 underwriting year, includes the frequency of claims and severity claims from policyholders of motor vehicle insurance products covered comprehensive expansion of Third Party Liability (TPL) category two in Indonesia.

Poisson Distribution

A discrete random variable K is said to be distributed in Poisson with a parameter $\lambda > 0$, and has the following probability mass function:

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

Poisson-Lindley Mixture Distribution

The Poisson-Lindley mixture distribution is formed from the Poisson distribution with the parameter λ following the Lindley distribution with the parameter $\delta > 0$. The probability density function of the Lindley distribution is:

$$\pi(\lambda) = \frac{\delta}{\delta+1}(\lambda+1)e^{-\delta\lambda}, \text{ for } \lambda > 0, \delta > 0 \tag{2}$$

The probability function of the Poisson-Lindley mixture distribution is as follows:

$$p_k = P(K = k) = \int_0^\infty P(k|\lambda)\pi(\lambda)d\lambda$$

$$p_k = \frac{\delta^2(k+\delta+2)}{(\delta+1)^{k+3}}; k = 0,1,2, \dots, \delta > 0 \tag{3}$$

The parameters of the Poisson-Lindley mixture distribution can be estimated using the Maximum Likelihood Estimation (MLE) method. The log-likelihood function of the Poisson-Lindley mixture distribution for random samples $k = k_1, k_2, \dots, k_n$ is:

$$l(\delta) = 2n \ln \delta + \sum_{i=1}^n \ln(k_i + \delta + 2) - \ln(\delta + 1) \sum_{i=1}^n (k_i + 3) \tag{4}$$

The first derivative of the log-likelihood function against a parameter δ equalized to zero is obtained as follows:

$$\frac{d}{d\delta} = \frac{2n}{\delta} + \sum_{i=1}^n \frac{1}{(k_i+\delta+2)} - \frac{1}{\delta+1} \sum_{i=1}^n (k_i + 3) = 0 \tag{5}$$

There is no analytical solution to obtain parameter estimates δ , therefore a numerical method is used, namely the Newton-Raphson iteration method, which requires a second derivative of the log-likelihood function for the parameter δ in the calculation process.

$$\frac{d^2}{d\delta^2} = -\frac{2n}{\delta^2} - \sum_{i=1}^n \frac{1}{(k_i+\delta+2)^2} + \frac{1}{(\delta+1)^2} \sum_{i=1}^n (k_i + 3) \tag{6}$$

The Newton-Raphson iteration equation is:

$$\hat{\delta}_{h+1} = \hat{\delta}_h - \frac{\frac{2n}{\hat{\delta}_h} + \sum_{i=1}^n \frac{1}{(k_i+\hat{\delta}_h+2)} - \frac{1}{\hat{\delta}_h+1} \sum_{i=1}^n (k_i+3)}{-\frac{2n}{\hat{\delta}_h^2} - \sum_{i=1}^n \frac{1}{(k_i+\hat{\delta}_h+2)^2} + \frac{1}{(\hat{\delta}_h+1)^2} \sum_{i=1}^n (k_i+3)}; h = 0,1,2, \dots \tag{7}$$

The iteration process is stopped if $|\hat{\delta}_{h+1} - \hat{\delta}_h| < \epsilon$, with the value $\epsilon = 1 \times 10^{-6}$. The initial value used in the iteration process is obtained from the estimation of the Lindley distribution parameter using the moment method which is [7]:

$$\hat{\delta}_0 = \frac{-(\bar{k}-1) + \sqrt{(\bar{k}-1)^2 + 8\bar{k}}}{2\bar{k}}, \text{ for } \bar{k} > 0 \tag{8}$$

where:

$$\bar{k} = \frac{\sum_{i=1}^n k_i}{n} \tag{9}$$

Exponential Distribution

A continuous random variable X is said to follow an exponential distribution with the parameter θ if it has the following density function [8]:

$$f(x|\theta) = \frac{e^{-x/\theta}}{\theta}, \text{ for } x > 0, \theta > 0 \tag{10}$$

The cumulative distribution function of the exponential distribution with the parameter θ :

$$F(x|\theta) = 1 - e^{-x/\theta}, \text{ for } x > 0, \theta > 0 \tag{11}$$

Exponential-Inverse Gamma Mixture Distribution

The exponential distribution of the inverse gamma mixture is formed from an exponential distribution with a parameter θ following an inverse gamma distribution with the parameter $\alpha > 0$ and $\beta > 0$. The probability density function of the inverse gamma distribution is [8]:

$$F\pi(\theta) = \frac{\left(\frac{\beta}{\theta}\right)^\alpha e^{-\beta/\theta}}{\theta\Gamma(\alpha)}, \text{ for } \theta > 0, \alpha > 1, \beta > 0 \tag{12}$$

