Jurnal Matematika dan Statistika serta Aplikasinya, Vol. 13, No. 2, 73-81, 2025
DOI:https://10.24252/msa.v13i2.60648

Premium Estimation Using a Spliced Gamma-Gamma Distribution for Long-Tail Insurance Claims

Erica Grace Simanjuntak[1]*, Nora Madonna[2], Ma'rufah Hayati[3]
[1] *Actuarial Science, Institut Teknologi Sumatera*
[2] *Actuarial Science, Institut Teknologi Sumatera*
[3] *Actuarial Science, Institut Teknologi Sumatera*
***Corresponding author**: erica.simanjuntak@at.itera.ac.id

***Submission date**: 08 August 2025, **Revision**: 20 August 2025, **Accepted**: 19 October 2025

ABSTRACT
Determining fair premiums that accurately reflect actual risks is a crucial element in insurance risk management, particularly when claim data exhibits long-tail characteristics that are challenging to model using a single distribution. This study aims to develop a premium estimation model using the spliced Gamma-Gamma distribution, which can capture the behavior of small to large claims more flexibly. This model is applied to a collective risk model framework, focusing on calculating the expected value and variance of aggregate claims as the basis for premium estimation. Premium estimation is conducted using three actuarial principles: the expected value principle, the variance principle, and the standard deviation principle. The research indicates that the standard deviation principle yields the most accurate premium estimation, as it accurately reflects the risk level while striking a balance between premium adequacy and affordability for policyholders. Based on the final calculations, the estimated premiums range from IDR 168,000 to IDR 197,000 per month for a five-year coverage period, which remains both sufficient to cover potential losses and competitive for policyholders. This approach considers both the expected loss and its volatility, making it more adaptive to extreme claim risks. This study demonstrates that claim modelling using splicing distributions, combined with volatility-based premium estimation principles, can be a practical and realistic approach to managing risk and estimating premiums more accurately .

KEYWORDS
Premium Estimation, Splicing, Gamma, Collective Risk Model.

1. INTRODUCTION

Splicing distribution refers to a modelling approach that combines two or more distinct probability distributions, each applied to specific intervals of the data domain [1]. This method offers enhanced flexibility in capturing complex data behaviors, especially when the data exhibit characteristics that cannot be adequately modelled by a single distribution [2],[3]. One notable application of spliced distributions is in the insurance industry, particularly for modelling claim severity data that displays long-tail behavior.

Long-tailed distributions are characterized by the probability of enormous values occurring in the data. In the context of insurance, this means that there may be a small number of claims with extremely high amounts, which can significantly impact premium estimation and risk management decisions [4]. Therefore, selecting a distribution model that accurately captures this behavior is essential in actuarial practice.

A common way to address this issue is by using a spliced two-component model, where different distributions are applied to different parts of the data, typically one distribution for the head segment (small to medium claim values) and another distribution for the tail segment (large claim values) [5]. In this study, the spliced Gamma–Gamma approach was used, which utilizes two Gamma distributions with different parameters to each segment. This model is selected due to its flexibility in fitting various data shapes and its ability to effectively model claim data with long tails.

In addition to fitting the data, it is also essential that the chosen model can be applied to practical applications, such as premium estimation that accurately reflects the actual risk profile. A commonly used approach is the collective risk model, where aggregate loss is modelled as a combination of claim frequency and claim severity. These two components are modelled separately, but both play a crucial role in estimating aggregate loss. Within this framework, selecting appropriate distribution models for both frequency and severity has a significant impact on the accuracy of premium estimates. Therefore, this study not only develops and analyses the spliced Gamma–Gamma model for claim severity but also derives its moment generating function, which is used in premium estimation under the collective risk model. The goal of this research is to provide a more accurate and realistic approach for estimating insurance premiums, based on actual claim data and the characteristics of long-tail distributions.

