

Value at Risk Estimation for Life Insurance Using Monte Carlo Simulation Based on an ARIMA Model with a Generalized Error Distribution (GED) Assumption

Andreas Timothy^{1*}, Dwi Susanti¹, Kankan Parmikanti¹

¹ Department of Mathematics, Universitas Padjadjaran, Sumedang, Indonesia

Corresponding Author e-mail: andreas22001@mail.unpad.ac.id

Article History:

Received: 24-09-2026

Revised: 26-09-2026

Accepted: 30-09-2026

Keywords: Value at Risk; Autoregressive Integrated Moving Average; Generalized Error Distribution; Monte Carlo simulation; life insurance claims

Abstract: Extreme increases in life insurance claims can threaten an insurer's solvency and liquidity. This study estimates the Value at Risk (VaR) of monthly life insurance claims in Indonesia using an Autoregressive Integrated Moving Average (ARIMA) model with a Generalized Error Distribution (GED) innovation, combined with Monte Carlo simulation. Monthly aggregate claims data published by the Financial Services Authority (OJK), comprising 108 monthly observations and 107 log-return observations, were analyzed. A three-stage procedure involving parameter significance, residual white-noise testing, and minimum Akaike Information Criterion (AIC) selected ARIMA(0,0,1) as the best mean model, with no remaining autocorrelation or conditional heteroscedasticity. Maximum Likelihood Estimation yielded a GED shape parameter of $\hat{\nu} = 0.731368$, indicating heavier tails than the normal distribution. Because the relevant risk for an insurer is a surge in claims, VaR was estimated from the upper tail of the one-month-ahead return distribution using Monte Carlo simulation. At the 99% confidence level, the estimated VaR was 0.91646 in log-return terms, corresponding to a potential claim increase of approximately IDR 21.07 trillion above the reference month's claim value. Kupiec's unconditional coverage test did not reject the model at the 5% significance level ($LR = 2.1508$; $p = 0.1425$). These findings indicate that the ARIMA-GED model combined with Monte Carlo simulation provides a useful statistical framework for quantifying extreme claim-surge risk and supporting risk management and reserve planning in the Indonesian life insurance industry.

How to Cite: Timothy, A., Susanti, D., & Parmikanti, K. (20xx). Value at Risk Estimation for Life Insurance Using Monte Carlo Simulation Based on an ARIMA Model with a Generalized Error Distribution (GED) Assumption. *Ambidextrous Journal of Innovation Efficiency and Technology in Organization*. 5(3). Pp 504-510 <https://doi.org/10.61536/ambidextrous.v5i03.895>



<https://doi.org/10.61536/ambidextrous.v5i03.895>

This is an open-access article under the CC-BY-SA License.



Introduction

Life insurance companies bear the risk that claims and benefits paid in a given period exceed what has been anticipated in premium pricing and technical reserves. A sudden, large increase in claims rather than a decline is the event that threatens an insurer's solvency, liquidity, and capacity to meet its obligations to policyholders. Quantifying the magnitude of such extreme increases is therefore essential for prudent reserve planning and for regulatory risk-based capital assessment.

Value at Risk (VaR) is a widely used risk measure that summarizes the maximum adverse movement of a variable that is not expected to be exceeded at a given confidence level over a specified horizon. While VaR was originally developed for market risk on tradable assets, the same quantile-based logic can be applied to any time series of interest, including insurance claims, provided the tail of interest is defined according to the risk that matters to the institution in this case, the upper (right) tail of the claims distribution, corresponding to claim surges.

Indonesian life insurance claims data published monthly by the Financial Services Authority (OJK) exhibit non-normal characteristics, including heavy tails and asymmetry, driven in part by extreme events such as the onset of the COVID-19 pandemic. Standard VaR approaches that assume normally distributed returns are therefore likely to understate tail risk. This study addresses this gap by combining an ARIMA mean model with a Generalized Error Distribution (GED) innovation which accommodates both leptokurtosis and, through its shape parameter, deviation from normality with Monte Carlo simulation to estimate one-month-ahead VaR for aggregate life insurance claims, and validates the resulting model through Kupiec's backtesting procedure.

