wjert, 2020, Vol. 6, Issue 5, 322-335. *Review Article* ISSN 2454-695X

World Journal of Engineering Research and Technology

WJERT

www.wjert.org SJIF Impact Factor: 5.924

PARAMETER ESTIMATION FROM NEGATIVE-LINDLEY BINOMIAL DISTRIBUTION WITH MAXIMUM POSSIBLE METHODS: APPLICATION IN MOTOR VEHICLE INSURANCE

Yudhistira Naufal Sulaeman[1], Berlian Setiawaty[1], I. Gusti Putu Purnaba[1] and Tita Rosita[2]*

[1]Department of Mathematics, Bogor Agricultural University, Bogor, Indonesia.

[2]Department of Mathematics and Statistics, AKA Bogor Polytechnic, Bogor, Indonesia.

Article Received on 30/07/2020 **Article Revised on 20/08/2020** **Article Accepted on 10/09/2020**

***Corresponding Author**
Tita Rosita
Department of Mathematics and Statistics, AKA Bogor Polytechnic, Bogor, Indonesia.

ABSTRACT

Frequency modeling of insurance claims plays an important role in an insurance company. This paper explains about negative binomial-Lindley distribution. This distribution has the greatest probability value at zero frequency. Negative binomial-Lindley distribution is usually used to model overdispersed data with large zero frequency probability which is often found in motor vehicle insurance data. For its application, the distribution parameters are estimated using the maximum likelihood method. The numerical calculation has done by using *Wolfram Mathematica 11.2* and *R 3.5.1* software. The result shows that negative binomial-Lindley distribution is suitable for estimating the frequency of motor vehicle insurance claims with a very good degree of accuracy.

KEYWORDS: Frequency of claims, maximum likelihood method, insurance, *Wolfram Mathematica*, negative binomial-Lindley distribution

INTRODUCTION

Over time, the community is wiser and more concerned to face the risks that occur in the future by entrusting it to insurance companies. This is a part of prevention and a sense of security to face the future. The insurance sector has an important role in a country's economy, especially its role as an intermediary and financial service provider by identifying community risks. Thus, the demand for insurance grows rapidly, resulting in insurance companies must be more intensive in financial planning and can provide reserve claims based on the frequency of claims that have occurred in the past. Modeling the frequency of claims is one of the important things in actuarial theory and practice, especially in motor vehicle insurance. Motor vehicle insurance is one type of insurance that is in high demand by consumers. The frequency of claims is modeled by discrete distributions such as Poisson, geometric and negative binomials. However, because the frequency of insurance claims with the number of zero claims frequencies is usually very large, a mixed distribution of Negative-Lindley Binomials (NB-L) is used. The Negative-Lindley Binomial mixture (NB-L) is also used to overcome data overdispersion (the range of data that exceeds the expected value).

In this study, the distribution of Negative Binomial and Lindley distribution is explained. This mixed distribution is considered as an alternative to modeling the number of insurance claims that have thick tails and are large at zero (Zamani and Ismail 2010). This means that the frequency of claims for each period cannot be predicted precisely but is assumed to always be of great value when no claims occur. This condition is very suitable with the distribution of the Negative-Lindley Binomial mixture (NB-L) which has thick tail characteristics and has a large value at zero.

To be able to apply this distribution in the insurance field, you must look for the parameters that will be estimated. There are several methods that can be used, namely the moment method and the maximum likelihood method. Estimation of parameters using the Maximum Likelihood method cannot always be done directly so that the solution can use the Newton-Raphson method. In the next process, the parameter estimator values from the Negative-Lindley Binomial distribution (NB-L) are searched by the Maximum Probability method. This study refers to an article entitled Negative Binomial-Lindley Distribution and Its Application written by Zamani and Ismail (2010).

MATERIALS AND METHODS

Expectation Value

Suppose X is a discrete random variable with a probability mass function $p_X(x)$ then the expected value of X, denoted by $E(X)$ is

$$E(X) = \sum_{\forall x} x p_X(x),$$

provided that number is absolute convergent. If X is a continuous random variable with a probability density function $f_X(x)$, then the expected value of X is

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx,$$

Provided that the integral is absolutely convergent (Hogg et al. 2014).