The probability function of the exponential–inverse gamma mixture distribution is as follows:

$$f(x) = \int_0^\infty f(x|\theta)\pi(\theta)d\theta$$

$$f(x) = \frac{\alpha\beta^\alpha}{(x+\beta)^{\alpha+1}}, \text{ for } x > 0, \alpha > 0, \beta > 0 \tag{13}$$

The cumulative distribution function of the exponential–inverse gamma mixture distribution is:

$$F(x) = 1 - \left(\frac{\beta}{x+\beta}\right)^\alpha, \text{ for } x > 0, \alpha > 0, \beta > 0 \tag{14}$$

The parameters of the exponential–inverse gamma mixture distribution can be estimated using the Maximum Likelihood Estimation (MLE) method. The log-likelihood function of the exponential–inverse gamma mixture distribution for random samples $x = x_1, x_2, \dots, x_n$ is:

$$l(\alpha, \beta) = n \ln \alpha + n\alpha \ln(\beta) - (\alpha + 1) \sum_{i=1}^n \ln(x_i + \beta) \tag{15}$$

The first derivative of the log-likelihood function to the parameter α and β which is equaled to zero is obtained as follows:

$$\frac{dl(\alpha, \beta)}{d\alpha} = \frac{n}{\alpha} + n \ln(\beta) - \sum_{i=1}^n \ln(x_i + \beta) = 0 \tag{16}$$

$$\frac{dl(\alpha, \beta)}{d\beta} = n \frac{\alpha}{\beta} - (\alpha + 1) \sum_{i=1}^n \frac{1}{x_i + \beta} = 0 \tag{17}$$

There is no analytical solution to obtain parameter estimates α and β therefore a numerical method is used, namely the Newton-Raphson iteration method, which requires a second derivative of the log-likelihood function to the parameter α and β in the calculation process.

$$\frac{d^2l(\alpha, \beta)}{d\alpha^2} = -\frac{n}{\alpha^2} \tag{18}$$

$$\frac{d^2l(\alpha,\beta)}{d\beta^2} = -n\frac{\alpha}{\beta^2} + (\alpha + 1) \sum_{i=1}^n \frac{1}{(x_i+\beta)^2} \tag{19}$$

$$\frac{d^2l(\alpha,\beta)}{d\alpha d\beta} = \frac{n}{\beta} - \sum_{i=1}^n \frac{1}{x_i+\beta} \tag{20}$$

Suppose $\gamma = (\alpha, \beta)T$ it is a parameter vector. At the $(h + 1)$ iteration, the updated parameter can be obtained:

$$\gamma^{(h+1)} = \gamma^{(h)} - \left[\frac{d^2l(\gamma)}{d\gamma d\gamma^T} \Big|_{\gamma=\gamma^{(h)}} \right]^{-1} \left[\frac{dl(\gamma)}{d\gamma} \Big|_{\gamma=\gamma^{(h)}} \right], h = 0, 1, 2, .. \tag{21}$$

where,

$$\frac{dl(\gamma)}{d\gamma} \Big|_{\gamma=\gamma^{(h)}} = \begin{pmatrix} \frac{dl(\alpha,\beta)}{d\alpha} \Big|_{\alpha=\alpha^{(h)}; \beta=\beta^{(h)}} \\ \frac{dl(\alpha,\beta)}{d\beta} \Big|_{\alpha=\alpha^{(h)}; \beta=\beta^{(h)}} \end{pmatrix} \tag{22}$$

and

$$\frac{d^2l(\gamma)}{d\gamma d\gamma^T} \Big|_{\gamma=\gamma^{(h)}} = \begin{pmatrix} \frac{d^2l(\alpha,\beta)}{d\alpha^2} \Big|_{\alpha=\alpha^{(h)}; \beta=\beta^{(h)}} & \frac{d^2l(\alpha,\beta)}{d\alpha d\beta} \Big|_{\alpha=\alpha^{(h)}; \beta=\beta^{(h)}} \\ \frac{d^2l(\alpha,\beta)}{d\alpha d\beta} \Big|_{\alpha=\alpha^{(h)}; \beta=\beta^{(h)}} & \frac{d^2l(\alpha,\beta)}{d\beta^2} \Big|_{\alpha=\alpha^{(h)}; \beta=\beta^{(h)}} \end{pmatrix} \tag{23}$$

