

Pemodelan Distribusi Pareto untuk Data Besar Klaim Asuransi Kendaraan Bermotor di Indonesia

Dede Yusuf*, Aceng Komarudin Mutaqin

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

*deyus0907@gmail.com, aceng.k.mutaqin@gmail.com

Abstract. This paper will discuss Pareto distribution modeling for data claims size on motor vehicle insurance in Indonesia. The Pareto distribution is a mixed distribution of the exponential distribution and the inverse gamma distribution. The Pareto distribution is a continuous random variable with parameters s and m . Pareto distribution modeling for data claims size of motor vehicle insurance in Indonesia was generated by the Kolmogorov-Smirnov fit test. The Kolmogorov-Smirnov test is one of the model fit tests for continuous random variables. The data used is secondary data recording results obtained from the insurance company PT. XYZ in 2014, the data contains the size of the partial loss claims of policyholders for comprehensive motor vehicle insurance product category 7 regions 1. The results of Pareto distribution modeling for large data of motor vehicle insurance claims in Indonesia using the Kolmogorov-Smirnov match test show that data claims size on Motor vehicles insurance category 7 region 1 in Indonesia comes from a population with a Pareto distribution. So that data on motor vehicle insurance category 7 region 1 in Indonesia can be used in the application of insurance premium calculations.

Keywords: Pareto Distribution, Kolmogorov Smirnov Fit Test, Partial Loss, Comprehensive.

Abstrak. Dalam makalah ini akan dibahas pemodelan distribusi Pareto untuk data besar klaim asuransi kendaraan bermotor di Indonesia. Distribusi Pareto merupakan distribusi campuran dari distribusi eksponensial dan distribusi inverse gamma. Distribusi Pareto merupakan peubah acak kontinu dengan parameter s dan m . Pemodelan distribusi Pareto untuk data besar klaim asuransi kendaraan bermotor di Indonesia dihasilkan dengan uji kecocokan Kolmogorov-smirnov. Uji Kolmogorov-Smirnov merupakan salah satu uji kecocokan model untuk peubah acak kontinu. Data yang digunakan adalah data sekunder hasil pencatatan yang diperoleh dari perusahaan asuransi PT. XYZ tahun 2014, data tersebut berisi besar klaim partial loss pemegang polis untuk produk comprehensive asuransi kendaraan bermotor kategori 7 wilayah 1. Hasil pemodelan distribusi Pareto untuk data besar klaim asuransi kendaraan bermotor di Indonesia dengan menggunakan uji kecocokan Kolmogorov-smirnov menunjukkan bahwa data besar klaim asuransi kendaraan bermotor kategori 7 wilayah 1 di Indonesia berasal dari populasi yang berdistribusi Pareto. Sehingga data asuransi kendaraan bermotor kategori 7 wilayah 1 di Indonesia dapat digunakan dalam penerapan perhitungan premi asuransi.

Kata Kunci: Distribusi Pareto, Uji Kecocokan Kolmogorov smirnov, *Partial Loss, Comprehensive.*

1. Pendahuluan

Kendaraan bermotor merupakan salah satu alat transportasi yang dimana angka pertumbuhannya terus meningkat setiap tahunnya khususnya di Indonesia. Banyaknya jumlah kendaraan bermotor membuat kondisi lalu lintas semakin padat dan terjadinya berbagai risiko yang dialami oleh manusia. Risiko yang terjadi seperti kecelakaan, kehilangan, dan lain hal

sebagainya yang dapat merugikan manusia itu sendiri. Maka dari itu pengendara harus mencari cara untuk mengalihkan risiko yang akan terjadi, salah satunya dengan cara menyerahkan pertanggung jawaban risiko kepada penanggungnya yaitu perusahaan asuransi. Menurut Kitab Undang-Undang Hukum Dagang (KUHD), bab 9, pasal 246 bahwa asuransi atau pertanggung jawaban adalah suatu perjanjian dengan mana seorang penanggung mengikatkan diri kepada seorang tertanggung, dengan menerima suatu premi, untuk memberikan penggantian kepadanya karena suatu kerugian, kerusakan, atau kehilangan keuntungan yang diharapkan, yang mungkin akan dideritanya karena suatu peristiwa yang tak tertentu.