Moment, Function Generating Moment and Function Generating Opportunities

The k moment of the random variable X is denoted by μ'_k is $\mu'_k = E(X^k)$

The moment generating function of the random variable X is defined as follows

$$M_X(t) = E(e^{tX}) = \begin{cases} \sum_x e^{tx} p_X(x); & X \text{ diskret} \\ \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx; & X \text{ kontinu} \end{cases}$$

(Bouk 2016).

The opportunity generating function of the random variable X is defined as follows

$$G_X(z) = E(z^X) = \begin{cases} \sum_x z^x p_X(x); & X \text{ diskret} \\ \int_{-\infty}^{\infty} z^x f_X(x)\, dx; & X \text{ kontinu} \end{cases}$$

(Grimmett dan Stirzaker 1992).

Factorial Moment

For natural numbers k, the k-factorial moment of the distribution denoted $\mu_{[k]}$ is $\mu_{[k]}(X) = E[X(X-1)(X-2)\ldots(X-k+1)]\,,\ \ k = 1, 2, 3, \ldots$ (Breton 2014).

Negative Binomial Distribution

Suppose X is a discrete random variable that spreads negative binomials with parameters r, p and denoted by $X \sim NB(r, p)$ has the opportunity mass function as follows

$$p_X(x) = P(X = x) = \binom{x + r - 1}{x} p^r (1 - p)^x,$$

with $0 \leq p \leq 1$, $r > 0$, and $x = 0,1,2,3,\ldots$ (Ghahramani 2005).

The expected value, the 2nd moment, and the kth factorial moment of X are:

$$E(X) = \frac{r(1-p)}{p}$$

$$E(X^2) = \frac{r(1-p)[1+r(1-p)]}{p^2}$$

$$\mu_{[k]}(X) = \frac{\Gamma(r+k)}{\Gamma(r)}\frac{(1-p)^k}{p^k}, \quad k = 1,2,3,\dots.$$

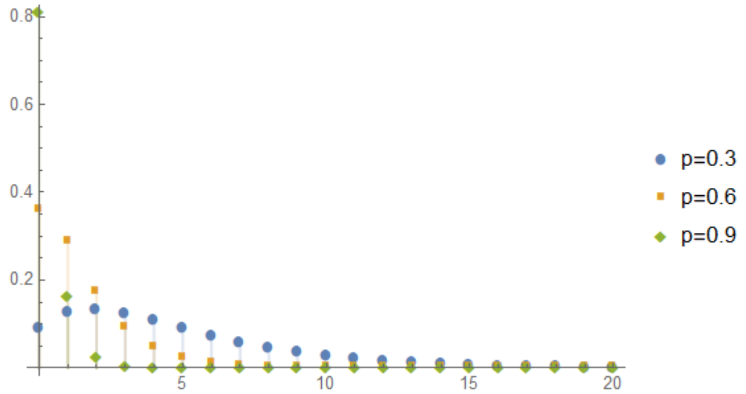

Figure 1: Illustration of opportunity mass function $X \sim NB(r,p)$ with *p* = 0.3,0.6,0.9 *r* = 2 and *x* = 20.

Lindley distribution

Suppose that Λ is a continuous random variable that spreads Lindley with the parameter θ and denoted by $\Lambda \sim \text{Lin}(\theta)$has the opportunity density function as follows:

$f_\Lambda(\lambda) = \frac{\theta^2}{\theta+1}(1+\lambda)e^{-\theta\lambda}$ dengan $\lambda > 0$ dan $\theta > 0$

The moment generating function is as follows: $M_\Lambda(z) = \frac{\theta^2}{\theta+1}\frac{\theta-z+1}{(\theta-z)^2}$

(Zamani dan Ismail 2010).

