Available online at https://ejournal.corespub.com/index.php/ijmsc/index

International Journal of Mathematics, Statistics, and Computing

e-ISSN 3025-0803

Vol. 2, No. 2, pp. 50-59, 2024

Determination of Insurance Premium Rates with Aggregation Claims at BPJS with Exponential and Gamma Distributions

Khalilah Razanah Zakirah[1*], Puspa Dwi Ayu Banowati[2]

[1,2]*Bachelor of Mathematics Study Program, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Sumedang, Indonesia, 45363*

**Corresponding author email: khalilah20001@mail.unpad.ac.id, puspa20004@mail.unpad.ac.id*

Abstract

Badan Penyelenggara Jaminan Sosial (BPJS) is a legal entity that has been provided by the government for the community with the aim of providing protection for all workers in Indonesia from certain socio-economic risks. National development, marked by planned and continuous strides, embodies a commitment to engage all societal, national, and state levels in fostering progress. Encompassing political, economic, socio-cultural, and defense and security realms, the development is meticulously designed to be comprehensive, targeted, integrated, gradual, and sustainable. The overarching objective is to catalyze an augmentation of national capabilities, align the standard of living for the Indonesian people with developed nations, and elevate overall welfare. To establish premium rates, a method involves multiplying the conditional expected value of claim frequency by the size of the claim, considering observed risk characteristics. A claim, in this context, constitutes a formal request to the insurance company, seeking payment in accordance with the terms of the agreement. The primary objective of this study is to establish the insurance premium rates applicable to policyholders (the insured) through the estimation of parameters in the distribution governing aggregate claims. This involves the distribution of both the number of claims and the size of the claims, and the estimation is performed using the moment method. Premium computations are executed based on two key principles: the pure premium principle and the expected value principle. This research produces the conclusion that the Poisson-Gamma aggregate claims distribution has a premium amount of 3.61 times greater than Poisson-Exponential due to the application of the anticipated value principle, namely IDR 4,403,542.94 per month and IDR 1,219,878.45 per month, respectively.

Keywords: distribution of aggregate claims, moment method, expected value principle, Poisson.

1. Introduction

The importance of national development as a series of planned and continuous steps reflects a commitment to involve all levels of society, nation and state in realizing progress. This development does not only cover political, economic, socio-cultural and defense and security aspects, but is also directed to be comprehensive, targeted, integrated, gradual and sustainable. The aim is to stimulate increased national capacity, bring the lives of Indonesian people on a par with developed countries, and improve welfare (Shaturaev, 2021; Khasandy & Badrudin, 2019).

In implementing national development, workers have a very important role as key actors in achieving development goals. Therefore, protection of workers and their families is a must in accordance with human values. Workers are defined as any individual who performs work and receives wages or compensation in various forms. This definition includes all levels of workers, whether individuals, partnerships, legal entities, or other entities, who receive wages or rewards as a form of appreciation for their contributions (Sagawa & Segal, 2000; Zelinsky, 2004; Rousseau & McLean Parks, 1993). This protection is an important foundation for ensuring sustainability and prosperity in the era of national development.

In the context of an Aggregation claims model, it is important to consider both the number and magnitude of risks. Claims can be differentiated into individual claims, which occur separately, and Aggregated claims, which are a combination of individual claims. Distribution of claims Aggregation can result from the distribution of the number of claims and the distribution of their magnitude (Zhang et al., 2019; Shi & Zhao, 2020). For the number of claims, the model generally uses discrete distributions such as the Binomial distribution, Negative Binomial distribution, Geometry distribution, and Poisson distribution. Meanwhile, for the size of the claim, the model generally adopts continuous distributions such as the Exponential distribution, Weibull distribution and Rayleigh distribution. This approach enables a holistic analysis of risks and claims in Aggregation scenarios, enriching the understanding of their distribution and characteristics.

One method that can be used to determine premium rates is to multiply the conditional expected value of the frequency of claims by the size of the claim and taking into account the observed risk characteristics (Denuit et al., 2007). A claim is a formal request to the insurance company to request payment based on the terms of the agreement.

This research aims to determine the insurance premium rates that must be paid by policy holders (insured) by estimating the parameters in the distribution that form aggregate claims (distribution of the number of claims and claim size) using the moment method. Premium calculations are carried out using the pure premium principle and the expected value principle.

