



# The use of monte carlo method to model the aggregate loss distribution

Rafika Septiany<sup>1</sup>, Berlian Setiawaty<sup>1</sup>, I Gusti Putu Purnaba<sup>1</sup>

<sup>1</sup> Institut Pertanian Bogor, Indonesia

✉ [rafika\\_2016@apps.ipb.ac.id](mailto:rafika_2016@apps.ipb.ac.id)

## Article Information

Submitted June 16, 2020

Revised June 20, 2020

Accepted June 24, 2020

## Keywords

Aggregate Loss; Monte Carlo; Lognormal distribution; Z12M–NBGE distribution.

## Abstract

Based on Law Number 24 of 2011, a state program was established to provide social protection and welfare for everyone, one of which is health insurance by the Social Insurance Administration Organization (BPJS). In its implementation, several important evaluations are needed. One that requires accurate evaluation is claim frequency and claim severity in determining premiums and reserved funds. This thesis provides one form of a method for selecting the distribution of claim frequency and claim severity. The data used in this study was taken from BPJS Health in the City of Tangerang in 2017. The distribution of opportunities chosen had been adjusted to the participant's claim data and parameter estimated using the Maximum Likelihood Estimation method. The chi-square test was used to check the goodness of fit for claim frequency distributions whereas the Anderson Darling tests were applied to claim severity distributions. The results of the chi-square test and the Anderson-Darling test showed that the model that matched the claim frequency distribution was the Z12M–NBGE distribution while the model that matched the claim severity was lognormal. The Z12M–NBGE distribution and the lognormal formed the aggregate loss distribution using the Monte Carlo method. Furthermore, the simulation results were obtained to the measurement of the Value in Risk (VaR) and Shortfall Expectations (ES). So, the Monte Carlo method is simple to implement the aggregate loss distributions and can easily handle various risks with dependency.

## INTRODUCTION

Government attention toward health and quality of the health of its citizens can be seen in the 1945 Constitution article 28 H paragraph (1) which reads "Every person has the right to live in physical and spiritual prosperity, to live and to have a good and healthy life and to have the right to health services". Referring to the law number 24 of 2011, a Social Insurance Administration Organization or BPJS was established as an institution to organize the national social security program. This state program aims to provide certainty of protection and social welfare for all people, namely BPJS Health and BPJS Employment. BPJS Health is a government program on health insurance with the principle of social insurance and the principle of equity. BPJS Employment covers the work accident insurance, old-age insurance, pension insurance, and life insurance.

A general approach to modeling claim data is to separate the claim frequency from the large claims. The claim severity refers to the total number of policyholder claims per time whereas the claim severity is the costs incurred per claim. Many policyholders do not submit claims. Cases like this produce a zero claim number with high probability. Poisson distribution is the archetype of modeling the claim frequency (Antonio et al., 2010) although, in practice, the amount of data observed often displays features such as overdispersion (variance values are greater than expectations) or less spread (variance values are less than expectations) that are

common in applied data analysis, (Reiner, 2000). The negative binomial distribution is better for excess data that are not always heavy-tailed (Wang, 2011). Very heavy tails imply excessive spread but the reverse does not apply. Negative Binomial Generalized Exponential (NBGE) distribution is an alternative distribution for Poisson distribution and negative binomial distribution, especially in analyzing discrete data that has over-dispersion with excess zeroes (Bodhisuwan, & Aryuyuen, 2013).

The actuarial model for claims is based on the continuous distribution (Klugman et al., 2012). The selection of continuous distribution is based on heavy and pointed tail samples (Pacakova, & Brebera, 2016). Lognormal and gamma distributions are the most commonly used distributions for large claims models. Another distribution for claim severity is logistic distribution.

Aggregate loss modeling is an important task for health insurance companies to predict future events (Hua, 2015). Aggregate loss modeling is a compound distribution of the claim frequency and the claim severity. Furthermore, there are three analytical solutions to determine the distribution of aggregate loss, namely Monte Carlo, recursive, Panjer, and Fast Fourier Transform (FFT) (Shevchenko, 2010).