2. LITERATURE REVIEW

2.1 Splicing

Splicing is a modelling approach that constructs a new probability distribution by combining several different probability density functions (pdf) over specific intervals of the data domain [6]. In other words, one distribution is used to represent the behavior of losses within a certain range of values, while a different distribution is applied to another range [7]. This method enables more accurate modelling, especially for the tail region of loss distributions, which is often difficult to capture using a single standard distribution.

Suppose there are k probability density functions, denoted by $f_1(x), \ldots, f_k(x)$, each defined on a distinct interval. Then the spliced probability density function can be formulated as follows

$$f_X(x) = \begin{cases} p_1 f_1^*(x), & 0 \le x < w_1 \\ p_2 f_2^*(x), & w_1 \le x < w_2 \\ \vdots & \vdots \\ p_k f_k^*(x), & w_{k-1} \le x < w_k \end{cases} \tag{1}$$

where $p_i \ge 0$ for $i = 1, \ldots, k$ with $\Sigma_{i=1}^{k} p_i = 1, 0 < w_1 < w_2 \ldots < w_k$, and

$$f_i^*(x) = \frac{f_i(x)}{\int_{w_{i-1}}^{w_i} f_i(x)\, dx} \tag{2}$$

for $w_{i-1} \le x < w_i$ and $\int_0^\infty f_X(x)\, dx = 1$ [6].

To derive the MGF of a spliced distribution, the MGFs of the component distributions are utilized. According to [8] and [9], the MGF of the spliced distribution, denoted by $M_X(t)$, can be expressed as follows

$$M_X(t) = p_1 \int_0^{w_1} \exp(tx) f_1^*(x)\, dx + \ldots + p_k \int_{w_{k-1}}^{w_k} \exp(tx) f_k^*(x)\, dx \tag{3}$$

2.2 Aggregate Claim

The aggregate claim of a group of insurance policies represents the total accumulation of losses that occur within the group. This can be modelled using two main approaches, one of which is the collective risk model. This model consists of two key components: claim frequency and claim severity. Claim frequency refers to the number of claims incurred and is defined over the set of non-negative integers, making it a discrete random variable. On the other hand, claim severity represents the size of the claims that occur and is defined over the set of non-negative real numbers, thus modelled as a continuous random variable [6].

Let $X_1, \ldots, X_N$ be independent and identically distributed random variables representing the claim amounts and let N denote the random variable for the number of claims. The sum of these variables is denoted by S, representing the aggregate claim amount, and is defined as

$$S = X_1 + \ldots + X_N \tag{4}$$

with $S = 0$ when $N = 0$. The distribution of the random variable S is referred to as a compound distribution [10]. According to [6], the mean, variance, and MGF of S are given by

$$E[S] = E[X]E[N] \tag{5}$$

$$Var(S) = E[N]Var(X) + Var(N)E[X]^2 \tag{6}$$

$$M_S(t) = M_N(\ln(M_X(t))) \tag{7}$$

2.3 Premium Calculation Model

Premium is the amount of money paid by the policyholder to an insurance company in exchange for financial protection against potential future losses. One of the main challenges in the insurance industry is determining the appropriate premium that is accurately aligned with the risk undertaken. To achieve this, insurers must estimate the total claims that may occur, commonly by calculating the expected value of aggregate claims. If the premium is lower than the long-term expected value of total claims, the company risks incurring financial losses. Therefore, premium determination must include a loading factor, which adds a certain margin above the expected claim. The margin ensures that the premium is sufficient to cover potential losses while maintaining the insurer's financial stability.

In actuarial practice, several premium calculation principles are used that consider the uncertainty associated with claims. The most basic principle is the expected value principle, where the premium is calculated as the expected value of total claims plus a proportional risk loading. It is given by

$$\pi_E = (1+\theta)E[S] \tag{8}$$

where π denotes the premium, $E[S]$ is the expected value of the aggregate claims, and $\theta > 0$ is the loading factor.