Asuransi menurut UU. RI Nomor 40 tahun 2014 tentang Perasuransian merupakan perjanjian antara dua pihak, yaitu perusahaan asuransi dan pemegang polis yang menjadi dasar bagi penerimaan premi oleh perusahaan asuransi sebagai imbalan untuk memberikan penggantian atas kerugian finansial yang dialami oleh tertanggung atau pemegang polis.

Kendaraan bermotor merupakan barang investasi untuk kehidupan, sehingga masyarakat banyak menggunakan perusahaan asuransi khususnya asuransi kendaraan bermotor sebagai pengalihan risiko yang terjadi pada kendaraan bermotor mereka dari kejadian yang tidak diinginkan.

Asuransi kendaraan bermotor adalah produk asuransi kerugian yang melindungi tertanggung dari risiko kerugian yang mungkin timbul sehubungan dengan kepemilikan dan pemakaian kendaraan bermotor (Peraturan Menteri Keuangan Nomor 74/PMK.010/2007 tentang penyelenggaraan pertanggung jawaban asuransi pada lini usaha asuransi kendaraan bermotor dalam pasal 1 ayat 2).

Terdapat berbagai macam sistem penetapan premi yang ditawarkan oleh setiap perusahaan asuransi. Salah satu perhitungan dalam menetapkan premi yaitu dengan melibatkan besar klaim sebagai komponen dalam sebuah sistem yang akan digunakan. Adapun pemodelan yang diterapkan pada besar klaim yaitu dapat dimodelkan oleh distribusi pareto, atau weibull. Dalam hal ini, pemodelan yang digunakan yaitu besar klaim yang di modelkan oleh distribusi Pareto.

1. Distribusi Pareto diperkenalkan pertama kali oleh Vilfredo Pareto pada tahun 1890 dalam bidang ekonomi. Distribusi Pareto merupakan distribusi peluang yang membahas mengenai fenomena sosial, ilmiah geofisika dan banyak jenis lainnya. Salah satu peran penting yang dimiliki distribusi Pareto yaitu dalam penentuan risiko asuransi. Distribusi Pareto merupakan distribusi campuran dari distribusi eksponensial dan distribusi inverse gamma.
2. Berdasarkan penjelasan sebelumnya di atas, maka pembahasan dalam makalah ini adalah pemodelan distribusi Pareto untuk data besar klaim asuransi kendaraan bermotor di Indonesia. Adapun tujuan dari penulisan makalah ini adalah menguji kecocokan pada data besar klaim asuransi kendaraan bermotor di Indonesia pada pemodelan distribusi Pareto.

2. Metodologi

Data yang digunakan adalah data sekunder hasil pencatatan yang diperoleh dari PT. XYZ tahun 2014. Data tersebut berisi frekuensi klaim dan besar klaim pemegang polis untuk produk asuransi kendaraan bermotor *comprehensive* terhadap perusahaan asuransi PT XYZ. Data yang akan dipakai sebagai bahan aplikasi dalam skripsi ini adalah data pemegang polis asuransi kendaraan bermotor kategori 7 dimana kategori ini merupakan jenis bus yang terdapat pada Wilayah 1 (Selain Wilayah 2 dan Wilayah 3) dan klaim yang diajukan adalah *partial loss*.

Tabel 1. Data Besar Klaim *Partial Loss* Produk Asuransi Kendaraan Bermotor *Comprehensive* Kategori 7 Wilayah 1.

No. Tertanggung	Klaim yang disetujui (Rupiah)
1	16.250
2	200.000
3	325.000

4	348.300
5	348.530
6	350.000
7	400.000
8	441.500
9	500.000
10	506.000
⋮	⋮
237	130.000.000

Sumber: Perusahaan Asuransi PT XYZ Tahun 2014

Metode Uji Kecocokan Distribusi Pareto

Langkah-langkah uji kecocokan distribusi Pareto untuk data besar klaim asuransi kendaraan bermotor dengan menggunakan uji kecocokan Kolmogorov-Smirnov:

