Desimal: Jurnal Matematika Vol 4 No 3 (2021) 339-348

Contents lists available at DJM

DESIMAL: JURNAL MATEMATIKA

p-ISSN: 2613-9073 (print), e-ISSN: 2613-9081 (online), DOI 10.24042/djm

http://ejournal.radenintan.ac.id/index.php/desimal/index

The determination of the aggregate loss distribution through the numerical inverse of the characteristic function using the trapezoidal quadrature rule

Aceng Komarudin Mutaqin, Khalimatus Sa'diah*

Bandung Islamic University, Indonesia

ARTICLE INFO

Article History

Received : 07-07-2021
Revised : 27-10-2021
Accepted : 09-11-2021
Published : 30-11-2021

Keywords:

Characteristic Function; Numerical Inversion; Aggregate Loss; Trapezoidal Quadrature.

*Correspondence: E-mail: sadiahkhalimatus4@gmail.com

Doi: 10.24042/djm.v4i3.9434

ABSTRACT

Aggregate loss is the total loss suffered by an insured in a certain period. The aggregate loss depends on the claim frequency and the amount of the claim each time the insured makes a claim. The distribution of aggregate losses must be known to calculate motor vehicle insurance premiums. In general, there are two methods that can be used in determining the distribution of aggregate losses, namely exact and numerical. When an exact solution is difficult to find, numerical methods such as Monte Carlo, Panjer Recursion, and Fast Fourier Transform can be used. This research will discuss the determination of the distribution of aggregate losses through the numerical inverse of the characteristic function using the trapezoidal quadrature rule, on the data of motor vehicle insurance category 7 in Indonesia. The estimated cumulative distribution function for the largest aggregate loss is 0.999993. When x=0, it means that if someone does not file a claim, the estimated value of the cumulative distribution function is 0.9293. This value is close to the percentage of the number of insured, which is 0.9241.

http://ejournal.radenintan.ac.id/index.php/desimal/index

INTRODUCTION

According to Salim (1985), insurance is generally divided into 3 categories: (i) Life insurance; (ii) Social insurance; and (iii) Loss insurance. Life insurance is economic protection against losses that may occur due to a possible event such as death, illness, or accident. Social insurance is insurance that is intended to provide basic insurance for the community and does not aim to obtain commercial benefits. Loss insurance is a transfer of risk for losses, loss of benefits, and legal liability to third parties arising from uncertain events.

One of the types of loss insurance in Indonesia that is in great demand is motor vehicle insurance, because it provides coverage for loss, damage, or loss that reduces the value of motorized vehicle objects. Products for motor vehicle insurance are Total Loss Only (TLO) and Comprehensive (Kementrian Keuangan

Republik Indonesia Badan Pengawasan Pasar Modal Dan Lembaga Keuangan, 2011). For TLO insurance products, the types of claims submitted are a total loss, while for comprehensive insurance products, there are two types of claims, namely total loss and partial loss. The type of total loss claim is that the insured can submit a claim only once in his insurance period to the insurer, while the type of partial loss claim is that the claim submitted to the insurer by the insured may be more than once during the insurance period. The total loss suffered by an insured in a certain period is often referred to as aggregate loss. Thus, the aggregate loss depends on the frequency of claims and the size of the claim each time the insured makes a claim.

In calculating motor vehicle insurance premiums, one of the things that must be known is the distribution of aggregate losses. In general, there are two methods that can be used in determining the distribution of aggregate losses, namely exact and numerical. When an exact solution is difficult to find, numerical methods such as Monte Carlo, Panjer Recursion (Utami et al., 2018), and Fast Fourier Transform can be used (Kartini et al., 2018). Witkovsky et al. (2017)discuss methods and algorithms for calculating the probability density function and the cumulative distribution function of the aggregate loss through the numerical inverse of its characteristic function using the trapezoidal quadrature rule, which is quite precise for most practical situations and can be implemented well in MATLAB software, via characteristic functions toolbox or CF Toolbox.

METHOD

To further explain the discussion in this research, materials and methods or steps are needed to determine the distribution of aggregate losses through the numerical inverse of its characteristic function using the trapezoidal quadrature rule on motor vehicle insurance data in Indonesia.

