Penerapan Distribusi Lindley pada Data Besar Klaim Asuransi Kendaraan Bermotor di Indonesia

Puput Aryanti*, Aceng Komarudin Mutaqin

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

*puput.aryanti99@gmail.com, aceng.k.mutaqin@gmail.com

Abstract. Motor vehicle insurance is a standard insurance policy used in Indonesia that includes combined coverage, namely coverage of the motor vehicle itself and legal liability coverage to third parties. One of the things that insurance companies need to pay attention to as risk-receiving institution is modeling the amount of the claim. A suitable distribution for modeling the amount of the claim is a continuous distribution that is non-negative, has a positive skew, and has a heavy tail. The Lindley distribution is a mixed exponential distribution and has criteria suitable for modeling the amount of the claim. The parameters were estimated using the maximum likelihood method. Testing the suitability of the distribution using the Kolmogorov-Smirnov test. The data used is secondary data of PT. X in 2019 which consists of data on large claims of motor vehicle insurance policyholders in category 7 region 2. Based on the results of the application of the Lindley distribution to model the 2019 PT. X motor vehicle insurance claim data using the Kolmogorov-Smirnov test, it can be concluded that the 2019 PT. X motor vehicle insurance claim data comes from a Lindley-distributed population with an estimated value of the Lindley distribution parameter of $\theta = 0.0000002995$.

Keywords: *Motor Vehicle Insurance, Lindley Distribution, Kolmogorov-Smirnov Test.*

Abstrak. Asuransi kendaraan bermotor adalah polis asuransi standar yang digunakan di Indonesia yang mencakup pertanggungan gabungan yaitu pertanggungan atas kendaraan bermotor itu sendiri dan pertanggungan tanggung jawab hukum kepada pihak ketiga. Salah satu hal yang perlu diperhatikan perusahaan asuransi sebagai lembaga penerima risiko adalah memodelkan data besar klaim. Distribusi yang cocok untuk memodelkan data besar klaim adalah distribusi kontinu yang non negatif, memiliki kemiringan yang positif, dan memiliki ekor yang tebal. Distribusi Lindley merupakan distribusi eksponensial campuran dan memiliki kriteria yang sesuai untuk pemodelan data besar klaim. Parameter yang ada ditaksir menggunakan metode kemungkinan maksimum. Pengujian kecocokan distribusi menggunakan uji Kolmogorov-Smirnov. Data yang digunakan merupakan data sekunder PT. X tahun 2019 yang terdiri dari data besar klaim pemegang polis asuransi kendaraan bermotor kategori 7 wilayah 2. Berdasarkan hasil penerapan distribusi Lindley untuk memodelkan data besar klaim asuransi kendaraan bermotor PT. X tahun 2019 dengan menggunakan uji Kolmogorov-Smirnov, dapat disimpulkan bahwa data besar klaim asuransi kendaraan bermotor PT. X tahun 2019 berasal dari populasi yang berdistribusi Lindley dengan nilai taksiran parameter distribusi Lindley sebesar θ=0,0000002995.

Kata Kunci: *Asuransi Kendaraan Bermotor, Distribusi Lindley, Uji Kolmogorov-Smirnov*

Corresponding Author
Email: aceng.k.mutaqin@gmail.com

A. Pendahuluan

Pada prinsipnya, setiap manusia selalu dihadapkan pada risiko, risiko sendiri berkaitan dengan kemungkinan kerugian dan ketidakpastian. Dalam bidang asuransi, terjadinya risiko dari pihak pemegang polis dapat menimbulkan klaim. Klaim adalah ganti rugi atas suatu risiko kerugian. Pengertian asuransi terdapat dalam Pasal 246 Kitab Undang-undang Hukum Dagang (KUHD).