The ARIMA model of Box and Jenkins provides a general framework for modelling the conditional mean of a stationary time series through autoregressive (AR) and moving-average (MA) components (Box, Jenkins, Reinsel, & Ljung, 2015). Model order is conventionally selected through a combination of parameter significance testing, residual diagnostic checking, and information criteria such as the Akaike Information Criterion (AIC).

Financial and actuarial time series frequently exhibit heavier tails than the normal distribution. The Generalized Error Distribution (GED), also known as the exponential power distribution, was introduced by Box and Tiao (1962) and later popularized in volatility modelling by Nelson (1991) as an innovation distribution capable of nesting the normal distribution as a special case while allowing heavier or lighter tails through a single shape parameter. Conditional heteroscedasticity in residuals, when present, is commonly tested using the ARCH-LM test of Engle (1982), which determines whether a constant-variance model is adequate or whether a conditionally heteroscedastic (e.g., GARCH-type) extension is required.

Value at Risk was formalized as a standard risk-management tool by Jorion (2007), who describes its estimation via parametric, historical, and Monte Carlo simulation approaches. Because a VaR model is only useful if its coverage is empirically accurate, backtesting is



required; Kupiec (1995) proposed an unconditional-coverage likelihood-ratio test that compares the observed frequency of VaR violations against the frequency implied by the chosen confidence level, providing a formal basis for accepting or rejecting a VaR model.

Research Methods

The data used in this study are aggregate monthly life insurance claims and benefits paid, published by the Indonesian Financial Services Authority (Otoritas Jasa Keuangan, OJK). The original series was published on a year-to-date cumulative basis and was converted into discrete monthly values prior to analysis. The resulting series comprises 108 monthly claim observations, from which 107 monthly log-return observations were computed. All computations were carried out in R, using the *tseries*, *forecast*, *rugarch*, *FinTS*, *moments*, and *lmtest* packages.

The claim series X_t was transformed into a log-return series to achieve stationarity in mean, as follows:

$$R_t = \ln\left(\frac{X_t}{X_{t-1}}\right) \quad (1)$$

Stationarity was formally verified using the Augmented Dickey-Fuller (ADF) test, and departure from normality was assessed with the Jarque-Bera test.

The best ARIMA(p,0,q) mean model, for $p, q \in \{0,1,2\}$, was selected through a three-stage procedure: (1) retaining only candidate models whose parameters were all statistically significant ($\alpha = 0.05$); (2) among those, retaining models whose residuals passed the Ljung-Box white-noise test at lag 12 ($\alpha = 0.05$); and (3) selecting the model with the smallest AIC among the models that passed both prior stages.

An ARCH-LM test was then applied to the residuals of the selected ARIMA model to determine whether conditional heteroscedasticity was present. If present, an ARIMA-GARCH(1,1) model with GED innovations would be estimated; otherwise, an ARIMA model with constant variance and a GED assumption would be estimated directly. Parameters were estimated by Maximum Likelihood Estimation (MLE).

For the selected MA(1) mean model, the one-month-ahead return is given by:

$$R_{t+1} = \mu + \theta \varepsilon_t + z_{t+1} \quad (2)$$

where ε_t is the last historical residual and $z_{t+1} \sim \text{GED}(0, \sigma, \nu)$ is an independently generated innovation. Because the risk of interest to an insurer is a claim surge, VaR was defined on the upper (right) tail of the simulated return distribution: $\text{VaR}_{1-\alpha}$ is the $(1 - \alpha)$ -quantile of the simulated R_{t+1} values. Monte Carlo simulation was performed with an initial 100,000 replications, increased in batches of 5,000 whenever the batch-based standard error (SE) of the quantile estimate remained at or above 0.01, until convergence ($\text{SE} < 0.01$) or a maximum of 300,000 replications was reached.

The resulting VaR estimate was converted to a nominal rupiah value using:

$$P_{t+1} = P_t \cdot e^{\text{VaR}} \quad (3)$$

where P_t is the claim value in the reference month. Model validity was assessed through Kupiec's (1995) unconditional-coverage backtest, in which a violation is defined as a realized return exceeding the time-varying upper VaR bound $\mu_t + \sigma \cdot q_{1-\alpha}$, where $q_{1-\alpha}$ is the $(1 - \alpha)$ -quantile of the standard GED; the resulting likelihood-ratio statistic was compared against a chi-square distribution with one degree of freedom.