In this research, the data used is secondary data recorded from the results of the insurance company PT. XYZ. The data contains vehicle insurance insureds who make claims to the insurance company that shelters them. The data used for the purposes of the application is data on the frequency and magnitude of claims made by the insured motor vehicle insurance category 7 (bus vehicles for all sum insured) who became the insured in 2011, while the claim used is a partial loss, which is the type of claim that can be submitted more than once by the insured.

Data on the frequency of partial loss claims from the insured motor vehicle insurance category 7 in Indonesia are presented in Table 1.

Table 1. Frequency of Motor Vehicle Insurance Claims Category 7

Claim Frequency	Number of Insured
0	1,911
1	115
2	21
3	15
4	3
5	3

Table 1 contains the frequency of claims of motor vehicle insurance category 7 policyholders at PT. XYZ. Based on Table 1, there are 1,911 policyholders who do not make a claim during the one-year insurance period. There are 115 policyholders who make a one-time claim during their one-year insurance period. There are 21 policyholders who make claims twice during one year of insurance and so on for the others. Furthermore, Table 2 contains the amount of the claim of the insured company of PT. XYZ which approved by insurance companies. For the number one insured, the amount of the approved claim is Rp. 3,100,000 the amount of approved claim for the second insured is Rp. 6,575,000 and so on for the others.

Data on the amount of claim insured for motor vehicle insurance category 7 in Indonesia are presented in Table 2.

Table 2. Amount of Motor Vehicle Insurance Claim Category 7

No.	Approved Claims (Rupiah)
1	3,100,000
2	6,575,000
3	11,756,000
4	2,544,300
5	2,760,000
6	6,533,000
⋮	⋮
226	1,671,725
227	1,021,250
228	4,250,000
229	1,975,000

The empirical characteristic function is a mixture with equal weights of the characteristic function of the Dirac random variable concentrated on the observed values x_j in X_j, that is, the mixture of the characteristic functions given by $\mathrm{cf}_{x_j}(t) = e^{\mathrm{i}tx_j}$,

$$\mathrm{cf}_{\hat{F}_X}(t) = \frac{1}{n}\sum_{j=1}^{n} e^{\mathrm{i}tx_j}. \qquad (2.10)$$

$\hat{F}_N$ expresses the empirical cumulative distribution function of the claim frequency $n_1, \dots, n_J$, in each of J historical years, with the empirical characteristic function given by

$$\mathrm{cf}_{\hat{F}_N}(t) = \frac{1}{J}\sum_{j=1}^{J} e^{\mathrm{i}tn_j}. \qquad (2.11)$$

Next, suppose $\hat{F}_X$ represents the empirical cumulative distribution function based on K observed values from claim $x_1, \dots, x_K$, with the empirical characteristic function being

$$\mathrm{cf}_{\hat{F}_X}(t) = \frac{1}{K}\sum_{k=1}^{K} e^{\mathrm{i}tx_k}. \qquad (2.12)$$

Then, the analogy of $\mathrm{cf}_S(t) = \mathrm{cf}_N(-\mathrm{i}\log(\mathrm{cf}_X(t)))$, an empirical characteristic function, let's say $\mathrm{cf}_{\hat{F}_S}(t)$, of the collective risk distribution S is

$$\mathrm{cf}_{\hat{F}_S}(t) = \mathrm{cf}_{\hat{F}_N}\left(-\mathrm{i}\log\left(cf_{\hat{F}_X}(t)\right)\right) = \frac{1}{J}\sum_{j=1}^{J}\left(\frac{1}{K}e^{\mathrm{i}tx_k}\right)^{n_j}. \qquad (2.13)$$

Gil-Pelaez (1951) derived the inverse formula of characteristic functions that can be integrated along $(-\infty, \infty)$, suitable for numerically evaluating probability density functions and/or cumulative distribution functions, which require the function to be integrated to have real values only. In particular, the probability density function of a continuous distribution, with the characteristic function $\mathrm{cf}_L(t)$, is given by

$$\mathrm{pdf}_L(\ell) = \frac{1}{\pi}\int_0^{\infty} \Re\left(e^{-\mathrm{i}t\ell}\mathrm{cf}_L(t)\right)dt, \qquad (2.14)$$

then if ℓ is the continuity point of the cumulative distribution function L, defined by $\mathrm{cdf}_L(\ell) = \Pr(L \le \ell)$, then the cumulative distribution function is given by

$$\mathrm{cdf}_L(\ell) = \frac{1}{2} - \frac{1}{\pi}\int_0^{\infty} \Im\left(\frac{e^{-\mathrm{i}t\ell}\mathrm{cf}_L(t)}{t}\right)dt, \qquad (2.15)$$

where $\Re(f(t))$ and $\Im(f(t))$ represent the real and imaginary parts of the complex function $f(t)$ respectively.