Terdapat dua jenis asuransi yaitu asuransi kerugian dan asuransi jiwa. Salah satu produk dari asuransi kerugian yaitu asuransi kendaraan bermotor. Menurut situs OJK (Otoritas Jasa Keuangan), asuransi kendaraan bermotor adalah polis asuransi standar yang digunakan di Indonesia dan dikeluarkan oleh AAUI (Asosiasi Asuransi Umum Indonesia), mencakup pertanggungan gabungan yaitu pertanggungan atas kendaraan bermotor itu sendiri dan pertanggungan tanggung jawab hukum kepada pihak ketiga. Asuransi ini memberikan manfaat berupa ganti rugi atas kerugian dan atau kerusakan pada kendaraan bermotor yang antara lain disebabkan oleh tabrakan, terperosok, perbuatan jahat, pencurian, atau kebakaran [8].

Salah satu hal yang perlu diperhatikan perusahaan asuransi sebagai lembaga penerima risiko adalah memodelkan data besar klaim. Besar klaim merupakan besar biaya yang perlu dibayarkan perusahaan asuransi yang timbul dari kerugian yang dialami oleh pemegang polis. Menurut penelitian Omari, dkk. (2018), distribusi yang cocok untuk memodelkan data besar klaim adalah distribusi kontinu yang non negatif, memiliki kemiringan yang positif, dan memiliki *heavy tail* [1]. Ada beberapa distribusi yang digunakan untuk memodelkan data besar klaim, di antaranya adalah distribusi eksponensial atau distribusi eksponensial campuran. Distribusi Lindley merupakan distribusi eksponensial campuran dan memiliki kriteria yang sesuai untuk pemodelan data besar klaim. Menurut Ghitany, dkk. (2008) dalam penelitiannya menyimpulkan bahwa distribusi Lindley lebih baik dalam pemodelan dibanding dengan distribusi eksponensial karena distribusi Lindley lebih fleksibel dalam berbagai sifat matematis [2]. Dalam penelitian ini akan dibahas bagaimana penerapan distribusi Lindley pada data besar klaim asuransi kendaraan bermotor di Indonesia dengan menggunakan uji Kolmogorov-Smirnov.

B. Metodologi Penelitian

Asuransi

Menurut Pasal 246 Kitab Undang-undang Hukum Dagang (KUHD), asuransi merupakan suatu perjanjian di mana perusahaan asuransi mengikatkan diri kepada pemegang polis, dengan menerima suatu premi untuk memberikan penggantian karena suatu kerugian, kerusakan, atau kehilangan keuntungan yang mungkin terjadi karena peristiwa tak tentu [3].

Asuransi Kendaraan Bermotor

Menurut Peraturan Menteri Keuangan Nomor 74/PMK.010/2007, asuransi kendaraan bermotor adalah produk asuransi kerugian yang melindungi pemegang polis dari risiko kerugian yang mungkin timbul berkaitan dengan kepemilikan dan penggunaan kendaraan bermotor [4]. Terdapat dua jenis pertanggungan asuransi kendaraan bermotor yaitu *Comprehensive (All Risk)* dan *Total Loss Only (TLO)*. Comprehensive (All Risk) menjamin risiko kerugian secara keseluruhan dengan jenis klaim yang diajukan yaitu total loss dan partial loss. Klaim total loss hanya bisa dilakukan sekali, sedangkan klaim partial loss bisa dilakukan lebih dari sekali dengan nilai klaim kurang dari 75% dari nilai kendaraan. Jenis yang kedua yaitu Total Loss Only, TLO hanya memberikan jaminan pergantian jika kendaraan mengalami kerusakan yang nilainya mencapai lebih dari 75% dari nilai kendaraan dan kerugian akibat kehilangan dengan jenis klaim yang diajukan adalah total loss.