Another method is the variance principle, which adjusts the premium based on the variance of the aggregate claims.

$$\pi_V = E[S] + \theta Var(S) \tag{9}$$

Here, the additional charge is proportional to the level of uncertainty in claims, measured by their variance.

Similarly, the standard deviation principle sets the premium based on the standard deviation, which serves as a measure of claim volatility

$$\pi_{STD} = E[S] + \theta\sqrt{Var(S)} \tag{10}$$

These three principles reflect different approaches to risk management, and the choice among them depends on the insurer's risk policy and the characteristics of the claim data being modelled [11], [12], [13].

3. METHODOLOGY

This study utilizes sample data from Advanced Referral Health Facilities (Fasilitas Kesehatan Rujukan Tingkat Lanjut, FKRTL) under National Health Insurance, BPJS Kesehatan, focusing on participants diagnosed with diabetes mellitus in the Special Region of Yogyakarta Province during the period 2015 to 2020. A total of 3760 participants met the inclusion criteria for analysis. The dataset includes information on the frequency of inpatient claims as well as the claim amounts (severity) submitted by each participant.

3.1 Research Procedure

This study was conducted through the following steps:

a. Data compilation was performed for participants diagnosed with diabetes mellitus in the Special Region of Yogyakarta . The dataset included the frequency and severity of inpatient claims.

b. Descriptive statistical analysis was conducted to explore the characteristics of the data. This included calculating summary measures such as the mean, median, variance, and standard deviation, as well as presenting tables and graphical illustrations for both claim frequency and severity.

c. Goodness-of-fit testing was carried out using the Kolmogorov-Smirnov test to determine the most suitable theoretical distributions for modelling claim frequency and severity.

d. Since the claim severity data exhibited long-tail characteristics, a spliced distribution model was applied. This involved identifying a threshold to separate small and large claims and fitting different distributions to each segment for more accurate modelling of both the body and the tail of the data.

e. Estimation of aggregate claims was performed by combining the selected models for claim frequency and severity. The expected value and variance of total claims were calculated, forming the basis for estimation.

f. Premiums were calculated using the expected value, variance, and standard deviation of aggregate claims, applying standard actuarial principles such as the expected value principle, variance principle, and standard deviation principle, with an appropriate loading factor to reflect risk.

g. Finally, conclusions were drawn based on the analysis results. The effectiveness of the spliced Gamma-Gamma distribution in modelling claim severity was assessed, and its implications for accurate premium estimation were discussed.

4. RESULT & DISCUSSION

4.1 Frequency and Severity Claim Distribution

The process of modelling data distributions involves identifying the probability distribution that best fits a given dataset. To determine the appropriate distributions to represent claim frequency and claim severity, the Kolmogorov-Smirnov test can be employed. This test evaluates the goodness of fit between the empirical data and a specified theoretical distribution. The null hypothesis (H_0) states that the data follows the specified distribution, while the alternative hypothesis (H_1) asserts that the data do not follow that distribution. At a predetermined significant level, the null hypothesis is rejected if the resulting p-value is less than the chosen significance threshold. Before performing distribution fitting, an overview of the distribution patterns of claim

Table 1. Descriptive Statistics of Claim Frequency

	Min	Q_1	Med	Mean	Q_3	Max
Claim Frequency	0.00	0.00	1.00	1.39	2.00	20.00

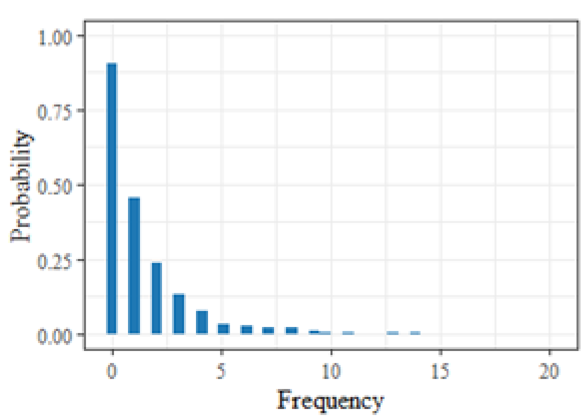

Figure 1. Histogram of Frequency Claim

frequency and claim severity is presented. **Table 1** and **Figure 1** provide a summary of the claim frequency data. It can be observed that some participants in the the Special Region of Yogyakarta diagnosed with diabetes did not make any inpatient claims during the period from 2015 to 2020. As a result, the claim frequency data are heavily concentrated around zero.