The iteration process is stopped when $|\gamma^{(h+1)} - \gamma^{(h)}| < \epsilon$, with the value $\epsilon = 1 \times 10^{-6}$. The initial values used in the iteration process can be obtained using the estimation of the parameters of the exponential-inverse gamma mixture distribution using the moment method, as follows:

$$\hat{\alpha}_0 = \frac{-2s^2}{(\bar{x}^2 - s^2)} \tag{24}$$

$$\hat{\beta}_0 = \bar{x}(\hat{\alpha}_0 - 1) \tag{25}$$

Goodness-of-Fit Testing for the Poisson-Lindley Mixture Distribution

The following are the steps in conducting a Poisson-Lindley mixture distribution compatibility test analysis of claim frequency data using the Chi-Squared test:

1. Formulating a test hypothesis.
 H_0 : Claim frequency data comes from a Poisson-Lindley mixture distributed population.
 H_1 : Claim frequency data does not come from a Poisson-Lindley mixture distributed population.
2. Estimating the parameters of the Poisson-Lindley mixture distribution using the Newton-Raphson iteration numerical method.
3. Calculates the probability for each claim frequency category, p_k , for $k = 0.1.2,...$ based on Equation (3).
4. Calculate the expected value for each claim frequency category.
5. Calculate the statistical value of the Chi-Square test using the equation:

$$\chi^2 = \sum_{i=1}^m \frac{(O_i - E_i)^2}{E_i} \tag{26}$$

Information:

E_i : The number of observations expected in the i th category based on the distribution function

O_i : Number of observations observed in the i th category

- Decide whether the null hypothesis is accepted or rejected. The null hypothesis is rejected if $\chi^2 \geq \chi^2_{(m-p-1)(1-\alpha)}$. According to Howell [9], the chi-square test can be invalid if the value of $E_i < 5$, but it can still be used if the sample total is large enough and the null hypothesis is not rejected.

Goodness-of-Fit Testing for the Exponential-Inverse Gamma Mixture Distribution

The following are the steps in conducting a match test of the exponential–inverse gamma mixture distribution match to claim severity data using the Kolmogorov-Smirnov test:

- Formulating a test hypothesis.
 H_0 : Claim severity data to comes from a population of exponential mix-inverse gamma.
 H_1 : Claim severity data does not come from a population with an exponential-inverse gamma mixture.
- Estimating the parameters of the exponential-inverse gamma mixture distribution parameters using the Newton-Raphson iteration numerical method.
- Sorting claim severity data on motor vehicle claims from smallest to largest value.
- Perform calculations based on the empirical cumulative distribution function on the data for each observation using the equation:

$$F_n(x_i) = \frac{\text{the number of observations } \leq x_i}{n} \tag{27}$$

- Calculate the value of the cumulative distribution function of the exponential-inverse gamma mixture using Equation (14).
- Calculating the statistical value of the Kolmogorov-Smirnov test using the equation:

$$D = \max_{1 \leq i \leq n} |F_n(x_i) - F^*(x_i)| \tag{28}$$

- Decide whether the null hypothesis is accepted or rejected. The null hypothesis is accepted if the test statistic D is smaller than the critical value.

Premium Calculation with the Optimal Bonus-Malus System

The following are the steps in calculating pure premium using the optimal bonus-malus system:

- Estimate the average number of claims for the year $(t + 1)$ using the equation:

$$\hat{\lambda}_{t+1,N} = \frac{(N+1)(N+2+t+\delta)}{(t+\delta)(N+1+t+\delta)} \tag{29}$$

- Estimate the average claim severity for the year $(t + 1)$ using the equation:

$$\hat{\theta}_{t+1,N} = \frac{\sum_{i=1}^N x_i + \hat{\beta}}{N + \hat{\alpha} - 1} \tag{30}$$

- Calculate the estimated motor vehicle insurance premium for the year $(t = 0$ and $N = 0)$ or for new policyholders, using the following equation:

$$\widehat{\text{premium}}_{1,0} = \frac{(2+\delta)}{(\delta)(1+\delta)} \cdot \frac{\hat{\beta}}{\hat{\alpha}-1} \tag{31}$$