Often the integrals in Equation (2.14) and (2.15) can be approximated efficiently by the trapezoidal quadrature numerical method, namely

$$\mathrm{pdf}_L(\ell) \approx \frac{\delta}{\pi}\sum_{j=0}^{N} w_j \Re\left(e^{-\mathrm{i}t_j\ell}\mathrm{cf}_L(t_j)\right), \qquad (2.16)$$

and/or

$$\mathrm{cdf}_L(\ell) \approx \frac{1}{2} - \frac{\delta}{\pi}\sum_{j=0}^{N} w_j \Im\left(\frac{e^{-\mathrm{i}t_j\ell}\mathrm{cf}_L(t_j)}{t_j}\right), \qquad (2.17)$$

where N is a very large integer, let's say $N = 2^{10}$, w_j is the weight for the trapezoidal quadrature (namely $w_0 = w_N = \frac{1}{2}$, and $w_j = 1$ for $j = 1, 2, \ldots, N - 1$), and $t_j = j\delta$ for $j = 0, \ldots, N$ are nodes equidistant (with a mutually beneficial distance δ) from the interval [0, T] for a sufficiently large T (let's say T is such that the integral function $\Re\left(e^{-\mathrm{i}t_j\ell}\mathrm{cf}_L(t_j)\right)$ and/or $\Im\left(\frac{e^{-\mathrm{i}t_j\ell}\mathrm{cf}_L(t_j)}{t_j}\right)$ is small enough for all $t > T$).

In this research, the determination of the distribution of aggregate losses will be discussed through the numerical inverse of the characteristic function using the trapezoidal quadrature rule. This rule is a rule that can be used for non-parametric data types, so it is not required to look for parameter values as in parametric data types in other research methods. In this research, researchers used MATLAB software. The steps taken are as follows.

1. Determine the estimated empirical characteristic function for the frequency data of motor vehicle insurance claims category 7 using Equation (2.11).
2. Determine the estimated empirical characteristic function for the data on the amount of motor vehicle insurance claims category 7 using Equation (2.12).
3. Determine the estimated empirical characteristic function for the aggregate losses of motor vehicle insurance category 7 using Equation (2.13).
4. Determine the estimated density function for the aggregate losses of motor vehicle insurance category 7 using Equation (2.16).
5. Determine the estimated cumulative distribution function for the aggregate losses of motor vehicle insurance category 7 using Equation (2.17).

To make it easier to understand, the authors present a flow chart to determine the distribution of aggregate losses through the numerical inverse of the characteristic function using the trapezoidal quadrature rule in Figure 1.

RESULTS AND DISCUSSION

The estimation of the empirical characteristic function of the aggregate loss of motor vehicle insurance category 7 depends on the estimation of the empirical characteristic function for the frequency data of motor vehicle insurance claim category 7 and the estimation of the empirical characteristic function for the aggregate loss of motor vehicle insurance category 7. The estimation of the empirical characteristic function for the frequency data of motor vehicle insurance claim category 7 can be calculated using Equation (2.11) based on the data on the claim frequency for motor vehicle insurance for category 7 in Table 1. With the help of MATLAB software, the estimation of the empirical characteristic function is obtained for the frequency data of motor vehicles insurance claims category 7.