Despite this concentration, the histogram in **Figure 1** reveals long and thin tail characteristics, caused by several cases with relatively high claim frequency, reaching up to 20 claims per participant. The quantiles located on the left side of the distribution indicate that the data are right-skewed, with most observations clustered at lower values and a gradual tapering toward higher frequencies. This distribution pattern suggests the presence of over-dispersion, where a few participants contribute disproportionately to the total number of claims. These characteristics should be considered in selecting an appropriate statistical model for claim frequency, especially one that can accommodate skewness and variability in the data.

Table 2. Descriptive Statistics of Claim Severity (in millions)

	Min	**Q_1**	**Med**	**Mean**	**Q_3**	**Max**
Claim Severity	0.75	3.50	4.85	6.89	7.09	103.95

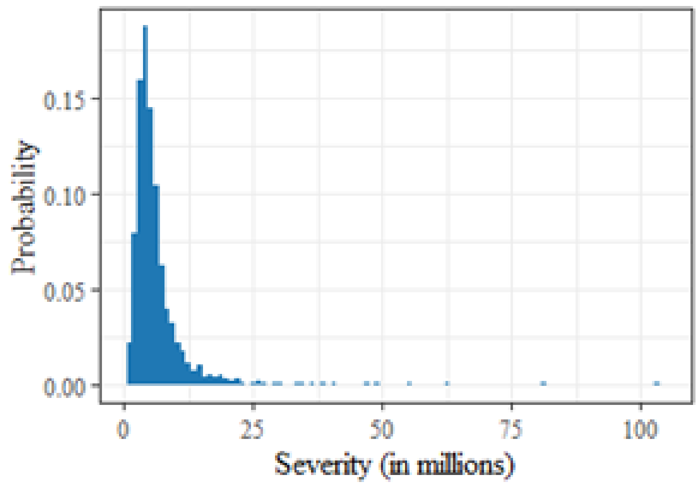

Figure 2. Histogram of Severity Claim

Table 2 and **Figure 2** present an overview of the claim severity data. Most claim values are concentrated on the left side of the distribution, representing relatively small claim amounts. However, the distribution also exhibits a long and thin right tail, indicating the presence of several claims with exceptionally high values. This pattern suggests that while most participants submitted low-value claims, a small number of participants filed claims with very high amounts, reaching IDR 103 million. Such a distribution pattern is characteristic of right-skewed, long-tailed data, which is commonly observed in insurance claim datasets.

The presence of a long tail makes it difficult to accurately fit the claim severity data using a single standard distribution. To address this, a splicing approach is applied in this study. The data are divided into two segments based on a selected threshold value, allowing each segment to be modelled with a different distribution. This method offers greater flexibility and accuracy in capturing the distinct behaviors of small and large claims. The result of this segmentation is shown in **Figure 3**, which displays a histogram of the claim severity data divided at the 0.9 quantile, corresponding to a threshold value of IDR 11.26 million.