Result and Discussion

Table 1 summarizes the descriptive statistics of the monthly log-return series. The negative skewness and kurtosis far above 3 indicate an asymmetric, heavy-tailed distribution, driven by two extreme observations coinciding with the onset of the COVID-19 pandemic in Indonesia (March 2020) and the sharp decline following the Delta-variant wave (September 2021). These two observations were retained in the analysis, as they reflect genuine national-scale events rather than data errors, and removing them would understate the tail risk that VaR is intended to capture.

Table 1. Descriptive statistics of the monthly claim log-return series

Statistic	Value
Number of observations (n)	107
Mean	0.004935
Standard deviation	0.505423
Skewness	-0.911574
Kurtosis	12.179433

The ADF test rejected the null hypothesis of a unit root, confirming that the log-return series is stationary, while the Jarque-Bera test strongly rejected normality, motivating the use of a non-normal innovation distribution (Table 2).

Table 2. Stationarity and normality test results

Test	Statistic	p-value	Conclusion
Augmented Dickey-Fuller	-8.0356	0.0100	Stationary
Jarque-Bera	390.49	< 0.0001	Not normally distributed

Among eight candidate ARIMA(p,0,q) models with $p, q \in \{0,1,2\}$, only ARIMA(0,0,1), ARIMA(1,0,0), and ARIMA(2,0,0) had fully significant parameters. Of these, only ARIMA(0,0,1) produced white-noise residuals (Ljung-Box $p = 0.7422$); the other two candidates were rejected at this stage ($p = 0.0481$ and $p = 0.0242$, respectively). ARIMA(0,0,1) was therefore selected as the best mean model, with the smallest AIC among all candidates that passed the significance and white-noise stages (Table 3).

Table 3. Selected ARIMA model and information criteria

Model	AIC	BIC	Status
ARIMA(0,0,1)	88.184	96.202	Selected

An ARCH-LM test on the ARIMA(0,0,1) residuals produced a chi-square statistic of 0.20285 ($p = 0.6524$), indicating no remaining conditional heteroscedasticity. A constant-variance ARIMA(0,0,1) model with a GED assumption was therefore estimated directly via Maximum Likelihood Estimation, without a GARCH extension. All four parameters were statistically significant (Table 4).

Table 4. Maximum Likelihood parameter estimates of the ARIMA(0,0,1) model with a GED assumption

Parameter	Estimate	p-value
$\hat{\mu}$ (mean)	0.007038	< 0.0001
$\hat{\theta}$ (MA(1))	-0.926742	< 0.0001
$\hat{\sigma}$ (scale)	0.323344	< 0.0001
$\hat{\nu}$ (shape)	0.731368	< 0.0001

The shape parameter $\hat{\nu} = 0.731368$ is well below 2, confirming that the fitted GED has substantially heavier tails than the normal distribution; this is consistent with a quantile-quantile comparison, in which the sum of squared deviations from the 45° line was smaller for the GED (25.2194) than for the normal distribution (38.2375), indicating a better overall fit. Diagnostic checks on the standardized residuals of the fitted model Ljung-Box tests on both the residuals and their squares, and ARCH-LM tests at multiple lags produced p-values well above 0.05 at every lag tested, confirming that the residuals are white noise with no remaining conditional heteroscedasticity, and that the ARIMA(0,0,1) model with a GED assumption is adequate for VaR estimation.

One-month-ahead ($k = 1$) VaR was estimated via Monte Carlo simulation on the upper tail of the return distribution, reflecting the risk of a claim surge. The simulation converged after 110,000 replications, at which point the batch-based standard error fell below the target precision of 0.01. Table 5 reports the simulated VaR at the 99% confidence level for horizon $k = 1$.

Table 5. VaR estimate at the 99% confidence level, horizon $k = 1$

Method	VaR (log-return)	SE	Claim increase (IDR trillion)
Monte Carlo (110,000 replications)	0.91646	0.00991	21.07

Applying the Monte Carlo VaR to the reference claim value of IDR 14.04 trillion (December 2024) implies a projected total claim value of approximately IDR 35.11 trillion under the 99% VaR scenario — an increase of IDR 21.07 trillion above the reference value.