Furthermore, to determine the distribution of compounds using the Monte Carlo method, the software was used called Mathematica version 12. The Monte Carlo method is easy to understand and quite accurate even for very complicated calculations. After the aggregate loss distribution had been obtained from the insurance company, the risk amount was calculated. The measurements of risk referred to were the Value at Risk (VaR) and the shortfall Expectations (ES). The amount of the risk was calculated using the Monte Carlo method.

The purposes of this study were to 1) determine the best discrete distribution for the distribution of the claim frequency spread from the data of BPJS Health in the City of Tangerang in 2017, 2) determine the best continuous distribution for the claim severity distribution, and 3) to determine the aggregate distribution loss from which is the distribution of compounds from  $X$  and  $N$  with the Monte Carlo method. Finally, to determine the Value at Risk (VaR) and Shortfall Expectation (ES) with the Monte Carlo method to measure the magnitude of the risk that might occur.

## **METHODS**

### **Data**

The data used to model the distribution of aggregate loss was the inpatient data of the participants of BPJS Health in the City of Tangerang in 2017. In this study, the number of participants was assumed to be around 70% of the total population of 1,500,000 people.

### **Research Procedure**

This research modeled the distribution of the frequency claim and severity claim. The best distribution of claims and the largest claim frequency had been chosen. There were three rules to adjust the statistical distribution (Achieng, & No, 2010) first, selected claims distribution family where the discrete claims distribution illustrated the distribution of the claim frequency and the continuous distribution represented the claim severity distribution. Second, we estimated the distribution parameters accordingly. Third, tested the goodness of fit of the

distribution. Chi-square test was used to check the goodness of fit-test for the claim frequency distribution and the Anderson – Darling test was applied to the claim severity distribution.

Next, determine the aggregate loss distribution for health insurance companies which is a compound distribution of the claim frequency and the claim severity. The distribution of compounds was determined by the Monte Carlo method. The software used was Mathematica version 12. After obtaining the aggregate loss distribution from the Monte Carlo method, the risk amount was calculated. The measures of risk referred to were the Value at Risk (VaR) and Shortfall Expectations (ES).

## RESULTS AND DISCUSSION

### Model of Data Distribution of the Claim frequency of BPJS Health in the City of Tangerang in 2017

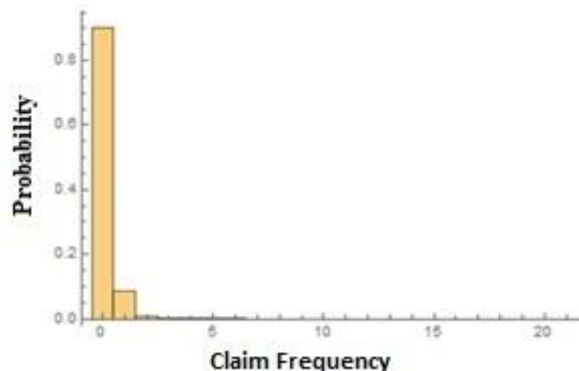
#### Data Description

The claim frequency data used in this study was the data on the numbers of inpatient claims of BPJS Health in the City of Tangerang in 2017. There were 1,500. 000 claims data activated at the BPJS Health in the City of Tangerang in 2017. The following is the statistical description of the claim frequency data of the BPJS Health in the City of Tangerang in 2017.

**Table 1.** Statistical Description of Participants Claims of BPJS Health in the City of Tangerang in 2017

Mean	Variance	Skewness	Kurtosis
0.11553	0.15767	5.85389	81.31400 Skewness

The positive value shown in the distribution is tilted to the left and has a long tail to the right. This can also mean that many BPJS Health participants did not submit claims or many BPJS participants were not hospitalized. Kurtosis values greater than three indicate that the distribution has a tapered curve. This can be seen in Figure 1 which shows the asymmetrical shape of the histogram indicating that the data are not normally distributed.