To evaluate the goodness of fit between the proposed distribution models and the observed claim frequency and severity data, the Kolmogorov–Smirnov test was applied to assess how well the theoretical distributions matched the empirical data. This test was applied to the Negative Binomial distribution, commonly used in modelling claim frequency, and to the Gamma distribution for claim severity data [14]. The hypotheses for the Kolmogorof-Smirnov test are defined as follows: the null hypothesis (H_0) states that the data follow the specified theoretical distribution, while the alternative hypothesis (H_1) states that the data do not follow the specified theoretical distribution. At a significant level of α, the null hypothesis is rejected if the p-value is less than α. The resulting p-values are presented in Table 4.3. Using a significance level of 0.01, it was found that all p-values were greater than the specified significance threshold. This indicates that there is insufficient evidence to reject the null hypothesis H_0, meaning that the data does not significantly deviate from the assumed distributions. Therefore, it can be

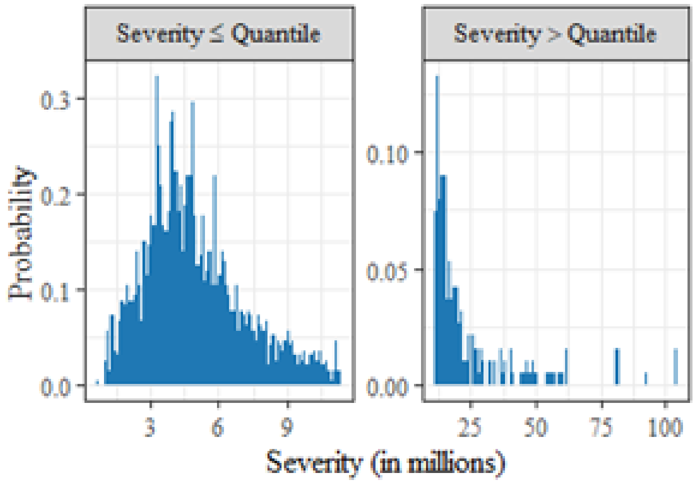

Figure 3. Histogram of Severity Claims Split at the 0.9 Quantile Threshold

Table 3. Distribution Fitting Results for Claim Frequency and Severity

	Distribution	**p-value**
Freq	Negative Binomial	0.38
	size = 0.66	
	μ = 1.39	
Sev $\leq Q_{0.9}$	Gamma	0.35
	shape = 5.15	
	rate = 1.04	
Sev $> Q_{0.9}$	Gamma	0.01
	shape = 2.97	
	rate = 0.12 (scale = 8.33)	

concluded that claim frequency follows the Negative Binomial distribution, while both the head and tail segments of the claim severity data follow Gamma distributions. After identifying the appropriate distributions for the head and tail of claim severity, these two segments were then combined to model the overall distribution of claim severity.

Let X_1 and X_2 represent the random variables corresponding to the two segments of claim severity data to be spliced. Suppose X_1 follows a Gamma distribution with parameters α_1 and β_1, and X_2 follows a Gamma distribution with parameters α_2 and β_2. Then, the probability density function (pdf) of the resulting spliced distribution based on **Equation** (1) is given by

$$f_X(x) = \begin{cases} \dfrac{p_1\beta_1^{-\alpha_1}x^{\alpha_1-1}\exp(-x/\beta_1)/\Gamma(\alpha_1)}{\displaystyle\int_0^w \beta_1^{-\alpha_1}x^{\alpha_1-1}\exp(-x/\beta_1)/\Gamma(\alpha_1)\,dx}, & 0 \leq x < w \\ \dfrac{p_2\beta_2^{-\alpha_2}x^{\alpha_2-1}\exp(-x/\beta_2)/\Gamma(\alpha_2)}{\displaystyle\int_w^\infty \beta_2^{-\alpha_2}x^{\alpha_2-1}\exp(-x/\beta_2)/\Gamma(\alpha_2)\,dx}, & x \geq w \end{cases}$$

or can be simplified to

$$f_X(x) = \begin{cases} \dfrac{p_1\beta_1^{-\alpha_1}x^{\alpha_1-1}\exp(-x/\beta_1)/\Gamma(\alpha_1)}{\gamma(\alpha_1, w/\beta_1)/\Gamma(\alpha_1)}, & 0 \leq x < w \\ \dfrac{p_2\beta_2^{-\alpha_2}x^{\alpha_2-1}\exp(-x/\beta_2)/\Gamma(\alpha_2)}{\Gamma(\alpha_2, w/\beta_2)/\Gamma(\alpha_2)}, & x \geq w \end{cases} \tag{11}$$