Distribusi Lindley

Lindley memperkenalkan distribusi Lindley pada tahun 1958. Ghitany dkk. (2008) menyatakan bahwa distribusi Lindley merupakan distribusi campuran dari distribusi eksponensial (θ) dan distribusi Gamma $(2, \theta)$ dengan bobot $w = \frac{1}{1+\theta}$ [2]. Misalkan X adalah peubah acak kontinu yang berdistribusi Lindley dengan parameter $\theta > 0$, maka X akan memiliki fungsi densitas

peluang sebagai berikut:

$$f(x) = wf_1(x) + (1-w)f_2(x)$$

di mana $f_1(x)$ merupakan fungsi densitas distribusi Gamma (2, θ) dan $f_2(x)$ merupakan fungsi densitas distribusi eksponensial (θ), maka diketahui fungsi densitas distribusi Lindley adalah sebagai berikut:

$$f(x) = \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}, \text{untuk } x > 0$$

Fungsi distribusi kumulatif dari distribusi Lindley yaitu:

$$F(x) = 1 - \frac{e^{-\theta x}(1+\theta+\theta x)}{\theta+1}$$

Ekspektasi dan varians dari distribusi Lindley masing-masing sebagai berikut [5]:

$$E(X) = \frac{\theta+2}{\theta(\theta+1)}$$

$$V(X) = \frac{\theta^2+4\theta+2}{\theta^2(\theta+1)^2}$$

Taksiran parameter θ dari distribusi Lindley dapat diperoleh menggunakan metode kemungkinan maksimum [6]. Misalkan $X_1, X_2, \dots, X_n$ adalah suatu sampel acak berukuran n berdistribusi Lindley dengan parameter θ, dengan nilai dari sampel acak tersebut adalah $x_1, x_2, \dots, x_n$. Fungsi *likelihood*-nya adalah sebagai berikut:

$$L(\theta) = \frac{\theta^{2n}}{(\theta+1)^n}\prod_{i=1}^{n}(1+x_i)e^{-\theta\sum_{i=1}^{n}x_i}$$

Fungsi *log-likelihood* untuk distribusi Lindley adalah sebagai berikut:

$$l(\theta) = 2n\ln\theta - n\ln(\theta+1) + \sum_{i=1}^{n}\ln(1+x_i) - \theta\sum_{i=1}^{n}x_i$$

Taksiran parameter distribusi Lindley tidak dapat diselesaikan secara analitik, maka digunakan metode numerik yaitu metode iterasi Newton-Raphson. Metode iterasi Newton-Raphson memerlukan nilai awal, turunan pertama dari fungsi *log-likelihood* terhadap parameter θ dan turunan kedua dari fungsi *log-likelihood* terhadap parameter θ. Proses iterasi Newton-Raphson dapat dihitung dengan bantuan Pemrograman RStudio, dengan nilai awal yaitu:

$$\theta_0 = -\frac{\bar{x} - 1 - \sqrt{\bar{x}^2 + 6\bar{x} + 1}}{2\bar{x}}$$

Uji Kolmogorov-Smirnov

Uji *Kolmogorov-Smirnov* merupakan salah satu uji kecocokan model untuk peubah acak kontinu. Misalkan sampel acak berukuran n, yaitu $X_1, X_2, \dots, X_n$, di mana realisasi dari sampel acak tersebut adalah $x_1, x_2, \dots, x_n$. Rumusan hipotesis untuk uji *Kolmogorov-Smirnov* adalah [7]:

H_0 : Data berasal dari suatu populasi yang berdistribusi tertentu.

H_1 : Data bukan berasal dari suatu populasi yang berdistribusi tertentu.

Statistik uji *Kolmogorov-Smirnov* untuk hipotesis di atas adalah:

$$D = \max_{1\le i\le n}|F_n(x_i) - F^*(x_i)|$$

di mana $F_n(x_i)$ merupakan fungsi distribusi empiris untuk data pengamatan ke-i yang dapat dihitung dengan persamaan berikut:

$$F_n(x_i) = \frac{\text{Banyaknya pengamatan yang } \le x_i}{n}$$

$F^*(x_i)$ merupakan fungsi distribusi kumulatif dari model yang diuji untuk data pengamatan ke-i. Statistik ini didasarkan pada perbedaan maksimum antara fungsi distribusi empirik dari data dengan fungsi distribusi teoritik dari distribusi yang dicocokkan. . Hipotesis nol diterima apabila statistik uji D lebih kecil dari nilai kritis pada taraf nyata α yang telah ditentukan dan dapat disimpulkan bahwa data berasal dari suatu populasi berdistribusi tertentu.