**Figure 1.** Histogram of the Claim Frequency of BPJS Health in the City of Tangerang in 2017

#### Estimator Value of Parameters and Mass Function Opportunities

Based on the data description, the researchers first chose a negative binomial distribution (NB) because the variance value was greater than the mean. The choice of the second distribution was the negative distribution of binomial generalized exponential (NBGE) this was because the data had a probability of no large claims. Furthermore, the estimator value for the parameters was the maximum likelihood estimator (MLE) method, the negative binomial probability

distribution (NB), and negative binomial generalized exponential (NBGE) probability function. The parameters can be seen in Table 2.

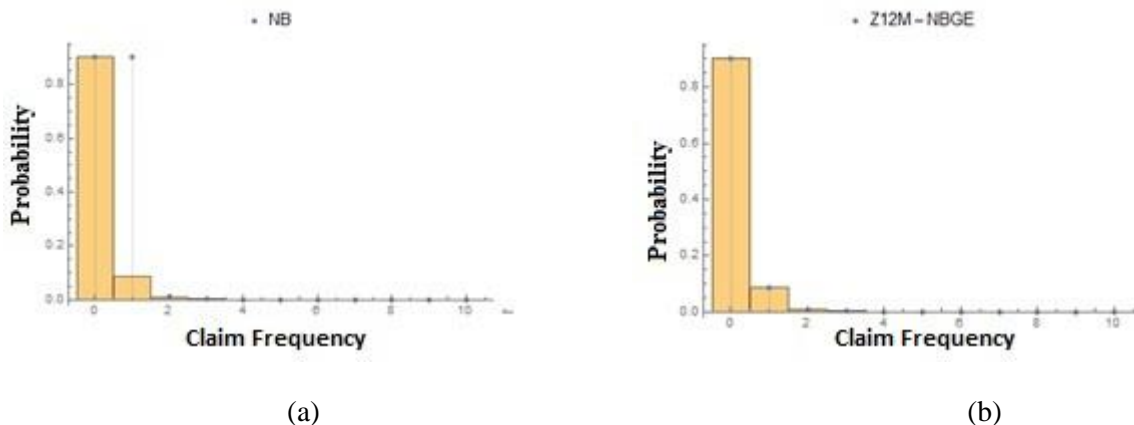
**Table 2.** The Estimated Parameter Values of the Negative Binomial Distribution (NB) and Negative Distribution of Binomial Generalized Exponential (NBGE).

Distribution	Mass Function Opportunities	Parameter Estimates
Negative Binomial	$p_k = \binom{k+r-1}{k} \left(\frac{1}{1+\beta}\right) \left(\frac{\beta}{1+\beta}\right)^k r,$ $k = 0,1,2,\dots, r > 0 \text{ and } \beta > 0$	$\hat{r} = 0.43998$ $\hat{\beta} = 0.26257$
Negative Binomial Generalized Exponential	$p_k = \binom{k+r-1}{k} \sum_{j=0}^k \binom{k}{j} (-1)^j \left( \frac{\alpha(+1)\Gamma\left(1 + \frac{r+j}{\beta}\right)}{\alpha\left(\Gamma + \frac{r+j}{\beta} +\right)} \right)_1$ $k = 0,1,2,\dots, r > 0, \alpha > 0 \text{ and } \beta > 0$	$\hat{r} = 1.06317$ $\hat{\alpha} = 1.48856$ $\hat{\beta} = 12.8549$

Furthermore, the researchers modified the distribution class on the negative binomial distribution (NB) and generalized exponential negative binomial distribution (NBGE), namely zero modified negative binomial (ZM-NB), zero one modified negative binomial distribution (Z1M – NB), zero modified negative binomial generalized exponential distribution (ZM-NBGE), zero one modified negative binomial generalized exponential distribution (Z1M – NBGE), and zero one two modified negative binomial generalized exponential distribution (Z12M)-NBGE).