Based on the **Equation** (11), the corresponding MGF can be derived using **Equation** (3).

$$M_X(t) = \frac{\int_0^w \exp(tx)\, p_1 \frac{\beta_1^{-\alpha_1}}{\Gamma(\alpha_1)} x^{\alpha_1-1} \exp\left(-\frac{x}{\beta_1}\right) dx}{\frac{1}{\Gamma(\alpha_1)}\gamma\left(\alpha_1, \frac{w}{\beta_1}\right)} + \frac{\int_0^w \exp(tx)\, p_2 \frac{\beta_2^{-\alpha_2}}{\Gamma(\alpha_2)} x^{\alpha_2-1} \exp\left(-\frac{x}{\beta_2}\right) dx}{\frac{1}{\Gamma(\alpha_2)}\gamma\left(\alpha_2, \frac{w}{\beta_2}\right)} \tag{12}$$

The **Equation** (12) will be solved separately for each component.

$$\int_0^w \exp(tx)\, p_1 \frac{\beta_1^{-\alpha_1}}{\Gamma(\alpha_1)} x^{\alpha_1-1} \exp\left(-\frac{x}{\beta_1}\right) dx = \frac{\beta_1^{-\alpha_1}}{\Gamma(\alpha_1)} \int_0^w x^{\alpha_1-1} \exp\left(-x\left(\frac{1}{\beta_1}-t\right)\right) dx \tag{13}$$

Let $y = x\left(\frac{1}{\beta_1}-t\right)$, then

$$\begin{aligned}
\int_0^w \exp(tx)\, p_1 \frac{\beta_1^{-\alpha_1}}{\Gamma(\alpha_1)} x^{\alpha_1-1} \exp\left(-\frac{x}{\beta_1}\right) dx &= \frac{\beta_1^{-\alpha_1}}{\Gamma(\alpha_1)} \int_0^{w\left(\beta_1^{-1}-t\right)} \left(\frac{y}{\beta_1^{-1}-t}\right)^{\alpha_1-1} \exp(-y) \frac{1}{\beta_1^{-1}-t} dy \\
&= \frac{\beta_1^{-\alpha_1}}{\Gamma(\alpha_1)} \left(\frac{1}{\beta_1^{-1}-t}\right)^{\alpha_1} \int_0^{w\left(\beta_1^{-1}-t\right)} y^{\alpha_1-1} \exp(-y)\, dy \\
&= \frac{1}{\Gamma(\alpha_1)} \left(\frac{1}{1-\beta_1 t}\right)^{\alpha_1} \gamma\left(\alpha_1, w\left(\beta_1^{-1}-t\right)\right)
\end{aligned} \tag{14}$$

Using the same approach, the following result is obtained:

$$\int_w^\infty \exp(tx)\, p_2 \frac{\beta_2^{-\alpha_2}}{\Gamma(\alpha_2)} x^{\alpha_2-1} \exp\left(-\frac{x}{\beta_2}\right) dx = \frac{1}{\Gamma(\alpha_2)} \left(\frac{1}{1-\beta_2 t}\right)^{\alpha_2} \Gamma\left(\alpha_2, w\left(\beta_2^{-1}-t\right)\right) \tag{15}$$

Therefore, the MGF of the spliced random variable is given as follows:

$$M_X(t) = \frac{p_1 \frac{1}{\Gamma(\alpha_1)} \left(\frac{1}{1-\beta_1 t}\right)^{\alpha_1} \gamma\left(\alpha_1, w\left(\beta_1^{-1}-t\right)\right)}{\frac{1}{\Gamma(\alpha_1)}\gamma\left(\alpha_1, \frac{w}{\beta_1}\right)} + \frac{p_2 \frac{1}{\Gamma(\alpha_2)} \left(\frac{1}{1-\beta_2 t}\right)^{\alpha_2} \Gamma\left(\alpha_2, w\left(\beta_2^{-1}-t\right)\right)}{\frac{1}{\Gamma(\alpha_2)}\Gamma\left(\alpha_2, \frac{w}{\beta_2}\right)} \tag{16}$$