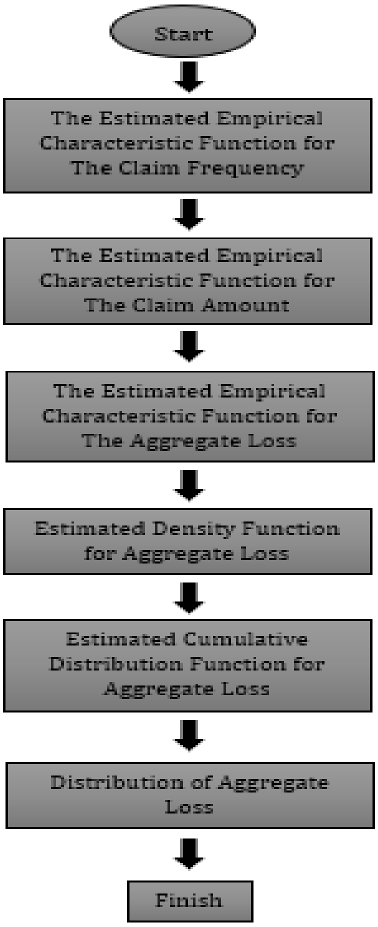

Figure 1. Flowchart of Determining the Distribution of Aggregate Loss

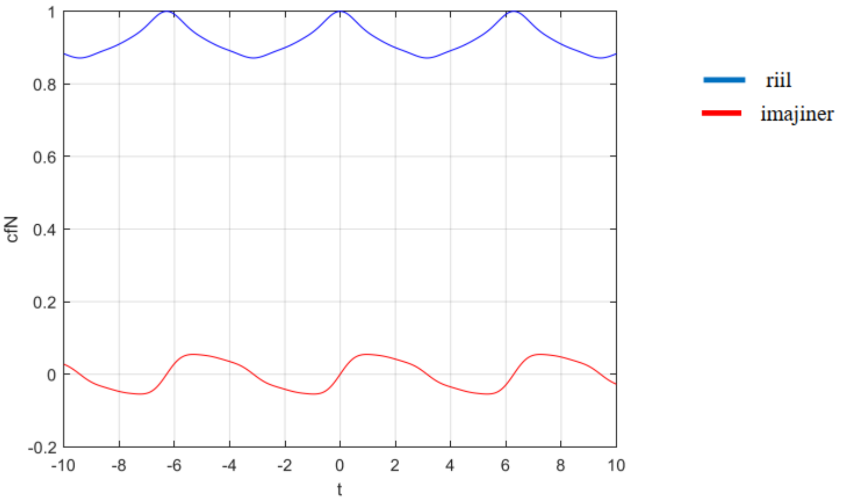

Figure 2. The Real (blue line) and Imaginary (red line) Parts of The Estimated Empirical Characteristic Function for The Claim Frequency

The estimation of the empirical characteristic function for the claim frequency is in the form of a complex number (containing the real and the imaginary part). The real and the imaginary part are separate. The real part is in the interval 0.8 to 1. While the imaginary part is in the interval -0.2 to 0.2. In the real and the imaginary part, the estimation of the empirical characteristic function for the claim frequency has a different pattern.

The estimation of the empirical characteristic function for the data of the amount of the motor vehicle insurance claim category 7 can be calculated using Equation (2.12) based on the data of the amount of motor vehicle insurance claim category 7 in Table 2.

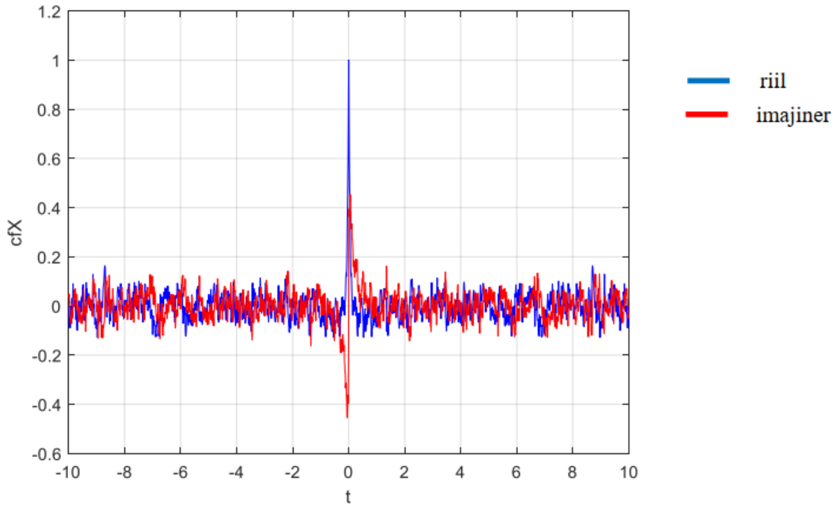

Figure 3. The Real (blue line) and Imaginary (red line) Parts of The Estimated Empirical Characteristic Function for The Claim Amount

The estimation of the empirical characteristic function for the amount of its claim value is in the form of a complex number (containing the real and the imaginary part). The real and the imaginary part of the estimation value of the empirical characteristic function for the amount of the claim have almost the same pattern, except at t=0.