2. Literature Review

In this research, statistical and actuarial theories are used is total compensation model, aggregate claim model, moment method, distribution for claim frequency, and pure premium principle.

2.1 Total Compensation Model

Total compensation per insured is denoted by the random variable S. for the large number of claims submitted by the insured in one period is denoted by N. Meanwhile each compensation is denoted by the random variable $X_1, X_2, \dots, X_N$. Total model Compensation per insured is given as in equation (1) below (Klugman, Panjer, & Wilmot, 208):

$$S = X_1 + X_2 + \cdots + X_N, \qquad N = 0,1,2,3, \dots \tag{1}$$

2.2 Aggregate Claim Model

An aggregate claims model is a mathematical approach or framework used in the insurance or financial risk industry to predict and manage the overall number of claims. This model helps insurance companies to understand potential risks and estimate the number of claims that may occur in the future.

Random variables for the number of claims and the size of the claim can be formed using a collective risk model with the following equation (2):

$$S = \sum_{i=1}^{N} X_i \tag{2}$$

The expectation value and variance of the aggregate claims model in equation (2) can be calculated with equation (3) and (4):

$$E[S] = E[X]E[N] \tag{3}$$

$$Var[S] = E[N]Var(X) + (E[X])^2 Var(N) \tag{4}$$

2.3 Moment Method

To determine an estimate or point estimate, there are several ways, one of which is to use moment method. The resulting equation is solved for get the estimated value of the parameter you are looking for. For example $X_1, X_2, \dots, X_N$ is a random sample from the population each with a function probability mass $f(x)$.

The population moment at k is:

$$\mu_k = E[X^k] \tag{5}$$

The sample moment at k is:

$$M_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \tag{6}$$

By equating equations (5) and (6), we will obtain an estimate of the parameters we are looking for.

2.4 Distribution for Claim Frequency

2.4.1 Poisson Distribution

The discrete random variable K can be said to have a Poisson distribution with $\lambda > 0$ if the probability function is as follows:

$$P(K = k) = p_k = \frac{e^{-\lambda}\lambda^k}{k!}, for\ k = 0,1,2,3, \dots \tag{7}$$

The expected value and variance of the Poisson distribution are:

$$E(K) = \lambda \tag{8}$$

$$Var(K) = \lambda \tag{9}$$

2.4.2 Exponential Distribution

The random variable X can be said to have an exponential distribution with parameter $\theta > 0$, so the probability density function is:

$$f(x) = \frac{1}{\theta}e; x > 0, \theta > 0 \tag{10}$$

The cumulative distributive function of the exponential distribution is:

$$F(x) = 1 - e^{-\frac{x}{\theta}}; x > 0, \theta > 0 \tag{11}$$

The expected value and variance of the exponential distribution are:

$$E(X) = \theta \tag{12}$$

$$Var(X) = \theta^2 \tag{13}$$

2.4.3 Gamma Distribution

The random variable N can be said to have an exponential distribution with parameter $\alpha\ and\ \beta$, so the probability function is:

$$f(N) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} N^{\alpha-1}; x > 0 \tag{14}$$

The expected value and variance of the exponential distribution are:

$$E(N) = \alpha\beta \tag{15}$$

$$Var(N) = \alpha\beta^2 \tag{16}$$

2.5 Pure Premium Principle

The pure premium principle is the basis for determining insurance premiums and is a key concept in the insurance industry. Based on the pure premium principle, the calculation of the premium with the symbol Π_s can be done by equation (17):

$$\Pi_s = E[S] \tag{17}$$

Based on the expected value principle, the calculation of the premium $\boldsymbol{\Pi_s}$ can be done by equation (18):

$$\Pi_s = (1 + \psi)E[S] \tag{18}$$

3. Materials and Methods

3.1 Materials

In this research, the data used comprises monthly claims for work accident insurance (JKK) from the BPJS Employment branch in Langsa for the period of January to December 2019. This data is sourced from a previous study conducted by Sumarni et al. From this data, information on the number of claims and the total claim amounts is extracted. The number of claims indicates the frequency of claim submissions made by insurance participants (policyholders) each year (Chrisan et al., 2019). On the other hand, the claim amount represents the compensation received by the insurance company for policyholders whose claims have been approved.