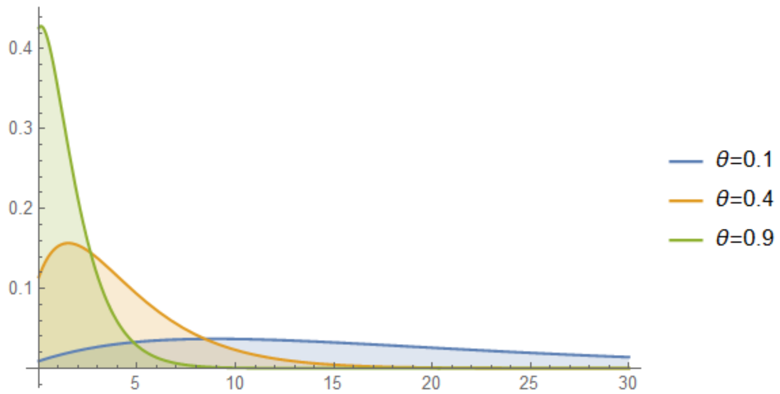

Figure 2. Illustration of probability density function $\Lambda \sim \text{Lin}(\theta)$ with θ = 0.1,0.4, and 0.9 with $0 \leq \lambda \leq 30$.

General Mixed Frequency Distribution

Let $G(x|\lambda)$ be an opportunity generating function with a random variable $X|\Lambda = \lambda$ where $U(\lambda) = P(\Lambda \leq \lambda)$ is a distribution function of Λ. Let $u(\lambda)$be the density density function of Λ. Then the function of generating opportunities *N* is as follows:

$$G_N(x) = \int G(x|\lambda)u(\lambda)d\lambda \text{ , if } \Lambda \text{ kontinu}$$

and the opportunity mass function for *N* as follows:

$$p_N(x) = \int p_N(x|\lambda)u(\lambda)d\lambda \text{ ,} if \ \Lambda \text{ kontinu}$$

(Klugman *et al* 2008*).*

Negative Binomial Mixed Distribution - Lindley

X is a random variable that spreads the Negative-Lindley Binomial with parameters *(r, θ),* denoted by $NB - L(r, \theta)$, if it meets the stochastic representation as follows

$X|\Lambda = \lambda \sim NB(r, p = e^{-\lambda})$ dan $\Lambda \sim Lin\ (\theta)$

with $r > 0$ and $\theta > 0$.

Characteristic 1

The mass function opportunity $X \sim NB - L(r, \theta)$ is

$$p_X(x) = \frac{\theta^2}{\theta+1}\binom{r+x-1}{x}\sum_{j=0}^{x}\binom{x}{j}(-1)^j\frac{(\theta+r+j+1)}{(\theta+r+j)^2} \text{ , dengan } x = 0,1,2 \ldots .$$

Characteristic 2

The k-factorial moment $X \sim NB - L(r, \theta)$ is

$$\mu_{[k]}(X) = \frac{\Gamma(r+k)}{\Gamma(r)}\frac{\theta^2}{\theta+1}\sum_{j=0}^{k}\binom{k}{j}(-1)^j\frac{(\theta-k+j+1)}{(\theta-k+j)^2}\text{, dengan } k = 0,1,2, \ldots$$

(Zamani dan Ismail 2010).

Characteristic 3

The Negative-Lindley Binomial Distribution has the characteristic that is the probability of $P(X = 0)$ which is always the greatest in all values of *r* and θ. So the Negative-Lindley Binomial distribution is assumed to produce a better estimation of claim frequency data if $P(X = 0)$ has the greatest value than the others. Since the Negative-Lindley Binomial distribution has $P(X = 0) > P(X = x)$, then the mass function chance of the Negative-Lindley Binomial distribution must be a monotonous function down. By using the Wolfram Mathematica 11.2 software, the results are obtained that the mass function of the Negative-

Lindley Binomial distribution opportunity is monotonous, meaning that $P(X=0) > P(X=x)$. For more details, the plot of the mass function opportunity is presented in Figures 3 to 8.