The estimation of the empirical characteristic function for the aggregate loss of motor vehicle insurance category 7 can be calculated using Equation (2.13) based on data on the frequency of claims and the data of the amount of motor vehicle insurance claim category 7 in Table 1 and Table 2. With the help of MATLAB software, the estimation of empirical characteristic function for the aggregate loss of motor vehicle insurance category 7 is obtained.

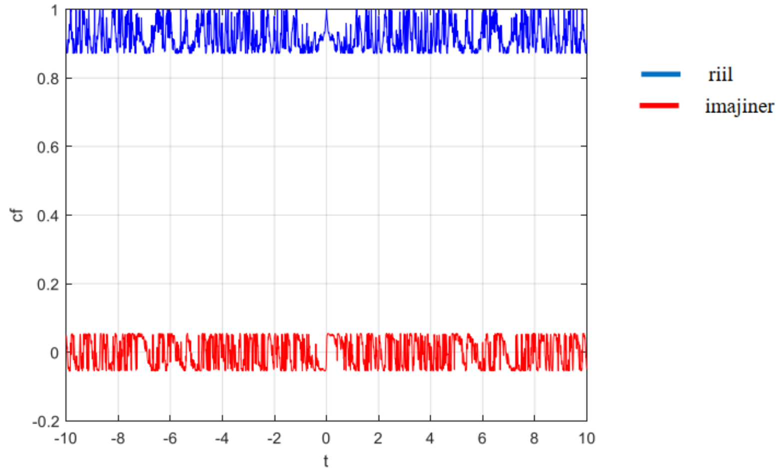

Figure 4. The Real (blue line) and Imaginary (red line) Parts of The Estimated Empirical Characteristic Function for The Aggregate Loss

The estimation of the empirical characteristic function for aggregate losses is in the form of complex numbers (containing real and imaginary parts). The real and the imaginary part are separate. The real part is in the interval 0.9 to 1. While the imaginary part is in the interval -0.05 to 0.05. In the real and the imaginary part, the estimation of the empirical characteristic function for the claim frequency has a different pattern.

The estimated density function for the aggregate loss of motor vehicle insurance category 7 can be calculated using Equation (2.16) based on the data on the claim frequency and the data of the amount of motor vehicle insurance claim category 7 in Table 1 and Table 2. With the help of MATLAB software, the estimated density function for aggregate losses of motor vehicle insurance category 7 is obtained.

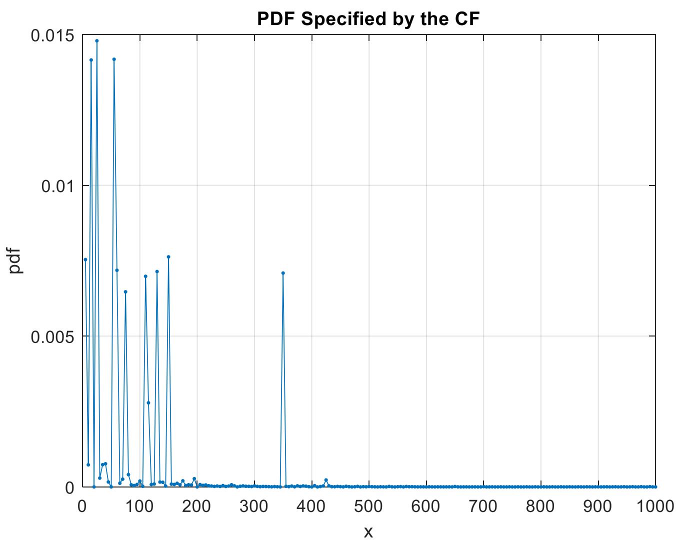

Figure 5. Estimated Density Function for Aggregate Loss

Based on Figure 5, it can be seen that the estimated value of the density function for aggregate losses is between 0 and 0.015, and has many peaks. It means that the estimated value of the density function for aggregate losses describes aggregate loss data that goes up and down.

The estimation of the cumulative distribution function for the aggregate loss of motor vehicle insurance category 7 can be calculated using Equation (2.17) based on the claim frequency data and the data of the amount of motor vehicle insurance claim category 7 in Table 1 and Table 2. With the help of MATLAB software, the estimation of the cumulative distribution function for aggregate losses of motor vehicle insurance category 7 is obtained.