1. Merumuskan Hipotesis
 H_0 : data besar klaim berasal dari populasi berdistribusi Pareto sebagai distribusi campuran.
 H_1 : data besar klaim bukan berasal dari populasi yang berdistribusi Pareto sebagai distribusi campuran..
2. Menaksir parameter distribusi Pareto sebagai distribusi campuran dengan menggunakan Persamaan (3.8) dan (3.9).
3. Mengurutkan data besar klaim kendaraan bermotor dari nilai terkecil ke terbesar.
4. Menghitung fungsi distribusi empiris pada data pengamatan ke- i .
5. Menghitung fungsi distribusi kumulatif untuk distribusi Pareto sebagai distribusi campuran menggunakan Persamaan (3.5).
6. Menghitung nilai statistik uji Kolmogorov-Smirnov menggunakan Persamaan (3.10).
7. Membandingkan nilai statistik uji dengan nilai kritis pada taraf nyata α . Kemudian memutuskan apakah hipotesis nol diterima atau ditolak.

3. Pembahasan dan Diskusi

Distribusi Pareto

Distribusi Pareto merupakan distribusi campuran dari distribusi eksponensial dan distribusi inverse gamma. Asumsikan bahwa besar klaim, $X|\theta$ berdistribusi eksponensial dengan parameter θ , dengan fungsi densitas peluang sebagai berikut:

$$f(X|\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \quad x > 0, \theta > 0 \quad \dots (3.1)$$

Fungsi distribusi kumulatif dari distribusi eksponensial adalah:

$$F(x|\theta) = 1 - e^{-\frac{x}{\theta}} \quad x > 0 \text{ dan } \theta > 0 \quad \dots (3.2)$$

Ekspektasi dan varians masing-masing adalah:

$$E(X) = \theta$$

$$V(X) = \theta^2$$

dimana θ memiliki distribusi inverse gamma dengan parameter s dan m dengan fungsi densitas peluangnya adalah sebagai berikut:

$$f(\theta) = \frac{\frac{1}{m} e^{-\frac{m}{\theta}}}{\left(\frac{\theta}{m}\right)^{s+1} \Gamma(s)}, \quad \theta > 0, s > 1, m > 0 \quad \dots (3.3)$$

dengan ekspektasi:

$$E(\theta) = \frac{m}{s - 1}$$

Dengan demikian distribusi tak bersyarat dari besar klaim, X yaitu campuran dari distribusi eksponensial dan distribusi inverse gamma adalah distribusi Pareto dengan parameter s dan m dengan fungsi densitas sebagai berikut:

$$f(x) = \frac{sm^s}{(x + m)^{s+1}}, \quad x > 0, s > 0, m > 0 \quad \dots (3.4)$$

Fungsi distribusi kumulatif dari distribusi Pareto adalah sebagai berikut:

$$F(X) = 1 - \left(\frac{m}{x + m}\right)^s, \quad x > 0, s > 0, m > 0 \quad \dots (3.5)$$

Taksiran parameter dari distribusi Pareto sebagai distribusi campuran menggunakan metode kemungkinan maksimum. Fungsi likelihood untuk data di atas adalah:

$$L(f(x)) = \prod_{i=1}^n \frac{sm^s}{(x_i + m)^{s+1}} \quad \dots (3.6)$$

Fungsi log-likelihood-nya adalah:

$$\begin{aligned} l(f(x)) &= \ln \left(\frac{sm^s}{(x_1 + m)^{s+1}} - \frac{sm^s}{(x_2 + m)^{s+1}} - \dots - \frac{sm^s}{(x_n + m)^{s+1}} \right) \\ &= \ln \left(\frac{(sm^s)^n}{\prod_{i=1}^n (x_i + m)^{s+1}} \right) \\ &= n(\ln s + s \ln m) - (s + 1) \sum_{i=1}^n \ln(x_i + m) \end{aligned} \quad \dots (3.7)$$

Turunan fungsi log-likelihood terhadap parameter s :

$$\begin{aligned} \frac{\partial l(s, m)}{\partial s} &= n \frac{1}{s} + n \ln m - \sum_{i=1}^n \ln(x_i + m) \\ \frac{n}{s} &= \frac{\sum_{i=1}^n \ln(x_i + m) - n \ln m}{n} \\ s &= \frac{n}{\sum_{i=1}^n \ln(x_i + m) - n \ln m} \end{aligned} \quad \dots (3.8)$$