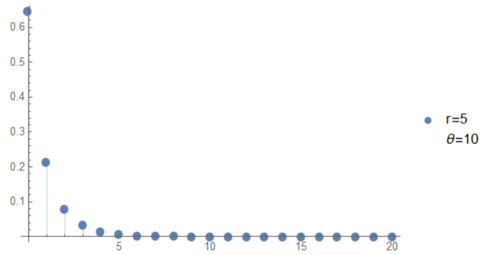

Figure 3: Plot $(\theta = 10, r = 5)$**.**

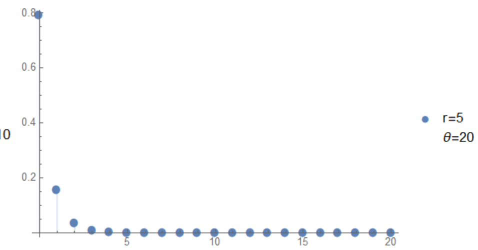

Figure 4: Plot $(\theta = 20, r = 5)$**.**

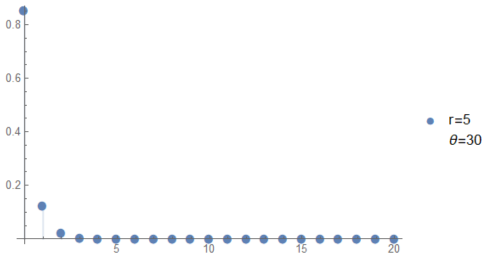

Figure 5: Plot $(\theta = 30, r = 5)$**.**

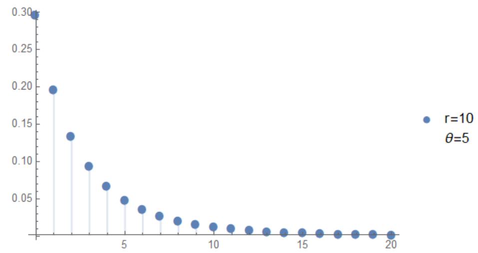

Figure 6: Plot $(\theta = 5, r = 10)$**.**

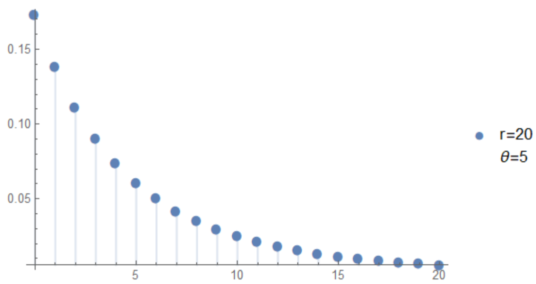

Figure 7: Plot $(\theta = 5, r = 20)$**.**

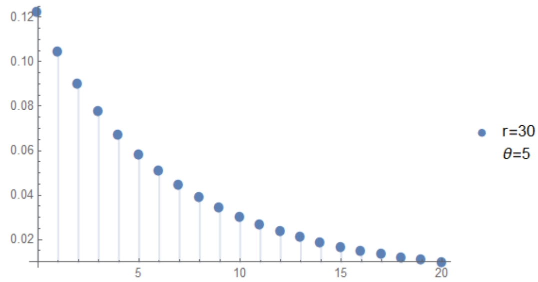

Figure 8: Plot $(\theta = 5, r = 30)$**.**

Estimation of Negative-Lindley Binomial Distribution Parameters with the Maximum Possible Method

Log Likehood distribution function NB-L(r, θ)**:**

$$L = \prod_{x=0}^{k} p_X(x)^{n_x}$$

$$\log L(r,\theta) = \mathcal{L}(r,\theta) = \sum_{x=0}^{k} n_x \log p_X(x)$$

$$= \sum_{x=0}^{k} n_x \log\left[\frac{\theta^2}{\theta+1}\binom{r+x-1}{x}\sum_{j=0}^{x}\binom{x}{j}(-1)^j \frac{\theta+r+j+1}{(\theta+r+j)^2}\right]$$

$$= \sum_{x=0}^{k} n_x \left(\log\left[\frac{\theta^2}{\theta+1}\right] + \log\binom{r+x-1}{x} + \log\sum_{j=0}^{x}\binom{x}{j}(-1)^j \frac{\theta+r+j+1}{(\theta+r+j)^2} \right)$$

where k = 0,1,2,3, ... and n_x is the frequency of the xth claim.