**Graphical Model Analysis**

The parameters had been estimated with the Probability mass function of each distribution. The researcher wanted to test the claim frequency model graphically. It can be seen in Figure 2.



**Figure 2.** Histogram of the Claim Frequency of BPJS Health in the City of Tangerang in 2017 with the Probability Mass Function Plot for the Distribution of (a) NB and (b) Z12M -NBGE

Figure 2 shows (b) the distribution of Zero One Two Modified Negative Binomial Generalized Exponential (Z12M-NBGE) that best matched the histogram with the probability mass function plot. This shows that graphically, the distribution of Zero One Two Modified

Negative Binomial Generalized Exponential (Z12M-NBGE) is most suitable for modeling the claim frequency distribution of BPJS Health in the City of Tangerang in 2017.

**Mathematical Model Analysis**

Zero One Two Modified Negative Binomial Generalized Exponential (Z12M – NBGE) was the most graphically suitable claim distribution model. Furthermore, the distribution was examined mathematically with the chi-square test. The formulated hypotheses were:

**H<sub>0</sub>**: Data on the claim frequency of BPJS Health in the City of Tangerang in 2017 followed a specific distribution.

**H<sub>1</sub>**: Data on the claim frequency of BPJS Health in the City of Tangerang in 2017 did not follow a specific distribution.

By using the estimated parameters and the probability mass function formula for each distribution, the results are presented in Table 3.

**Table 3.** Expectations of the Claim frequency from the NB Distribution and the Z12M – NBGE Distribution.

<i>k</i>	Claim frequency	Expectations of the Claim frequency (NB)	Expectations of the Claim frequency (Z12M-NBGE)
0	1352907	1353757	1352907
1	128662	123869	128662
2	13940	18547	13940
3	2902	3137	3393
4	873	561	785
5	343	104	213
6	160	20	65
7	83	4	22
8	47	1	8
9	29	0	3
10	15	0	1
11	13	0	1
12	8	0	0
13	6	0	0
14	3	0	0
15	4	0	0
16	3	0	0
17	0	0	0
18	1	0	0
19	0	0	0
20	0	0	0
21	1	0	0
Total	1500000	1500000	1500000
Khi-squared		2524.27	6.56 x 10 <sup>-5</sup>
Degrees of freedom		3	2
<i>p-value</i>		0	0.99

Based on Table 3, it can be seen that the p-value distribution of Zero One Two Modified Negative Binomial Generalized Exponential (Z12M-NBGE) is 0.99, which means the opportunity not to reject the incorrect hypothesis  $H_0$  is 0.01. It is certainly not to reject the hypothesis  $H_0$  because the probability of error is quite small. Therefore, the hypothesis testing conducted concluded that the distribution that was most suitable for modeling the claim frequency of BPJS Health in the City of Tangerang in 2017 was the Zero One Two Modified Negative Binomial Generalized Exponential (Z12M-NBGE) distribution.

**DATA DISTRIBUTION MODEL OF THE CLAIM SEVERITY OF BPJS HEALTH IN THE CITY OF TANGERANG IN 2017**

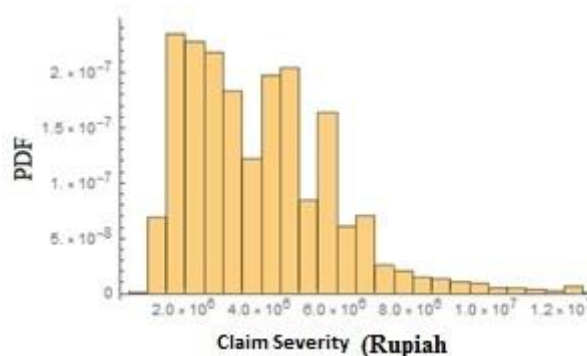
**Description of Data**

Data on the claim severity used in the study was the claims data of BPJS Health in the City of Tangerang in 2017. They were 173.001 selected data which covered the claim severity of the inpatient coverage. The following is the statistical description of the claim severity data of BPJS Health in the City of Tangerang in 2017.