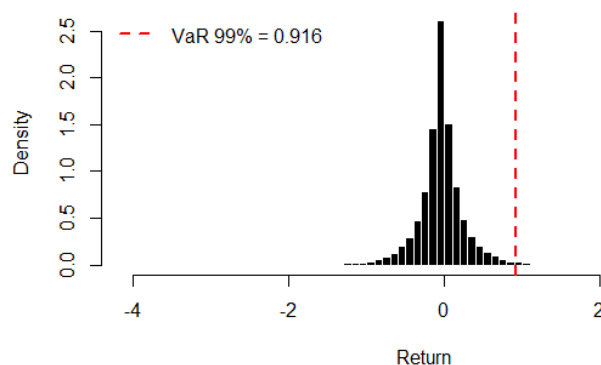

Figure 1. Distribution of one-month-ahead simulated returns and the 99% VaR threshold

Figure 1 shows the histogram of simulated one-month-ahead returns together with the 99% VaR threshold. The distribution is sharply peaked near zero with a right tail extending toward large positive values, consistent with the heavy right tail implied by the estimated GED shape parameter.

Model validity was assessed through Kupiec's (1995) unconditional-coverage backtest

Ambidextrous : Journal of Innovation Efficiency and Technology in Organization

<https://doi.org/10.61536/ambidextrous.v5i03.895>

This open-access article is distributed under a Creative Commons Attribution (CC-BY-SA) 4.0 license



Value at Risk Estimation for Life Insurance (Timothy, Susanti, & Parmikanti) Page | 509
at the 99% confidence level, comparing realized returns against the time-varying upper VaR bound over the full 107-observation sample (Table 6).

Table 6. Kupiec backtesting results at the 99% confidence level, horizon $k = 1$

n	LR statistic	p-value	Decision
107	2.1508	0.1425	Valid

Zero violations were observed against 1.07 expected under the null hypothesis, yielding a likelihood-ratio statistic of 2.1508 ($p = 0.1425$), which does not reject the null hypothesis at $\alpha = 0.05$. The model is therefore considered valid at the 99% confidence level. It should be noted that, with only 107 observations and an expected violation count close to one, the backtest has limited statistical power; the "valid" decision reflects an absence of evidence against the model rather than strong positive evidence of its accuracy, and should be interpreted accordingly.

Conclusion and Recommendation

This study estimated the Value at Risk of monthly life insurance claims in Indonesia using an ARIMA(0,0,1) mean model with Generalized Error Distribution innovations, combined with Monte Carlo simulation. A three-stage selection procedure identified ARIMA(0,0,1) as the best mean model, with white-noise residuals and no remaining ARCH effect, justifying a constant-variance specification. Maximum Likelihood Estimation produced a GED shape parameter well below 2, confirming heavier tails than the normal distribution.

Because the relevant risk to a life insurer is a surge rather than a decline in claims, VaR was estimated on the upper tail of the one-month-ahead return distribution. At the 99% confidence level, the estimated VaR corresponds to a potential claim increase of approximately IDR 21.07 trillion above the reference month's value, a result confirmed valid through Kupiec's backtesting procedure. These findings demonstrate that the ARIMA model with a GED assumption, combined with Monte Carlo simulation, provides a statistically defensible tool for quantifying extreme claim-surge risk, with practical relevance for reserve adequacy and risk-based capital assessment in the Indonesian life insurance industry. Future research is recommended to extend the model to multi-period horizons with out-of-sample backtesting, and to consider skewed innovation distributions such as the Skewed Generalized Error Distribution to better accommodate the asymmetry observed in the data.

Acknowledge

The authors thank the Department of Mathematics, Universitas Padjadjaran, for its support during this research, and the Indonesian Financial Services Authority (OJK) for the publicly available insurance statistics used in this study.

References

- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.
- Box, G. E. P., & Tiao, G. C. (1962). A further look at robustness via Bayes's theorem. *Biometrika*, 49(3/4), 419–432.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the



- variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007.
- Jorion, P. (2007). *Value at risk: The new benchmark for managing financial risk* (3rd ed.). McGraw-Hill.
- Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, 3(2), 73–84.
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2), 347–370.
- Otoritas Jasa Keuangan. (2024). *Statistik perasuransian Indonesia*. Otoritas Jasa Keuangan.