Table 1: The Total Number of Claims and The Claim Amounts of the Work Accident Insurance (JKK) of the BPJS Employment branch in Langsa of the year 2019

The nth Month	Number of Claims	Claim Amounts (IDR)
1	16	50,107,293
2	20	222,740,605
3	15	73,430,153
4	18	65,516,716
5	19	21,386,586
6	6	25,834,898
7	48	73,418,203
8	21	237,761,812
9	13	68,844,799
10	10	31,280,495
11	13	117,892,959
12	10	168,479,298
Total	209	1,156,693,817

Referring to Table 1, there were 209 claim submissions during the research period from January to December 2019. The total overall claim amount for that period was IDR 1,156,693,817. The largest claim submitted occurred in August, amounting to IDR 237,761,812, while the smallest claim took place in May, totaling IDR 21,386,586.

3.2 Methods

This research applies a methodology by integrating a literature review with the utilization of secondary data. Premium calculations are conducted by applying the principles of pure premium and the principle of expected value. Data is processed using statistical software with the aim of identifying the distribution in the data of claim numbers and claim amounts. With the results of data processing, distributions for claim number and claim amount data can be concluded.

Next, parameters from the data of claim numbers and claim amounts are estimated. Parameter estimation is done by applying the method of moments. The estimated parameter values are used to determine the expected value and variance of the distribution of claim numbers and total claims.

Using the expected value and variance from the distribution of claim numbers and claim amounts, the expected value and variance of the aggregate claim distribution are determined. The calculation of premium size is done based on the principles of pure premium and the principle of expected value. Establishing the premium size with the principle of expected value involves applying the Central Limit Theorem to determine the premium loading factor or the relative security loading factor ψ. This calculation utilizes the standard normal distribution approach.

The final step involves determining the monthly insurance premium to be charged to each participant. This approach is carried out by dividing the results of premium size calculations using the principles of pure premium and the principle of expected value by the number of claims that occurred during the research period. The assumption applied is that not all insurance participants will file claims, so the total premium received may exceed the calculated premium amount using both principles.

4. Results and Discussion

4.1 Number of Claims Distribution Model

This section discusses the types of distributions for the number of claims and the distribution for the claim amounts. The identification of a suitable distribution for the available data is carried out through hypothesis testing using the Kolmogorov-Smirnov and Anderson-Darling tests. This testing is conducted using statistical software.

To determine the distribution shape for the number of claims N, it can be tested using the Kolmogorov-Smirnov test (Rillifa and Aceng, 2021). The formulated hypothesis for the Kolmogorov-Smirnov test is:

H_0 : the number of claims follows a Poisson distribution.

H_1 : the number of claims does not follow a Poisson distribution.

The null hypothesis is accepted if the asymptotic significance value (*p-value*) is greater than the α level. Data processing was conducted with the assistance of SPSS software, and the output results can be observed in Figure 1.

One-Sample Kolmogorov-Smirnov Test

		Jumlah Klaim
N		12
Poisson Parameter[a,b]	Mean	17.42
Most Extreme Differences	Absolute	.242
	Positive	.242
	Negative	-.083
Kolmogorov-Smirnov Z		.838
Asymp. Sig. (2-tailed)		.484

Figure 1: The Output of the Kolmogorov-Smirnov Test Result

Based on the calculation output in Figure 1, the *p-value* obtained is 0.484. Therefore, H_0 is not rejected as the *p-value* $> \alpha = 0.05$. The Kolmogorov-Smirnov test results indicate that the number of claims follows a Poisson distribution.

To obtain the parameter values for the Poisson distribution, the method of moments can be applied. The first moment is calculated using the adjusted equation (6), yielding:

$$M_1 = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$M_1 = \frac{1}{12}(16 + 20 + 15 + 18 + 19 + 6 + 48 + 21 + 13 + 10 + 13 + 10) = 17.42$$

The expectation for a random variable X that follows a Poisson distribution with parameter λ is $E[X] = \lambda$. Therefore, based on the method of moments estimation for the parameter λ, the estimated value $\hat{\lambda}$ is given by $\hat{\lambda} = M_1 = 17.42$.