Based on the constructed spliced distribution, the expectation and variance of claim severity were calculated to provide insight into the average and variability of losses per claim. The expectation can be obtained by evaluating the first derivative of the MGF at zero or directly from the pdf. Similarly, the variance can be determined through the second derivative of the MGF or directly from the pdf. The results of these calculations show that the expected claim severity is IDR 7.14 million with a variance of IDR 117.60 million.

4.2 Aggregate Claim

Claim frequency and claim severity are the two primary components in calculating aggregate claims. After obtaining the expectation and variance of each component, the expected value and variance of aggregate claims can be calculated using **Equation** (5) and **Equation** (6). The results of these calculations are presented in **Table 4**.

Table 4. Expected Value and Variance of Aggregate Claims

	N	X	S
μ	1.39	7.14	9.92
σ^2	4.31	117.60	383.19

4.3 Premium Estimation

Premiums represent a form of compensation paid by policyholders in exchange for protection against financial risks, including those arising from natural disasters, which the insurance company may be bear. The amount of premium is determined based on collective risk, which consists of claim frequency and claim severity. In actuarial practice, premiums are calculated based on the expected value of total claims, to which a safety margin or loading factor is added. This ensures that the insurer can cover all claims while maintaining financial stability.

The loading factor typically ranges between 1% and 10%, depending on company policy and market competition [11]. Setting a loading factor that is too high may result in premiums becoming uncompetitive, potentially reducing customer interest and weakening the insurer's market position. Conversely, setting it too low increases the risk of long-term financial losses. Therefore, determining an optimal premium level is crucial to balancing a business's sustainability and market competitiveness. The calculated premiums, based on **Equation** (8), **Equation** (9), and **Equation** (10), are presented in **Table 5**.

Table 5. Premium Estimation (in millions)

θ	π_E	π_V	π_{STD}
1%	10.02	13.75	10.12
2%	10.12	17.58	10.31
3%	10.22	21.42	10.51
4%	10.32	25.25	10.70
5%	10.42	29.08	10.90
6%	10.52	32.91	11.09
7%	10.61	36.74	11.29
8%	10.71	40.58	11.49
9%	10.81	44.41	11.68
10%	10.91	48.24	11.88

Premium estimation was conducted by applying three actuarial principles: the expected value principle, the variance principle, and the standard deviation principle. The results show that premiums calculated under the expected value principle are relatively stable, ranging from IDR 10.02 million to IDR 10.91 million. In contrast, premiums derived using the variance principle exhibit a substantial increase, ranging from IDR 13.75 million to IDR 48.24 million, depending on the level of loading applied. This reflects the fact that the variance principle is susceptible to the level of claim uncertainty, resulting in significantly higher premiums, especially when a large loading factor is applied.

On the other hand, the application of the standard deviation principle results in more moderate and controlled premium estimates, ranging from IDR 10.12 million to IDR 11.88 million. These values are relatively close to those obtained under the expected value principle but incorporate a more realistic risk margin, as they account for claim volatility. Moreover, under the standard deviation principle, the increase in premiums across different loading levels is more proportional and does not exhibit the sharp jumps observed under the variance principle.

Considering the balance between premium adequacy to cover risk and competitiveness in the insurance market, the standard deviation principle appears to be the most appropriate approach in this context. The estimated premiums are calculated for a coverage period of five years, which, when converted to a monthly payment, amount to IDR 168,000 to IDR 197,000 . This range is considered reasonable and remains competitive from the policyholder's perspective, while still providing adequate risk protection for the insurance provider.