By taking partial derivatives of the Log Likelihood function respectively, r and θ will be obtained, namely by equating the two partial derivatives to zero, so we get an equation

$$\frac{\partial}{\partial\theta}\mathcal{L}(r,\theta) = n\left(\frac{2}{\theta} - \frac{1}{\theta+1}\right) + \sum_{x=0}^{k} n_x \left(\frac{\sum_{j=0}^{x}\binom{x}{j}(-1)^{j+1}\frac{\theta+r+j+2}{(\theta+r+j)^3}}{\sum_{j=0}^{x}\binom{x}{j}(-1)^{j}\frac{\theta+r+j+1}{(\theta+r+j)^2}} \right) = 0$$

where, $n = \sum_{x=0}^{k} n_x$ and

$$\frac{\partial}{\partial r}\mathcal{L}(r,\theta) = \frac{\partial}{\partial r}\sum_{x=0}^{k} n_x log\binom{x+r-1}{x}$$

$$+ \sum_{x=0}^{k} n_x \left(\frac{\sum_{j=0}^{x}\binom{x}{j}(-1)^{j+1}\frac{\theta+r+j+2}{(\theta+r+j)^3}}{\sum_{j=0}^{x}\binom{x}{j}(-1)^{j}\frac{\theta+r+j+1}{(\theta+r+j)^2}} \right) = 0\,.$$

According to Klugman et al (2008), it was shown that

$$\frac{\partial}{\partial r}\sum_{x=0}^{k} n_x \log\binom{x+r-1}{x} = \frac{\partial}{\partial r}\sum_{x=0}^{k} n_x \log\frac{(x+r-1)(x+r-2)\ldots(r)(r-1)!}{x!\,(r-1)!}$$

$$= \frac{\partial}{\partial r}\sum_{x=0}^{k} n_x \log\left(\prod_{m=0}^{x-1}(r+m)\right)$$

$$= \sum_{x=0}^{k} n_x \frac{\partial}{\partial r}\sum_{m=0}^{x-1}\log(r+m)$$

$$= \sum_{x=0}^{k} n_x \sum_{m=0}^{x-1}\frac{1}{r+m}\,.$$

So we get results

$$n\left(\frac{2}{\theta} - \frac{1}{\theta+1}\right) = \sum_{x=0}^{k} n_x \sum_{m=0}^{x-1}\frac{1}{r+m}$$

di mana, $n = \sum_{x=0}^{k} n_x$.

Assume

$$B=\sum_{x=0}^{k} n_x \sum_{m=0}^{x-1} \frac{1}{r+m}\ .$$

From these equations we can know the roots of the equation are two different real roots. However, because *θ*> *0*, the root that satisfies $\hat{\theta}(r)$ is as follows

$$\hat{\theta}(r)=\frac{n-B+\sqrt{(B-n)^2+8nB}}{2B}$$

where $n=\sum_{x=0}^{k} n_x$.

$H(\hat{r})$ is assumed to be the first derivative of $\mathcal{L}(r,\theta)$ with respect to the parameter *r*

$$H(\hat{r})=\sum_{x=0}^{k} n_x \sum_{m=0}^{x-1} \frac{1}{\hat{r}+m}+\sum_{x=0}^{k} n_x \left(\frac{\sum_{j=0}^{x}\binom{x}{j}(-1)^{j+1}\frac{\hat{\theta}+\hat{r}+j+2}{(\hat{\theta}+\hat{r}+j)^3}}{\sum_{j=0}^{x}\binom{x}{j}(-1)^{j}\frac{\hat{\theta}+\hat{r}+j+1}{(\hat{\theta}+\hat{r}+j)^2}}\right).$$

With the first derivative $H(\hat{r})$ is

$$H'(\hat{r})=-\sum_{x=0}^{k} n_x \sum_{m=0}^{x-1} \frac{1}{(\hat{r}+m)^2}+\sum_{x=0}^{k} n_x$$