Data dan Sumber Data

Data yang digunakan merupakan data sekunder hasil pencatatan yang diperoleh dari PT. X pada tahun 2019. Data berisi besar klaim dari setiap klaim yang diajukan oleh pemegang polis untuk produk asuransi kendaraan bermotor *comprehensive* di PT. X, kategori 7 wilayah 2. Klaim yang diajukan adalah *partial loss*. Tabel 1 menampilkan data besar klaim pemegang polis untuk produk asuransi kendaraan bermotor *comprehensive* selama satu periode asuransi (tahun).

Tabel 1. Data Besar Klaim Pemegang Polis Asuransi Kendaraan Bermotor

Nomor Klaim	Besar Klaim (Rupiah)
1	66.000
2	854.000
3	1.000.000
4	1.768.000
5	20.00.000
⋮	⋮
29	11.100.000
30	13.200.000
31	22.400.000
32	27.100.000
33	39.300.000

Sumber: Perusahaan Asuransi PT. X Tahun 2019

Tahapan Analisis

1. Menaksir parameter distribusi Lindley dengan menggunakan metode iterasi Newton-Raphson pada data besar klaim asuransi kendaraan bermotor di Indonesia.
2. Melakukan uji kecocokan distribusi dengan uji Kolmogorov-Smirnov untuk distribusi Lindley pada data besar klaim asuransi kendaraan bermotor di Indonesia.

C. Hasil Penelitian dan Pembahasan

Taksiran Parameter Distribusi Lindley

Nilai taksiran parameter parameter θ dari distribusi Lindley diperoleh berdasarkan iterasi Newton-Raphson. Sebelumnya perlu ditentukan terlebih dahulu nilai awal untuk iterasi Newton-Raphson yaitu:

$$\theta_0 = -\frac{6677355{,}788 - 1 - \sqrt{6677355{,}788^2 + 6677355{,}788 + 1}}{2(6677355{,}788)} = 0{,}0000002995$$

maka diperoleh nilai awal sebesar $\theta_0 = 0{,}0000002995$.

Dengan bantuan pemrograman RStudio, iterasi Newton-Raphson berhenti pada iterasi ke-1 dengan nilai $\left|\hat{\theta}^{(s+1)} - \hat{\theta}^{(s)}\right| = 1{,}4692 \times 10^{-14}$, di mana nilai tersebut lebih kecil dari toleransi kesalahan yang ditetapkan yaitu $\varepsilon = 10^{-10}$. Diperoleh nilai taksiran parameter θ distribusi Lindley sebesar $\theta = 0{,}0000002995$.

Uji Kecocokan Distribusi untuk Data Besar Klaim

Uji kecocokan distribusi Lindley pada data besar klaim pemegang polis asuransi kendaraan bermotor PT. X tahun 2013 dengan menggunakan uji kecocokan *Kolmogorov-Smirnov*. Hipotesis untuk pengujian tersebut adalah:

H_0: Data besar klaim asuransi kendaraan bermotor PT. X tahun 2013 berasal dari populasi yang berdistribusi Lindley.

H_1: Data besar klaim asuransi kendaraan bermotor PT. X tahun 2013 bukan berasal dari populasi yang berdistribusi Lindley.