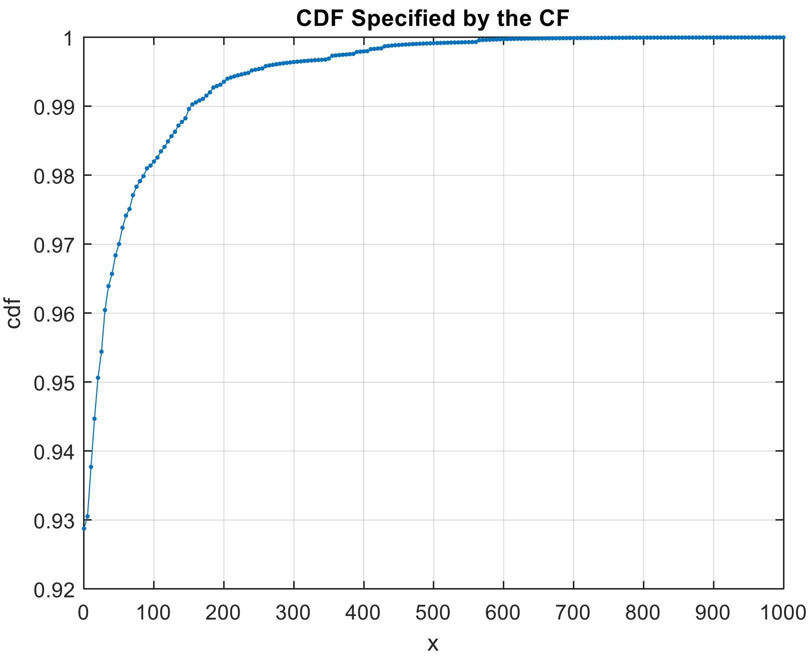

Figure 6. Estimated Cumulative Distribution Function for Aggregate Loss

Based on Figure 6, it can be seen that the estimated value of the cumulative distribution function for the largest aggregate loss is 0.999993. When *x*=0, it means that if someone does not file a claim, the estimated value of the cumulative distribution function is 0.9293. This value is close to the percentage of the number of the insured, which is 1.911/2068=0.9241. It means that the estimated value is close to the real value, that there are as many as 1,911 insured who did not file a claim.

CONCLUSIONS AND SUGGESTIONS

In this research, it has been proposed to determine the distribution of the insured's aggregate loss through the numerical inverse of the characteristic function using the trapezoidal quadrature rule. Based on the results of the application of the numerical inverse method of the characteristic function using the trapezoidal quadrature rule on aggregate loss data for insured motor vehicle insurance category 7 in Indonesia, it can be concluded that the estimated cumulative distribution function for the largest aggregate loss is 0.999993. When *x*=0, it means that if someone does not file a claim, the estimated value of the cumulative distribution function is 0.9293. This value is close to the percentage of the number of insured, which is 0.9241.

Insurance companies are advised to consider the numerical inverse method of the characteristic function using this trapezoidal quadrature rule in determining the distribution of aggregate losses. For other researchers, it is recommended to use other methods besides the numerical inverse method of the characteristic function using the trapezoidal quadrature rule, namely a numerical method such as Monte Carlo to determine the distribution of aggregate losses.

REFERENCES

Gil-Pelaez, J. (1951). Note on the inversion theorem. *Biometrika*, *38*(3–4). https://doi.org/10.1093/biomet/38.3-4.481

Kartini, N., Sunendiari, S., & Mutaqin, A. K. (2018). Penentuan distribusi kerugian agregat tertanggung asuransi kendaraan bermotor di indonesia menggunakan metode fast fourier transform. *Prosiding Statistika Unisba*.

Kementrian Keuangan Republik Indonesia Badan Pengawasan Pasar Modal Dan Lembaga Keuangan. (2011). *Peraturan ketua badan pengawasan pasar modal dan lembaga keuangan nomor: PER-04/BL/2011*.

Salim, A. A. (1985). *Dasar-dasar asuransi*. Tarsito.

Utami, D. A., Mutaqin, A. K., & Wachidah, L. (2018). *Penentuan distribusi kerugian agregat tertanggung asuransi kendaraan bermotor di indonesia menggunakan metode rekursif panjer*. Universitas Islam Bandung.

Witkovsky, V., Wimmer, G., & Duby, T. (2017). *Computing the aggregate loss distribution based on numerical inversion of the compound empirical characteristic function of frequency and severity*. https://arxiv.org/pdf/1701.08299.pdf