$$\left(\frac{\left(\sum_{j=0}^{x}\binom{x}{j}(-1)^{j+1}\frac{-2\left(\hat{\theta}+\hat{r}+j+3\right)}{\left(\hat{\theta}+\hat{r}+j\right)^4}\right)\left(\sum_{j=0}^{x}\binom{x}{j}(-1)^{j}\frac{\hat{\theta}+\hat{r}+j+1}{\left(\hat{\theta}+\hat{r}+j\right)^2}\right)-\left(\sum_{j=0}^{x}\binom{x}{j}(-1)^{j}\frac{-\left(\hat{\theta}+\hat{r}+j+2\right)}{\left(\hat{\theta}+\hat{r}+j\right)^2}\right)\left(\sum_{j=0}^{x}\binom{x}{j}(-1)^{j+1}\frac{\hat{\theta}+\hat{r}+j+2}{\left(\hat{\theta}+\hat{r}+j\right)^3}\right)}{\left[\sum_{j=0}^{x}\binom{x}{j}(-1)^{j}\frac{\hat{\theta}+\hat{r}+j+1}{\left(\hat{\theta}+\hat{r}+j\right)^2}\right]^2}\right)=0$$

The solution to find the parameter $\hat{r}$ is the Newton Raphson method numerically. The equation needed for the *k*-th interaction, namely:

$$r_k=r_{k-1}-\frac{H(r_{k-1})}{H'(r_{k-1})}$$

and to find $\hat{\theta}$, substitute $\hat{r}$ which has been obtained to the following equation:

$$n\left(\frac{2}{\theta}-\frac{1}{\theta+1}\right)=\sum_{x=0}^{k} n_x \sum_{m=0}^{x-1} \frac{1}{r+m}$$

Application In Motor Vehicle Insurance

Motor vehicle insurance is an interesting thing to study because it is one type of loss insurance that is of interest to consumers. In everyday life, modeling the frequency of claims data is one of the most important topics, especially in the insurance field. The frequency of claims is modeled by discrete distributions such as Poisson, geometric, and negative binomial distributions. Let X be a random variable whose value is in the form of census data. The distribution commonly used to model X is the Poisson distribution. Poisson distribution has a special characteristic, which has the same average value and variety.

In fact, census data is often found with a greater variety of conditions than it should be, this event is called overdispersion. When overdispersion occurs, the assumption of average similarity and variance in Poisson is violated, so it is necessary to look for other distributions that can be used to analyze the chopped data when overdispersion occurs. One distribution that can be used to overcome the overdispersion problem is negative binomials. However, negative binomial distribution is not suitable if used for insurance claim frequency data with a large number of frequencies that do not submit claims. Therefore the negative binomial distribution is modified by mixing it with the Lindley distribution. The Negative-Lindley Binomial mixture distribution can overcome this problem with very good accuracy.

Data

The data used in this study are data sourced from Klugman et al (2008).

Table 1: Data on the frequency of motor vehicle insurance claims.

No	Claim Frequency	Number of Policy Holders
1	0	81714
2	1	11306
3	2	1618
4	3	250
5	4	40
6	5	7
Total		94935

Data Overdispersion

Overdispersion is a condition when the range of response variables is greater than the mean value of the response variable. Overdispersion can occur because of the large number of zero observations on the response variable. The presence or absence of overdispersion can be seen from the Pearson Chi-square value divided by the degree of freedom. If the Pearson Chi-square value divided by degrees of freedom is greater than 1, this shows the value of variance greater than the expected value.

By using Mathematica 11.2 software, it was found that the data from Table 1 overdispersed the data with an average value of 15822.5 and a variance value of 8.84×10^8.

Negative Binomial Fit Test

In this section, the Negative Binomial (r, p) distribution matching test on the category 1 claims of motor vehicle insurance policy holders in Table 1 where the percentage of frequency data that does not submit a claim is (81714/94935) * 100% = 86,074%. A percentage of 86,074% indicates that the number who did not file a claim was very large. With a hypothesis:

H0 : Frequency data for category 1 motor vehicle insurance claims comes from the population of Negative Binomial distribution

H1 : Category 1 motor vehicle insurance claim frequency data does not originate from the Negative Binomial distribution population

The next stage is the estimation of the Negative-Lindley Binomial distribution parameters. With the help of software R 3.5.1 obtained the results of estimating the parameters $\hat{r} = 0.8992$ and $\hat{p} = 0.846$. Based on the estimated values of the parameters *r* and *p*, the expected value for each claim frequency can be calculated.

Table 2. Estimating the opportunity value and expectation of claims for negative Binomial distribution.