**Table 4.** Statistical Description of the Claim Severity of BPJS Health in the City of Tangerang in 2017

Mean	Variance	Skewness	Kurtosis
Rp4,548,066.00	$2,304 \times 10^{13}$	7,435	83,026 Positive

The skewness values indicate that the distribution of skewness is tilted to the left and has a long tail to the right as can be seen in the histogram Figure 3. This can also mean the severity of the claims was mostly at a value less than the average covered. The kurtosis value was greater than three. It indicated that the distribution had a tapered curve. It can be seen in Figure 3 which shows the asymmetrical shape of the histogram so that it indicates that the data was not normally distributed.



**Figure 3.** Histogram of the Severity Claim of BPJS Health in the City of Tangerang in 2017

**Estimator Value of Parameters and Probability Density Functions**

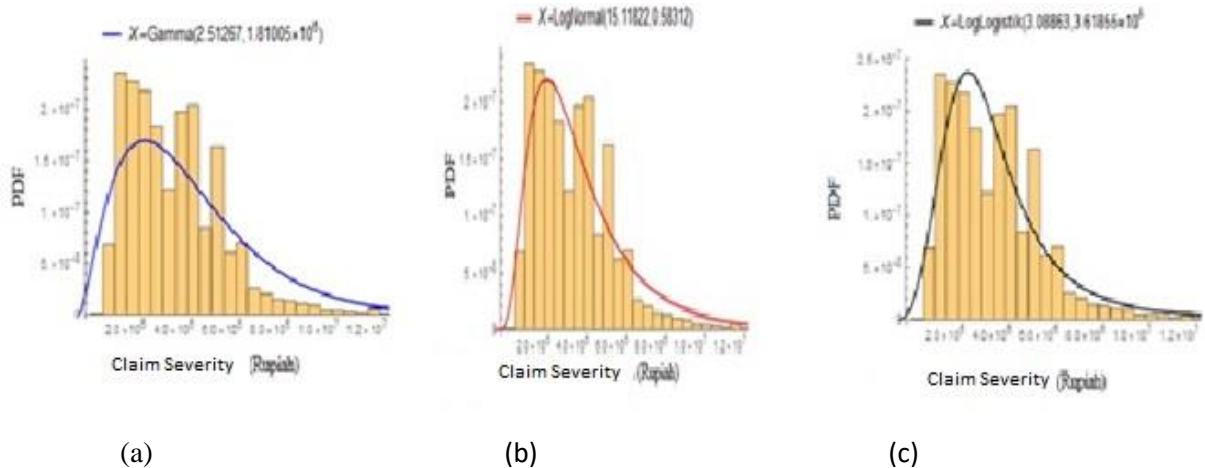
The claim severity is modeled using a positive continuous opportunity distribution (Klugman et al. 2012). In this study, the researchers chose the gamma distribution, lognormal distribution, and logistical distribution to estimate the claim severity, distribution model. Furthermore, the estimator values for the parameters were the maximum likelihood estimation method (MLE) and the probability density function of the gamma distribution, lognormal, distribution, and logistical distribution. The data can be seen in Table 5.

**Table 5.** The Estimator Value of Parameters and Probability Density Functions of Gamma Distribution, Lognormal Distribution, and Logistical Distribution.