Untuk membantu menghitung statistik uji Kolmogorov-Smirnov, dihitung nilai-nilai sebagaimana yang ada pada Tabel 2. Tabel 2 kolom (1) menjelaskan urutan data, kolom (2) memuat data besar klaim pemegang polis asuransi kendaraan bermotor PT. X tahun 2013 yang telah diurutkan dari nilai terkecil hingga terbesar. Kolom (3) merupakan nilai fungsi distribusi kumulatif dari distribusi Lindley untuk data pada kolom (2). Untuk menghitung nilai pada kolom (3) dibutuhkan nilai taksiran parameter θ dari distribusi Lindley, diperoleh nilai taksiran parameter θ distribusi Lindley sebesar $\theta = 0{,}0000002995$.

Berdasarkan nilai taksiran θ di atas dapat dihitung nilai taksiran fungsi distribusi kumulatif dari distribusi Lindley. Misal untuk data urutan pertama, $x_1 = 66.000$, dengan bantuan *software* Microsoft Excel diperoleh nilai fungsi distribusi empiris yaitu:

$$F_{33}(x_1) = F_{33}(66.000) = \frac{\text{Banyaknya pengamatan yang } \leq 66.000}{33} = \frac{1}{33} = 0{,}0303$$

dengan nilai fungsi distribusi kumulatif dari distribusi Lindley adalah:

$$F^*(x_1) = F^*(66.000)$$
$$= 1 - \frac{e^{-(0{,}0000002995)(66000)}\left(1 + 0{,}0000002995 + (0{,}0000002995)(66000)\right)}{0{,}0000002995 + 1}$$
$$= 0{,}0001928$$

Nilai taksiran fungsi distribusi kumulatif dari distribusi Lindley untuk data lainnya dihitung dengan cara yang sama sebagaimana di atas. Selanjutnya kolom (4) merupakan fungsi distribusi empiris, dan kolom (5) merupakan nilai absolut dari fungsi distribusi empiris yang terdapat pada kolom (4) dikurangi dengan nilai fungsi distribusi kumulatif distribusi Lindley pada kolom (3).

Tabel 2. Hasil Perhitungan Uji Kolmogorov-Smirnov pada Data Besar Klaim

i	x	$F_n(x_i)$	$F^*(x_i)$	$\lvert F_n(x_i) - F^*(x_i)\rvert$
(1)	(2)	(3)	(4)	(5)
1	66.000	0,03030	0,000193	0,030110188
2	854.000	0,06061	0,027636	0,032969952
3	1.000.000	0,09091	0,036830	0,054079393
4	1.768.000	0,12121	0,099294	0,021917719
5	2.000.000	0,15152	0,121585	0,029929873
⋮	⋮	⋮	⋮	⋮
29	11.100.000	0,87879	0,844379	0,034408602
30	13.200.000	0,90909	0,904967	0,004123657
31	22.400.000	0,93939	0,990598	0,051203959
32	27.100.000	0,96970	0,997279	0,027582274
33	39.300.000	1	0,999901	0,000098649

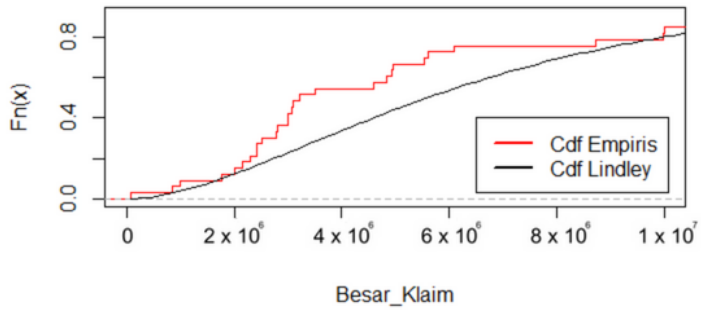
Berdasarkan hasil pada Tabel 2 dapat diperoleh nilai statistik uji Kolmogorov-Smirnov yaitu sebesar:

$$D = \max_{1 \leq i \leq n} |F_n(x_i) - F^*(x_i)| = 0{,}2652$$

Dari hasil perhitungan diperoleh nilai statistik uji Kolmogorov-Smirnov sebesar $0{,}2652$. Dengan taraf nyata $\alpha = 10\%$ dan $n = 33$, diperoleh nilai kritis berdasarkan tabel nilai kritis uji Kolmogorov-Smirnov sebesar $0{,}277$. Terlihat bahwa nilai statistik uji lebih kecil dibandingkan dengan nilai kritis $(0{,}2652 < 0{,}277)$, sehingga hipotesis nol diterima dan dapat disimpulkan bahwa data besar klaim pemegang polis asuransi kendaraan bermotor PT. X tahun 2019 berasal dari populasi berdistribusi Lindley.