4.2 Claim Amounts Distribution Model

In determining the distribution model for claim amounts (variable X), the Anderson-Darling test can be conducted. Assumptions tested for the claim amounts distribution include exponential and gamma distributions. Calculations are performed using Minitab software, and the results output can be observed in Figure 2.

Descriptive Statistics

N	N*	Mean	StDev	Median	Minimum	Maximum	Skewness	Kurtosis
12	0	96391151	74688039	71131501	21386586	237761812	1,04843	-0,171935

Goodness of Fit Test

Distribution	AD	P
Exponential	0,584	0,360
Gamma	0,378	>0,250

Figure 2: The Output of the Anderson-Darling Test Result

Based on the calculation output in Figure 2, the *p-values* for the exponential and gamma distributions are found to be greater than the significance level $\alpha = 0.05$ .Therefore, the results of the hypothesis test indicate that the distribution of claim amounts data follows both the exponential and gamma distributions.

In estimating the values of parameters in the exponential and gamma distributions, the method of moments is applied. The parameter estimation (θ) for the exponential distribution is as follows:

$$\hat{\theta} = M_1 = \frac{1}{209}(50{,}107{,}293 + 222{,}740{,}605 + \cdots + 168{,}479{,}298) = 5{,}534{,}420.18$$

To estimate the parameters α and β in the gamma distribution, they can be obtained by using the first and second moments according to equation (2.6). The calculations for the first and second moments are as follows:

$$M_1 = \frac{1}{209}\sum_{i=1}^{n} X_i = \frac{1}{209}(50{,}107{,}293 + 222{,}740{,}605 + \cdots + 168{,}479{,}298) = 5{,}534{,}420.18$$

$$M_1 = \frac{1}{209}\sum_{i=1}^{n} X_i^2 = \frac{1}{209}(50{,}107{,}293^2 + 222{,}740{,}605^2 + \cdots + 168{,}479{,}298^2) = 8.27064 \times 10^{14}$$

The estimation for both parameters of the gamma distribution is

$$\hat{\alpha} = \frac{M_1^2}{M_2} = 0.0370343835 \text{ dan } \hat{\beta} = \frac{M_1}{\alpha} = 149{,}440{,}044{,}945$$

The results of the parameter estimation calculations for the distribution of the number of claims and claim amounts can be seen in Table 2

Table 2: The Parameter Estimation for the Number of Claims and Claim Amounts Distribution

Distribution	Parameter	Value
Poisson	λ	17.42
Exponential	θ	5,534,420.18
Gamma	α	0.0370343835
	β	149,440,044.945

4.3 The Estimation of the Aggregate Claims Model

Based on the results of calculations using the method of moments, it can be concluded that the number of claims follows a poisson distribution with a parameter $\lambda = 17.42$, while the claim amounts follow an exponential distribution with a parameter $\theta = 5{,}534{,}420.18$, and the claim amounts also follow a Gamma distribution with parameters $\alpha = 0.0370343835$ and $\beta = 149{,}440{,}044.945$.

The expected value and variance of the Poisson distribution (Bain and Engelhardt, 1992), are

$$E[X] = \lambda = 17.42$$

$$Var(X) = \lambda = 17.42$$

The expected value and variance of the exponential distribution are

$$E[N] = \theta = 5{,}534{,}420.18$$

$$Var(N) = \theta^2 = 3.06298 \times 10^{13}$$

The expected value and variance of the gamma distribution are

$$E[N] = \alpha\beta = 553{,}420.177$$

$$Var(N) = \alpha\beta^2 = 8.27064 \times 10^{14}$$

Referring to Equations (6) and (7), the expected value and variance of the aggregate claims model for both the poisson distribution and the exponential distribution are

$$E[S] = 17.42 \times 553{,}420.177 = 96{,}391{,}151.42$$

$$Var(S) = (553{,}420.177 \times 17.42) + (17.42^2 \times 3.06298 \times 10^{13}) = 9.29125 \times 10^{15}$$

Meanwhile, the expected value and variance for the poisson and gamma distributions are

$$E[S] = 17.42 \times 553{,}420.177 = 96{,}391{,}151.42$$

$$Var(S) = (553{,}420.177 \times 17.42) + (17.42^2 \times 8.27064 \times 10^{14}) = 2.50882 \times 10^{17}$$

with $E[S]$ representing the mean of the losses incurred by the insurance company each month, and $Var(S)$ as the measure of the distribution of losses experienced by the insurance company each month. The results achieved in subsection 4.3 are summarized in Table 3.