Category	Claim Frequency	Number of Policy Holders	Opportunity for Claims	Expectation Value of Claims	Level of accuracy
1	0	81714	0.86078	81718.450	99.995%
2	1	11306	0.11886	11283.952	99.805%
3	2	1618	0.01733	1645.401	98.335%
4	3	250	0.00257	244.171	97.668%
5	4	40	0.00038	36.549	91.372%
6	5	7	0.00006	5.499	78.558%
Total		94935	1	94934	99.998%
Parameter				$r = 0.8992$; $p = 0.846$	
Degree of Freedom				5	
p-*value*				97.5%	
Log likelihood				-44764.49	

Can be seen from Table 3 column 5 of the last row, which is 1.37414. With a real level of 1% the quantile value of the chi-square distribution with free degrees 5 is 15,086. Thus the conclusion does not reject H0, which means that the frequency data for category 1 motor vehicle insurance claims comes from the population of the Negative Binomial distribution.

Table 3: Values needed to calculate test statistics.

Claim Frequency (x)	**Number of Policy Holders** (O_i)	**Opportunity for Claims** (p_i)	**Expectation Value of Claims** (E_i)	$\frac{(O_i - E_i)^2}{E_i}$
0	81714	0.86078	81718.450	0.00024
1	11306	0.11886	11283.952	0.04289
2	1618	0.01733	1645.401	0.45628
3	250	0.00257	244.171	0.13915
4	40	0.00038	36.549	0.32591
5	7	0.00006	5.499	0.40967
Total	94935	1	94934	1.37414

Negative-Lindley Binomial Fit Test

This section tests the suitability of the Negative-Lindley Binomial distribution *(r, θ)* on the frequency 1 claim of motorized insurance policy holders in Table 1 where the percentage of frequency data that does not submit a claim is (81714/94935) * 100% = 86,074% . A percentage of 86,074% indicates that the number who did not file a claim was very large. With a hypothesis:

H0 : frequency data for category 1 motor vehicle insurance claims comes from the Negative-Lindley Binomial distribution population

H1 : category 1 motor vehicle insurance claim frequency data does not originate from the Negative-Lindley Binomial distribution population.

The next stage is the estimation of the Negative-Lindley Binomial distribution parameters. With the help of software R 3.5.1 obtained the results of estimating the parameters $\hat{r} = 19.10969$ and $\hat{\theta} = 119.35747$. Based on the estimated values of the parameters r and θ, the expected value for each claim frequency can be calculated.

Table 4: Estimating the opportunity value and expectation of the occurrence of the Negative-Lindley Binomial distribution.

Category	Claim Frequency	Number of Policy Holders	Opportunity for Claims	Expectation Value of Claims	Level of accuracy
1	0	81714	0.86100	81739.298	99.969%
2	1	11306	0.11881	11279.589	99.766%
3	2	1618	0.01713	1626.152	99.499%
4	3	250	0.00257	244.333	97.733%
5	4	40	0.00040	38.177	95.443%
6	5	7	0.00006	6.191	88.442%
Total		94935	1	94933.7	99.998%
Parameter				$r = 19.10969$ $\theta = 119.35747$	
Degree of Freedom				5	
p-*value*				99.9%	
Log likelihood				-44764.32	

Can be seen from Table 5 the last 5 rows, i.e. 0.43466. With a real level of 1% the quantile value of the chi-square distribution with free degrees 5 is 15,086. Thus the conclusion does not reject H0, which means that the frequency data for motor vehicle insurance claims for category 1 originates from the Negative-Lindley Binomial distribution population.

Table 5: Values needed to calculate test statistics.

Claim Frequency (x)	**Number of Policy Holders** (O_i)	**Opportunity for Claims** (p_i)	**Expectation Value of Claims** (E_i)	$\frac{(O_i - E_i)^2}{E_i}$
0	81714	0.86100	81739.298	0.00783
1	11306	0.11881	11279.589	0.06179
2	1618	0.01713	1626.152	0.04085
3	250	0.00257	244.333	0.13144
4	40	0.00040	38.177	0.08703
5	7	0.00006	6.191	0.10572
Jumlah	94935	1	94933.7	0.43466

Calculation of Percentage of Policyholders in Filing a Claim

In the Negative-Lindley Binomial Fit Test section it has been discussed that frequency 1 category of motor insurance claims data comes from the Negative-Lindley Binomial distribution population. Therefore, estimating the opportunity of the Negative-Lindley Binomial distribution can be used to calculate the percentage of category 1 motor vehicle insurance policy holders who submit claims. The opportunity values are presented in Table 4 and Table 5.