Distribution	Probability Density Functions	Parameter Estimates
Gamma	$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\lambda x), x \geq 0$	$\hat{\alpha} = 2.51267$ $\hat{\lambda} = 1.81005 \times 10^6$
Lognormal	$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right),$ $x > 0, -\infty < \mu < \infty, \sigma > 0$	$\hat{\mu} = 15.11822$ $\hat{\sigma} = 0.58312$ Logistical
analysis	$f(x) = \frac{\left(\frac{\beta}{\alpha}\right) \left(\frac{x}{\alpha}\right)^{\beta-1}}{\left(1 + \left(\frac{x}{\alpha}\right)^\beta\right)^2},$ $x \geq 0, \alpha > 0, \beta > 0$	$\hat{\alpha} = 3.08863$ $\hat{\beta} =$ 3.61855Estimated $\times$ $10^6$

**Graphical Model Analysis**

The researcher tested the model for the claim frequency graphically which can be seen in Figure 4.



**Figure 4.** Histogram Data on the Claim severity and the Curve of the Probability Density Function for the distribution of (a) Gamma, (b) Lognormal, (c) Logistical

Figure 4 shows that (b) lognormal distribution was the most appropriate for the plot of the probability density function. This shows that graphically, the lognormal distribution was the most suitable for modeling the distribution of participants’ claims of BPJS in the City of Tangerang in 2017.

**Mathematical Model Analysis**

The lognormal distribution is a distribution model for the claim severity that is most graphically appropriate. Furthermore, the distribution will be examined mathematically using the Anderson Darling test. The formulated hypotheses are:

$H_0$ : Data on the claim severity of BPJS Health in the City of Tangerang in 2017 followed a specific distribution.

$H_1$ : Data on the claim severity of BPJS Health in the City of Tangerang in 2017 did not follow a specific distribution.

**Table 6.** Anderson Darling Value and p-value for the Claim severity of BPJS Health in the City of Tangerang in 2017

Distribution	Anderson-Darling Value	p-value
Gamma	0.81624	0.06458
Lognormal	0.64303	0.60827
Logistical	Logistic	2.28236 0.46947

From the results of the calculations in Table 6, using the Anderson Darling test can be seen that the distribution of lognormal with an AD value of 0.64303 and p-value of 0.60827. The probability of not rejecting the incorrect  $H_0$  hypothesis is 0.39173. This gives the decision not to reject the hypothesis  $H_0$  because the probability of error is quite small between the gamma distribution and the loglogistic distribution. Therefore, the hypothesis testing conducted concluded that the distribution that was most suitable for modeling the magnitude of the claims of BPJS Health participants in the City of Tangerang in 2017 was the lognormal distribution.

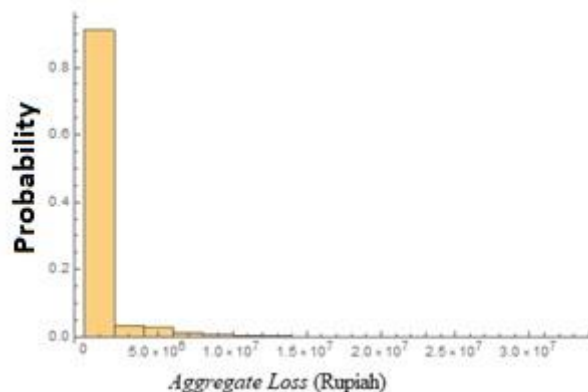
**Model Of Aggregate Loss Distribution of Bpjs Health of Tangerang City In 2017**

The statistical description of aggregate loss is presented in table 7.

**Table 7.** The Statistical Description of Aggregate Loss of BPJS Health of the City of Tangerang in 2017

Mean	Variance	Skewness	Kurtosis
Rp. 508,053.00	$3.49 \times 10^{12}$	5.51803 48.5338	Skewness

The far from zero values can be seen in Figure 5. The distribution has a very long tail to the right. This shows the distribution of aggregate loss mostly at a value of less than Rp508.053.00. The kurtosis values are greater than three, as can be seen in Figure 5, that the data is more pointed to the right. This shows that the distribution of aggregate loss experienced frequent high fluctuations.