Table 3: Expected Value and Variance for the Distribution of the Number of Claims, Claim Amounts, and Aggregate Claims

Statistical Measures	Number of Claims Distribution	Claim Amounts Distribution		Aggregate Distribution	
	Poisson	Exponential	Gamma	Poisson Exponential	Poisson Gamma
Eexpected Value	17.42	553,420.177	553,420.177	96,391,151.42	96,391,151.42
Variance	17.42	3.06298×10^{13}	8.27064×10^{14}	9.29125×10^{15}	2.50882×10^{17}

4.4 Premium Calculation

In this study, the premium is calculated using two principles, namely the pure premium principle and the expected value principle. The premium calculation based on the pure premium principle can be determined using Equation (10) (Dickson, 2005), as follows

$$\Pi_s = E[S] = 96{,}391{,}151.42$$

The results of the calculations apply equally to both the poisson-exponential and poisson-gamma aggregate claims distributions.

By approximating the aggregate claims, the premium loading factor (ψ) is calculated using equation (12). Based on the standard normal distribution table at a significance level of $\alpha = 0.05$, the 95th percentile value of Z is obtained as 1.645. For the poisson-exponential aggregate claims distribution with $E[S] = 96{,}391{,}151.42$ dan $Var(S) = 9.29125 \times 10^{15}$, the calculation result are as follows:

$$\psi = 1.645 \times \frac{\sqrt{9.29125 \times 10^{15}}}{96{,}391{,}151.42} = 1.645$$

Meanwhile, for the poisson-gamma aggregate claims distribution with $E[S] = 96{,}391{,}151.42$ and $Var(S) = 2.50882 \times 10^{17}$, the results obtained are:

$$\psi = 1.645 \times \frac{\sqrt{2.50882 \times 10^{17},}}{96{,}391{,}151.42} = 8.548$$

Based on the expected value principle according to equation (11), the premium amount for the Poisson-Exponential and Poisson-Gamma aggregate claims distributions are as follows:

$$\Pi_s = 96{,}391{,}151.42 + (1.645 \times 96{,}391{,}151.42) = 254{,}954{,}596.3$$

$$\Pi_s = 96{,}391{,}151.42 + (8.548 \times 96{,}391{,}151.42) = 920{,}340{,}474.3$$

The outcomes of computing premiums using equation (10) are outlined in Table 3. Table 3 illustrates the premium loading factor (ψ) and the individual premium amount borne by insurance users, adhering to the principles of pure premiums and the expected value for 100 portfolios based on equation (10). The figures in Table 3 are derived from the equation. The calculation of premiums using the pure premium principle excludes the involvement of the premium loading factor, resulting in Table 4 lacking any premium loading factor values.

This study utilizes the total number of claims data that occurred during the research period from January to December 2019, totaling 209 claims. The claim amount in Table 4 column (4) is calculated on a monthly basis, so if divided by 209, it will yield the claim amount per person per month (column (5) with an asterisk).

Table 4: The Value of The Premium Loading Factor and The Premium Amount per Person

Aggregate Claims Distribution	Premium Calculation Principles	Premium Loading Factor (ψ)	Premium Amount (IDR)	Premium Amount per person per month*) per year **) (IDR)
(1)	(2)	(3)	(4)	(5)
Poisson and Exponential	The Principle of Pure Premium	-	96,391,151.42	461,201.68*) 5,534,420.18**)
	The Principle of Expected Value	1.645	254,954,596.3	1,219,878.45*) 14,638,541.41**)
Poisson and Gamma	The Principle of Pure Premium	-	96,391,151.42	461,201.68*) 5,534,420.18**)
	The Principle of Expected Value	8.548	920,340,474.3	4,403,542.94*) 52,842,515.27**)

Note that the premium amount in column (4) is divided by the total number of claims (209 claims) to determine the premium amount per month or per year. If the premium amount in column (4) is divided by the number of participants, it will be less because, in reality, there will be more insurance participants than claims. The best course of action is to presume that all premiums are used to cover claims, as this would put a strain on the corporation. As a result, the premium amount in column (4) needs to be divided by 209, the total number of claims.