Based on Table 4 and Table 5 shows that the opportunity for a policyholder not to submit a claim is 0.86100. If this data represents the frequency of insurance claim data in Indonesia, then if there are 1,000,000 policyholders, there would be a possibility that approximately 861,000 policyholders would not submit claims.

Comparison of Estimating Parameters by Using Negative Binomial and Negative-Lindley Binomials

The difference between the estimation of the expected value of claims using the Negative Binomial distribution and the Negative-Lindley Binomial is not significant. It can be seen in Table 2 and Table 4 that the overall accuracy is the same, with a value of 99,998% with an average difference of 96,809% -94,289% = 2.52%. However, the accuracy of each category will be compared in this case. Table 4 shows the accuracy level of Negative-Lindley Binomial distribution in each category is better than the accuracy level of Negative Binomial distribution. For more details the differences in the level of accuracy are presented in Table 6.

Table 6: Differences in the accuracy of Negative Binomial and Negative Binomial-Lindley.

Category	Claim Frequency	Negative Binomial Accuracy Rate	Negative-Lindley Binomial Accuracy Level
1	0	99.995%	99.969%
2	1	99.805%	99.766%
3	2	98.335%	99.499%
4	3	97.668%	97.733%
5	4	91.372%	95.443%
6	5	78.558%	88.442%
Rata-rata		94.289%	96.809%

Table 6 it is found that the estimated probability of the occurrence of claims with the Negative-Lindley Binomial distribution is better than the Negative Binomial distribution.

CONCLUSION

The Negative-Lindley Binomial Distribution is a new mixed distribution between the Negative Binomial distribution and the Lindley distribution. The Negative-Lindley Binomial Distribution is suitable for modeling motor vehicle insurance claim frequency data. Estimation of parameters in the Negative-Lindley Binomial uses the Maximum Likelihood Method.

Estimation of parameters using software R 3.5.1 and Wolfram Mathematica 11.2. The Negative-Lindley Binomial Distribution is better than the Negative Binomial Distribution in modeling claim frequency data. Lindley's Negative Binomial Distribution has an accuracy rate of 99.998%. It can be ascertained that the Negative-Lindley Binomial distribution is very suitable for modeling claim frequency data.

REFERENCES

1. Bouk MA. Pendugaan Model Regresi Binomial Negatif dengan Metode Kemungkinan Maksimum[skripsi]. Yogyakarta (ID): Universitas Sanata Dharma, 2016.
2. Bowers NL, Gerber HU, Hickman JC, Jones DA, Nesbitt CJ. *Actuarial Mathematics*. 2nd Edition. Schaumburg (US): The Society of Actuaries, 1997.
3. Breton JC, Privault N. Factorial Moments of Point Processes. *Stochastic Processes and their Applications*, 2014; 124(10): 3412-3428.
4. Ghahramani S. *Fundamentals of Probability*. 3th Edition. New York (US): Prentice-Hall, 2005.
5. Grimmett GR, Stirzaker DR. *Probability and Random Processes.* 2nd Edition. Oxford: Clarendon Press, 1992.
6. Hogg RV, Craig AT, McKean JW. *Introduction to Mathematical Statistics.* 7th Edition. New Jersey (US): Prentice-Hall, 2014.
7. Klugman SA, Panjer HH, Willmot GE. *Loss Models: From Data to Decision*. 3th Edition. New Jersey (US): John Wiley and Sons, 2008.
8. Pemerintah Republik Indonesia. Undang-Undang Republik Indonesia Nomor 40 Tahun 2014 tentang Usaha Perasuransian. Jakarta (ID): Sekretariat Negara, 2014.
9. Zamani H, Ismail N. Negative Binomial-Lindley Distribution and Its Application. *Journal of Mathematics and Statistics*, 2010; 6(1): 4-9.