5. CONCLUSION

This study examined the modelling of claim severity using the spliced Gamma-Gamma distribution, which proved effective in capturing the long-tail characteristics commonly observed in claim data. The model was applied within the framework of the collective risk model for premium calculation, utilizing three actuarial principles: the expected value principle, the variance principle, and the standard deviation principle.

The results indicate that the standard deviation principle yields a more balanced premium estimate, ranging from IDR 168,000 to IDR 197,000 per month for a five-year coverage period. In contrast, the variance principle produced excessively high premiums, while the expected value principle did not adequately account for claim uncertainty.

However, this study has certain limitations. The data were limited to a single province and focused on a single disease; thus, the findings may not fully reflect broader conditions. Moreover, the splicing approach used in this study was restricted to a two-component Gamma-Gamma model; exploring alternative splicing forms may provide a more comprehensive framework for modeling extreme claims.

REFERENCES

[1] G. Gan and E. A. Valdez, “Fat-Tailed Regression Modeling with Spliced Distributions,” *North American Actuarial Journal*, vol. 22, no. 4, pp. 554–573, Oct. 2018.

[2] M. V. Wüthrich and M. Merz, *Statistical Foundations of Actuarial Learning and its Applications*, ser. Springer Actuarial. Cham: Springer Nature, 2023.

[3] T. Reynkens, R. Verbelen, J. Beirlant, and K. Antonio, “Modelling censored losses using splicing: A global fit strategy with mixed Erlang and extreme value distributions,” *Insurance: Mathematics and Economics*, vol. 77, pp. 65–77, Nov. 2017.

[4] J. Beirlant, Y. Goegebeur, J. Teugels, and J. Segers, *Statistics of Extremes: Theory and Applications*. England: John Wiley & Sons, Ltd, 2004.

[5] Z. Li, F. Wang, and Z. Zhao, “A new class of composite GBII regression models with varying threshold for modeling heavy-tailed data,” *Insurance: Mathematics and Economics*, vol. 117, pp. 45–66, Jul. 2024.

[6] Y.-K. Tse, *Nonlife Actuarial Models: Theory, Methods and Evaluation*, 2nd ed., ser. International Series on Actuarial Science Series. Cambridge: Cambridge University Press, 2023.

[7] S. A. Klugman, H. H. Panjer, and G. E. Willmot, *Loss Models: From Data to Decisions*, 5th ed. John Wiley and Sons, Inc, 2019.

[8] A. Bâcă and R. Vernic, “On the three-spliced Exponential-Lognormal-Pareto distribution,” *Analele Universitatii ”Ovidius” Constanta - Seria Matematica*, vol. 30, no. 3, pp. 21–35, Sep. 2022.

[9] A. Bâcă and R. Vernic, “Modeling Data with Extreme Values Using Three-Spliced Distributions,” *Axioms*, vol. 13, no. 7, p. 473, Jul. 2024.

[10] R. J. Gray and S. M. Pitts, *Risk Modelling in General Insurance: From Principles to Practice*, 1st ed. Cambridge University Press, Jun. 2012.

[11] F. I. Yusuf, P. Z. R. Adi, and T. E. T. Saragih, “Penerapan Collective Risk Model dalam Penentuan Premi Asuransi Bencana Alam,” *Euler J. Ilm. Mat. Sains dan Teknol.*, vol. 12, no. 2, pp. 213–219, Dec. 2024.

[12] W. H. Pramujati, “Pemodelan Premi Asuransi Bencana Banjir di Kabupaten Cilacap Jawa Tengah,” *juses*, vol. 8, no. 1, pp. 44–52, Feb. 2025.

[13] T. Mikosch, *Non-Life Insurance Mathematics: An Introduction with Stochastic Processes*, ser. Universitext. Dordrecht: Springer, 2006.

[14] P. Parodi, *Pricing in general insurance*, second edition ed. Boca Raton: Chapman & Hall/CRC, 2023.