**Figure 5.** Histogram of Aggregate Loss Distribution Data



**Determining the Aggregate Loss Claim Distribution Model with the Monte Carlo Simulation Method**

The aggregate loss claim distribution model is a combination of the claim frequency and the claim severity. From the previous discussion, the distribution used for the claim frequency was the Zero One Two Modified Negative Binomial Generalized Exponential (Z12M-NBGE) distribution whereas the claim severity was lognormal distribution.

The value of the claim frequency with the distribution of Zero One Two Modified Negative Binomial Generalized Exponential (Z12M-NBGE) was simulated by randomly generating  $n = 10,000$ . The value of the claim severity with lognormal distribution was simulated by randomly generating the parameter  $\hat{\mu} = 15.11822$  and  $\hat{\sigma} = 0.58312$  as much data as the claim frequency obtained for each  $n$ .

**Determining the Expectations and Variants of Distribution Model of Aggregate Loss with Monte Carlo Simulation Method**

Random variable  $N$  that expressed the claim frequency was a random variable from the Zero One Two Modified Negative Binomial Generalized Exponential (Z12M-NBGE) distribution with expectations and variants that can be obtained using numerical methods. Expectations and values of the variants can be seen in Table 8.

Random variable  $X$  that expressed the claim severity was a random variable from the lognormal distribution with the successive lognormal distribution. The values of expectations and variants based on the lognormal distribution.

$$E(X) = \exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

$$Var(X) = [\exp(\sigma^2) - 1]\exp(2\mu + \sigma^2).$$

Random variable  $S$  states the aggregate loss of the combination of the claim frequency and the claim severity with the expectations and the variants obtained from Monte Carlo simulations. The values of these expectations and variances can be seen in Table 8.

**Table 8.** Expectations and Variants of the Claim frequency ( $N$ ), the Claim severity ( $X$ ), and the Distribution of Aggregate Loss ( $S$ ).

$N$		$X$		$S$	
$E(N)$	$Var(N)$	$E(X)$	$Var(X)$	$E(S)$	$Var(S)$
0.11401	0.14184	Rp. 4.350.182.00	$7.69 \times 10^{12}$	Rp. 508.053.00	$3.49 \times 10^{12}$

**Errors on the Aggregate Loss Distribution Obtained Using Monte Carlo Simulation Method**

An error is a difference between the estimated value and the actual value. There are two types of errors, absolute and relative errors. Absolute error is the absolute value of an error obtained from the difference between the analytical results and the results of the simulation while the relative error is the comparison between the absolute error and the analytical value. Each combined distributions possessed its expectations and variances as the results of simulation and

analytic. The simulation results were obtained using the Monte Carlo simulations. Meanwhile, the analytic results were obtained by counting manually, namely

$$E(S) = E(N)E(X) \text{ and } Var(S) = E(N)Var(X) + Var(N)(E(X))^2.$$

**Table 9.** Expectations, Variances, Absolute Errors, and Relative Errors of the Distribution of Aggregate Loss

	$E(S)$	$Var(S)$
Simulation Results	Rp. 508,053.00	$3.49 \times 10^{12}$
Analytical Results	Rp. 495,964.00	$3.56 \times 10^{12}$
Absolute Error	12089	$0.07 \times 10^{12}$
Relative Error	2.44%	1.97%

Based on Table 9, the absolute and relative errors obtained were relatively small. It can be said that the Monte Carlo simulation method estimated the aggregate loss quite well.

**Determining the VaR Value and ES Value for the Aggregate Loss Claim Distribution with the Monte Carlo Simulation Method**

The final stage was calculating the risk measure namely Value at Risk (VaR) and Shortfall Expectation (ES). The estimation of the amount of risk was done after generating random variables on the distribution of aggregate loss. Furthermore, the estimated results of VaR and ES values on the aggregate loss distribution by the Monte Carlo method are presented in Table 10.