- Calculate the estimated motor vehicle insurance premium for the year $(t + 1)$ for the claim frequency of 0 to 4 times and the period of the year of coverage from the 1st to the 5th year using the equation:

$$\widehat{\text{premium}}_{t+1,N} = \frac{(N+1)(N+2+t+\hat{\delta})}{(t+\hat{\delta})(N+1+t+\hat{\delta})} \cdot \frac{\sum_{i=1}^N x_i + \hat{\beta}}{N + \hat{\alpha} - 1} \tag{32}$$

RESULT AND DISCUSSION

Goodness-of-Fit Test for the Poisson-Lindley Mixture Distribution

The test hypothesis is formulated as follows:

H_0 : The claim frequency data of motor vehicle insurance with comprehensive coverage extended by TPL category 2 for the 2019 underwriting year comes from a population that follows a Poisson-Lindley mixture distribution.

H_1 : The claim frequency data of motor vehicle insurance with comprehensive coverage extended by TPL category 2 for the 2019 underwriting year does not come from a population that follows a Poisson-Lindley mixture distribution.

Estimation of the parameters of the Poisson-Lindley mixture distribution was sought using the Newton-Raphson iteration method with the help of RStudio software version 4.2.3. The initial value of the estimated Poisson-Lindley distribution parameter is calculated based on Equation (8), with the average value of the data as follows:

$$\bar{k} = \frac{\sum_{i=1}^n k_i}{n} = \frac{226}{56488} = 0.004$$

Thus, the initial value for the estimation of the parameters of the Poisson-Lindley mixture distribution is obtained:

$$\hat{\delta}_0 = \frac{-(\bar{k}-1) + \sqrt{(\bar{k}-1)^2 + 8\bar{k}}}{2\bar{k}} = \frac{-(0.004-1) + \sqrt{(0.004-1)^2 + 8(0.004)}}{2(0.004)} = 250.9925$$

Based on the results of calculations using the Newton-Raphson iteration method, in the 3rd iteration, an estimate of the Poisson-Lindley mixture distribution parameters for category 2 motor vehicle insurance claim frequency data of . This value can be used to calculate the estimated value of the probability for each category of motor vehicle insurance claim frequency using Equation (3). $\hat{\delta} = 250.939$. Table 1. presents the values required to calculate the statistical value of the test. The final result of calculating the chi-squared statistical value based on Equation (26) is found in column (5) of the last line, which is 0.0129.

Table 1. The values required in the statistical calculation of the test.

Claim frequency (k)	Number of Policies (O_i)	Claim Frequency Probability (p_k)	Expected Claim Frequency (E_i)	$\frac{(O_i - E_i)^2}{E_i}$
(1)	(2)	(3)	(4)	(5)
0	56263	0.9960	56262.901	1.76×10^7
1	224	0.0039	224.2024	0.0002
≥ 2	1	1.5816×10^{-5}	0.8934	0.0127
Sum	56488	1	56488	0.0129

With a significance level of 5%, the chi-squared quantile value with a degrees of freedom $1 = (3 - 1 - 1)$ is 3.8414 [9]. Because the statistical value of the test is smaller than the quantile value, the null hypothesis is accepted and it is concluded that the claim frequency data of motor vehicle insurance with comprehensive coverage extended by TPL category 2 for the 2019 underwriting year, comes from a population that follows a Poisson-Lindley mixture distribution.

Goodness-of-Fit Test for the Exponential-Inverse Gamma Mixture Distribution

The hypotheses used in this test are formulated as follows:

H_0 : The claim severity data of motor vehicle insurance with comprehensive coverage extended by TPL category 2 for the 2019 underwriting year comes from a population that follows a exponential-inverse gamma mixture distribution.

H_1 : The claim severity data of motor vehicle insurance with comprehensive coverage extended by TPL category 2 for the 2019 underwriting year does not come from a population that follows a exponential-inverse gamma mixture distribution. Estimation of the parameters of the exponential-inverse gamma mixture distribution is sought using the Newton-Raphson iteration method with the help of software Rstudio version 4.2.3. The initial value of the estimated parameter of the exponential-inverse gamma distribution parameter calculated based on Equation (24) for $\hat{\alpha}_0$ and Equation (25) for $\hat{\beta}_0$, with the mean values and variances of the claim severity data which are respectively $\bar{x} = 5760171.668$ and $s^2 = 5.1307 \times 10^{13}$. Thus, the initial value is obtained for the estimation of the parameters of the exponential-inverse gamma mixture distribution:

$$\hat{\alpha}_0 = \frac{-2s^2}{(\bar{x}^2 - s^2)} = \frac{-2(5.1307 \times 10^{13})}{((5.760.171,668)^2 - 5.1307 \times 10^{13})} = 5.66$$