Gambar 3 menampilkan kurva fungsi distribusi kumulatif empiris dan fungsi distribusi

kumulatif dari distribusi Lindley untuk data besar klaim asuransi kendaraan bermotor PT. X tahun 2019.



Gambar 1. Kurva Fungsi Distribusi Empiris dan Fungsi Distribusi Kumulatif dari Distribusi Lindley

Berdasarkan Gambar 1 terlihat bahwa kurva fungsi distribusi empiris berdekatan dengan fungsi distribusi kumulatif dari distribusi Lindley untuk data besar klaim asuransi kendaraan bermotor PT. X tahun 2019.

D. Kesimpulan

Berdasarkan hasil penerapan distribusi Lindley untuk memodelkan data besar klaim asuransi kendaraan bermotor PT. X tahun 2019 dengan menggunakan uji Kolmogorov-Smirnov, dapat disimpulkan bahwa data besar klaim asuransi kendaraan bermotor PT. X tahun 2019 berasal dari populasi yang berdistribusi Lindley dengan nilai taksiran parameter distribusi Lindley sebesar $\theta = 0{,}0000002995$.

Acknowledge

Penulis mengucapkan terimakasih kepada keluarga yang telah memberikan doa dan juga dukungan untuk menyelesaikan penelitian ini. Terimakasih juga kepada Bapak Dr. Aceng Komarudin Mutaqin, M.T., M.Si, selaku dosen pembimbing yang telah memberikan ilmu pengetahuan, motivasi, saran, waktu, dan kemudahan kepada penulis selama penelitian ini.

Daftar Pustaka

[1] Omari, C. O., Nyambura, S. G., & Mwangi, J. M. (2018). Modeling the Frequency and Severity of Auto Insurance Claims Using Statistical Distributions. *Journal of Mathematical Finance*, Vol.8, 137-160.

[2] Ghitany, M. E., Atieh, B., & Nadarajah, S. (2008). Lindley distribution and its application. *ScienceDirect: Mathematics and Computers in Simulation*, Vol.78, 493-506.

[3] Kitab Undang-undang Hukum Dagang (KUHD). Bab IX Tentang Asuransi atau Pertanggungan pada Umumnya, Pasal 246.

[4] Kementerian Keuangan. (2007). *Peraturan Menteri Keuangan Nomor 74/PMK.010/2007: Tentang Penyelenggaraan Pertanggungan Asuransi pada Lini Usaha Asuransi Kendaraan Bermotor*. Jakarta: Kemenkeu

[5] Zamani, H., & Ismail, N. (2010). Negative Binomial-Lindley Distribution and Its Application. *Journal of Mathematics and Statistics*, Vol.6, No.1, 4-9.

[6] Shanker, R., & Fesshaye, H. (2016). On modeling of lifetime data using aradhana, sujatha, lindley, and exponential distribution. *Biometric & Biostatistics International Journal*, Vol.4, 28-38.

[7] Nugroho, S. (2008). Metode Statistika Nonparametrika. Bengkulu: UNIB Press.

[8] M. R. N. Alam and A. K. Mutaqin, "Pemodelan Distribusi Poisson-Sujatha pada Data Frekuensi Klaim Asuransi Kendaraan Bermotor di Indonesia," *Jurnal Riset Statistika*, pp. 71–78, Jul. 2023, doi: 10.29313/jrs.v3i1.1944.