**Table 10.** Estimated VaR values and ES values for the Aggregate Loss distribution

Level of Confidence	VaR ( $\hat{Q}_\alpha$ )	ES ( $\hat{\Omega}_\alpha$ )
95.0%	Rp. 4.279.005,00	Rp. 7.353.810,00
97.5%	Rp. 6.193.182,00	Rp. 9.534.680,00
99.0%	Rp. 9.060.822,00	Rp. 12.785.011,00

It can be seen in Table 10 that the potential claims that can be tolerated at a 95% confidence level are Rp. 4,279,005.00. In other words, the amount of reserved funds to cover the claims of one person per year is Rp. 4, 279,005.00.

ES value of IDR 7,353,810.00 at a confidence level of 95% shows that the claims that can be tolerated with a 95% confidence level are Rp. 7, 353,810.00. The value of ES that is greater than the value of VaR indicates that the maximum amount of reserved funds for the claims of one person per year is Rp. 7, 353,810.00. This explanation applies to other levels of trust. Thus, the Monte Carlo method in its implementation can easily handle many risks with a degree of trust.

**CONCLUSIONS**

It can be concluded that the data distribution for the claim frequency of BPJS in the city of Tangerang in 2017 is the distribution of Z12M-NBGE with some parameters, namely  $\hat{\tau} = 1.06317$ ,  $\hat{\alpha} = 1.48856$  and  $\hat{\beta} = 12.8549$ . The data of the claim severity distribution is the lognormal distribution with  $\hat{\mu}$  parameter equals to 15.11822 and  $\hat{\sigma} 0.58312$ . The combination of Z12M – NBGE distribution and lognormal distribution with the Monte Carlo method which

is  $n = 10,000$  is the aggregate loss distribution. The amount of reserve funds (VaR) that must be prepared at a confidence level of 95%, 97.5%, and 99% are Rp. 4,279,005.00, Rp 6,193,182.00 and Rp. 9,060,822.00 respectively. The maximum reserve funds (ES) that must be prepared at the 95%, 97.5%, and 99% confidence levels are Rp. 7,353,810.00, Rp. 9,534,680.00, and Rp. 12,785,011.00, respectively.

Suggestions that can be developed from the aggregate loss distribution data on BPJS Health in the Tangerang city in 2017 is that it can be done using the Panjer and Fast Fourier Transform methods.

## **AUTHOR CONTRIBUTIONS STATEMENT**

RS as the main drafter in this study. IGPP worked as design of research articles.

## **REFERENCES**

- Achieng, O. M., No, I. (2010). [Actuarial modeling for insurance claims severity in motor comprehensive policy using industrial statistical distributions](#). *International Congress of Actuaries* 712: 5.
- Antonio, K., Frees, E. W., & Valdez, E. A. (2010). [A multilevel analysis of intercompany claim counts](#). *ASTIN Bulletin* 40: 151–177.
- Bodhisuwan, W., & Aryuyuen, S. (2013). [The negative binomial-generalized exponential \(NB-GE\) distribution](#). *Applied Mathematical Sciences* Vol. 7, No.22.
- Hua,L. (2015). [Tail negative dependence and its applications for aggregate loss modeling](#). *Insurance Mathematics and Economics*.
- Klugman SA, Panjer HH, Willmot GE. 2012. *Loss models from data to decisions*. Ed ke-4. New Jersey (US): John Wiley & Sons, Inc.
- Pacakova, V., & Brebera, D. (2016). [Modelling of extreme losses in natural disasters](#). *International Journal of Mathematical Models and Methods in Applied Sciences*,10: 171–178.
- Reiner, W. (2000). *Economic analysis of count data*. Ed ke-3. Germany: Springer Verlag.
- Shevchenko, P. V. (2010). [Calculation of aggregate loss distribution](#). *The Journal of Operational Risk*. 5(2). pp. 3-40.
- Wang, Z. (2011). [One mixed negative binomial distribution with application](#). *Statistical Planning and Inference*. 141: 1153–1160.