According to the calculation results, applying the moment technique results in the same premium amount, or IDR 461,201.68 each month or IDR 5,534,420.18 for the year 2019, for both the Poisson-Exponential and Poisson-Gamma aggregate claims distributions when based on the pure premium principle. Even if the moment approach is still used to derive parameter estimations, applying the expected value principle yields very different results. The premium amount for the Poisson-Gamma aggregate claims distribution is IDR 4,403,542.94 each month, which is 3.61 times more than the Poisson-Exponential, which has a value of IDR 1,219,878.45 each month, according to the expected value principle calculations

5. Conclussion

According to the study's findings, the claim quantity data is distributed gamma and exponentially, whereas the claim frequency data has a Poisson distribution. For aggregate claims, this means that Poisson-Exponential and Poisson-Gamma distributions apply. The Poisson, Exponential, and Gamma distribution parameters are computed using information from earlier studies by Sumarni et al. The moment method is the technique used to estimate parameters. The pure premium concept applied to both constructed aggregate claims distributions yields the same amount of premium. In the meantime, the Poisson-Gamma aggregate claims distribution has a premium amount that is 3.61 times bigger than the Poisson-Exponential due to the application of the anticipated value principle. Insurance firms may take these premium estimates into account while managing the BPJS Employment workers' accident insurance fund (JKK).

References

Alfaridzi, S. I. R., & Prabowo, A. (2023). Determination of Premium Rates in Vehicle Insurance with Exponential and Gamma Distributed Claims. Premium Insurance Business Journal, 10(1), 29-41.

Chrisan, K.W., Altien, J.R., and Tohap, M. (2019). Distribution Model of Car Insurance Claim Data to Determine Pure Premium. d'Cartesian, Journal of Mathematics and Applications, Vol. 8, No. 2, 108-113.

Denuit, M., Maréchal, X., Pitrebois, S., & Walhin, J. F. (2007). *Actuarial modelling of claim counts: Risk classification, credibility and bonus-malus systems*. John Wiley & Sons.

Khasandy, E. A., & Badrudin, R. (2019). The influence of zakat on economic growth and welfare society in Indonesia.

Rillifa, I.A., and Aceng, K.M. (2021). Calculation of Pure Premium in Malus Bonus System for Negative Binomial Distributed Claim Frequency and Weibull Distributed Claim Amount on Motor Vehicle Insurance Data in Indonesia. Gaussian Journal, Vol. 10, No. 2, 171-179.

Rousseau, D. M., & McLean Parks, J. (1993). The contracts of individuals and organizations. *Research in organizational behavior*, *15*, 1-1.

Sagawa, S., & Segal, E. (2000). Common interest, common good: Creating value through business and social sector partnerships. *California management review*, *42*(2), 105-122.

Shaturaev, J. (2021). indonesia: superior policies and management for better education (Community development through Education). *Архив научных исследований*, *1*(1).

Shi, P., & Zhao, Z. (2020). Regression for copula-linked compound distributions with applications in modeling aggregate insurance claims. *The Annals of Applied Statistics*, *14*(1), 357-380.

Sumarni, Muliani, F., & Amelia. (2022). Number of Poisson-Distributed Claims and Rayleigh-Distributed Claims with an Aggregation Claim Model at BPJS Ketenagakerjaan Langsa Branch. UNAND Journal of Mathematics, 11(4), 258-270.

Utami, D. A., Mutaqin, A, K., & Wachidah, L. (2017). Determination of Aggregate Loss Distribution of Motor Vehicle Insurance Insured in Indonesia Using the Panjer Recursive Method. Proceedings of Statistics, 3(11), 50-57.

Zelinsky, E. A. (2004). Defined Contribution Paradigm, The. *Yale LJ*, *114*, 451.

Zhang, Y., Zhao, P., & Cheung, K. C. (2019). Comparisons of aggregate claim numbers and amounts: A study of heterogeneity. *Scandinavian Actuarial Journal*, *2019*(4), 273-290.