Substitute the value in Equation (25), to obtain: $\hat{\alpha}_0$

$$\hat{\beta}_0 = \bar{x}(\hat{\alpha}_0 - 1) = 5760171.668(5.66 - 1) = 26843584.47$$

Based on the results of the calculation using the Newton-Raphson iteration method, in the 7th iteration, an estimate of the parameters of the mixture distribution was obtained exponential-inverse gamma for the claim severity data category 2 motor vehicle insurance claims are equal to $\hat{\alpha} = 6.4909$ and $\hat{\beta} = 31524867$. Table 2. presents the values needed to calculate the statistics of the Kolmogorov-Smirnov test. Based on the results of the calculation in Table 2., then the statistics of the Kolmogorov-Smirnov test can be calculated using Equation (28), namely:

$$D = \max_{1 \leq i \leq n} |F_n(x_i) - F^*(x_i)| = 0.0625.$$

With a significance level of 5%, the critical value is 0.0905 [10]. Since the statistical value of the Kolmogorov-Smirnov test is smaller than its critical value, the null hypothesis is accepted and it is concluded that the claim severity data of motor vehicle insurance with comprehensive coverage extended by TPL category 2 for the 2019 underwriting year comes from a population that follows a exponential-inverse gamma mixture distribution.

Table 2. Calculation Results for the Kolmogorov-Smirnov Test Claim Severity Data

<i>i</i>	<i>x_i</i>	<i>F_n(x_i)</i>	<i>F*(x_i)</i>	<i> F_n(x_i) - F*(x_i) </i>
(1)	(2)	(3)	(4)	(5)
1	300000	0.0044	0.0596	0.0552

2	325000	0.0088	0.0644	0.0556
3	385000	0.0133	0.0758	0.0625
4	386000	0.0221	0.0760	0.0583
5	386000	0.0221	0.0760	0.0538
:	:	:	:	:
226	50000000	1	0.9979	0.0021

Optimal Bonus-Malus System Premium Calculation

The determination of the optimal bonus-malus system premium is obtained using two components, namely the frequency of claims and the amount of claims. It was found that the claim frequency follows a Poisson–Lindley mixture distribution with an estimated parameter value of $\hat{\delta} = 250.939$, and the claim severity follows an exponential–inverse gamma mixture distribution with estimated parameter values of $\hat{\alpha} = 6.4909$ and $\hat{\beta} = 31524867$. For new policyholders ($t = 0, N = 0$):

$$\widehat{prem}_{1,0} = \frac{(2+\hat{\delta})}{(\hat{\delta})(1+\hat{\delta})} \cdot \frac{\hat{\beta}}{\hat{\alpha}-1} = \frac{(2+250.939)}{(250.939)(1+250.939)} \cdot \frac{31524867}{6.4909-1} = 22970$$

This value indicate that the contribution to the pure premium for the TPL extension that must be paid by new policyholders in the first year is Rp22,970. For policyholders who did not file a claim in the first year ($t = 1, N = 0$):

$$\widehat{prem}_{2,0} = \frac{(N+1)(N+2+t+\hat{\delta})}{(t+\hat{\delta})(N+1+t+\hat{\delta})} \cdot \frac{\sum_{i=1}^N x_i + \hat{\beta}}{N+\hat{\alpha}-1} = \frac{(0+1)(0+2+1+250.939)}{(1+250.939)(0+1+1+250.939)} \cdot \frac{0+31524867}{0+6.4909-1} = 22878$$

This value indicate that the contribution to the pure premium for the TPL extension that must be paid by policyholders who do not file a claim in the first year is Rp22,878. For policyholders who make a claim in the first year once ($t = 1, N = 1, x_1 = 6000000$):

$$\widehat{prem}_{2,1} = \frac{(1+1)(1+2+1+250.939)}{(1+250.939)(1+1+1+250.939)} \cdot \frac{6000000+31524867}{1+6.4909-1} = 46074$$

This value indicate that the contribution to the pure premium for the TPL extension that must be paid by policyholders who submit a claim in the first year once with a claim amount of Rp6,000,000 is Rp46,074. For policyholders who make a claim in the first year twice ($t = 1, N = 2, x_1 = 6000000, x_2 = 300000$):

$$\widehat{prem}_{2,2} = \frac{(2+1)(2+2+1+250.939)}{(1+250.939)(2+1+1+250.939)} \cdot \frac{(6000000+300000)+31524867}{2+6.4909-1} = 60362$$

Table 3. presents the amount of premiums that must be paid by policyholders, starting from the new policyholder to the amount of premiums for the following years. The severity of the claim is for example: the first claim is Rp6,000,000, the second claim is Rp300,000, the third claim is Rp10,000,000, and the fourth claim is Rp50,000,000.

Table 3. Premium Value Based on Claim Frequency and Claim Amount (Rupiah)

Year (t)	Number of Claims (N)				
	0	1	2	3	4
0	22,970				
1	22,878	46,074	60,362	89,775	205,353
2	22,788	45,891	60,123	89,419	204,538
3	22,697	45,709	59,885	89,065	203,730

4	22,608	45,529	59,649	88,715	202,928
5	22,519	45,351	59,416	88,367	202,132

CONCLUSION

This study applies the calculation of pure premiums within the bonus-malus system using a Poisson–Lindley mixture distribution for claim frequency and an exponential–inverse gamma mixture distribution for claim severity, based on motor vehicle insurance data with comprehensive coverage extended by TPL category 2 in Indonesia. The results indicate that the Poisson–Lindley distribution fits the claim frequency data well, and the exponential–inverse gamma distribution fits the claim severity data appropriately. The estimated pure premium for new policyholders in the first year is Rp22,970. The premium amounts in subsequent years are determined by the frequency and severity of the claims submitted. Premiums increase if claims occur and decrease if no claims are made. It is recommended that insurance companies offering TPL coverage consider implementing the bonus-malus system in premium calculations. Future research may explore alternative distributions for claim frequency and severity to evaluate more optimal approaches within the bonus-malus system.

REFERENCES

- [1] GoodStats, “Perkembangan jumlah kendaraan bermotor Indonesia, sepeda motor terbanyak,” *GoodStats.id*, 2024. [Online]. Available: <https://data.goodstats.id/statistic/perkembangan-jumlah-kendaraan-bermotor-indonesia-sepeda-motor-terbanyak-KC4IR>. [Accessed: Feb. 6, 2025].
- [2] Kompas.id, “Menyoal rencana wajib asuransi kendaraan bermotor di Indonesia,” *Kompas*, Jul. 26, 2024. [Online]. Available: <https://www.kompas.id/baca/riset/2024/07/26/menyoal-rencana-wajib-asuransi-kendaraan-bermotor-di-indonesia>. [Accessed: Feb. 6, 2025].
- [3] Y.-L. Grize, “Applications of statistics in the field of general insurance: An overview,” *Int. Stat. Rev.*, vol. 83, pp. 135–159, 2015, doi: 10.1111/insr.12066.
- [4] F. Eygenio and D. T. Qoyyimi, “Sistem bonus malus tergeneralisasi dengan frekuensi klaim berdistribusi binomial negatif dan besar klaim berdistribusi Pareto,” Undergraduate thesis, Dept. Stat., Univ. Gadjah Mada, Yogyakarta, 2019.
- [5] R. I. Adisti and A. K. Mutaqin, “Perhitungan premi murni pada sistem bonus malus untuk frekuensi klaim berdistribusi binomial negatif dan besar klaim berdistribusi Weibull pada data asuransi kendaraan bermotor di Indonesia,” *J. Gaussian*, vol. 10, no. 2, pp. 170–179, 2021, doi: 10.14710/j.gauss.v10i2.30084.
- [6] G. R. Sevina and J. Purwadi, “Bonus malus system for motorized vehicle insurance using geometric distributions and Weibull distributions,” *J. Fundam. Math. Appl.*, vol. 6, no. 1, pp. 71–79, 2023, doi: 10.14710/jfma.v6i1.16505.
- [7] M. E. Ghitany, B. Atieh, and S. Nadarajah, “Lindley distribution and its application,” *Math. Comput. Simul.*, vol. 78, no. 4, pp. 493–506, 2008, doi: 10.1016/j.matcom.2007.06.007.
- [8] S. Klugman, H. H. Panjer, and G. E. Willmot, *Loss Models: From Data to Decisions*, 4th ed. Hoboken, NJ: Wiley, 2019.
- [9] Sudjana, *Metode Statistika*, 6th ed. Bandung: Tarsito, 2005.
- [10] S. Nugroho, *Statistika Nonparametrika*. Yogyakarta: Graha Ilmu, 2008.