Turunan fungsi log-likelihood terhadap parameter m :

$$\begin{aligned} \frac{\partial l(s, m)}{\partial m} &= 0 + ns \frac{1}{m} - (s + 1) \sum_{i=1}^n \frac{1}{(x_i + m)} \\ \frac{ns}{m} &= (s + 1) \sum_{i=1}^n \frac{1}{(x_i + m)} \\ m &= \frac{ns}{(s + 1) \sum_{i=1}^n \frac{1}{(x_i + m)}} \end{aligned} \quad \dots (3.9)$$

Uji Kecocokan Kolmogorov-Smirnov

Uji Kolmogorov-Smirnov merupakan salah satu uji kecocokan model untuk peubah acak kontinu. Hipotesis uji Kolmogorov-Smirnov adalah:

H_0 : Data berasal dari suatu populasi berdistribusi $F(\cdot)$

H_1 : Data tidak berasal dari suatu populasi berdistribusi $F(\cdot)$

Misalkan suatu sampel acak berukuran n , yaitu X_1, X_2, \dots, X_n , dimana realisasi dari sampel acak tersebut adalah x_1, x_2, \dots, x_n . Statistik uji Kolmogorov-Smirnov untuk hipotesis di atas adalah:

$$D = \max_{1 \leq i \leq n} |F_n(x_i) - F^*(x_i)| \quad \dots (3.10)$$

dimana $F_n(x_i)$ adalah fungsi distribusi empiris pada data pengamatan ke- i , sedangkan $F^*(x_i)$ adalah fungsi distribusi kumulatif dari model yang diuji untuk data pengamatan ke- i . Dengan demikian, hipotesis nol diterima apabila nilai statistik uji D lebih kecil dari nilai kritisnya. Nilai kritis uji hipotesis di atas disajikan dalam Tabel 2.

Tabel 2. Nilai Kritis Uji Kolmogorov-Smirnov

Tingkat Signifikansi (α)	0,10	0,05	0,01
Nilai kritis	$\frac{1,22}{\sqrt{n}}$	$\frac{1,36}{\sqrt{n}}$	$\frac{1,63}{\sqrt{n}}$

Sumber: Klugman dkk. (2012)

Uj Kecocokan Besar Klaim Distribusi Pareto

Uji kecocokan distribusi Pareto sebagai distribusi campuran untuk data besar klaim tertanggung asuransi kendaraan bermotor Kategori 7 Wilayah 1 di Indonesia menggunakan uji Kolmogorov-Smirnov. Hipotesis untuk pengujian tersebut adalah:

H_0 : Data besar klaim asuransi kendaraan bermotor kategori 7 wilayah 1 di Indonesia berasal dari populasi yang berdistribusi Pareto.

H_1 : Data besar klaim asuransi kendaraan bermotor kategori 7 Wilayah 1 di Indonesia bukan berasal dari populasi yang berdistribusi Pareto.

Statistik uji untuk hipotesis di atas ada dalam Persamaan (3.10). Untuk membantu menghitung statistik uji Kolmogorv-Smirnov, dihitung nilai-nilai sebagaimana yang ada dalam Tabel 3.2 pada kolom (1) menjelaskan urutan data, kolom (2) memuat data besar klaim tertanggung asuransi kendaraan bermotor di Indonesia kategori 7 Wilayah 1 yang sudah diurutkan dari yang terkecil hingga terbesar, kolom (3) merupakan nilai fungsi distribusi empiris dari data besar klaim tertanggung, kolom (4) merupakan nilai fungsi distribusi kumulatif dari distribusi Pareto sebagai distribusi campuran. Untuk menghitung nilai yang ada di kolom (4) digunakan Persamaan (3.5) dan dibutuhkan nilai taksiran parameter s dan m dari distribusi Pareto sebagai distribusi campuran. Dengan menggunakan *software EasyFit 0.5*, maka nilai taksiran parameter $\hat{s} = 2,6213$ dan $\hat{m} = 13.880.000$.

Berdasarkan nilai taksiran parameter s dan m , dapat dihitung nilai taksiran fungsi distribusi kumulatif dari distribusi Pareto sebagai distribusi campuran. Misalkan untuk data urutan pertama, $x_{(1)} = 16.250$, dengan bantuan perangkat *software Microsoft Excel 2013* diperoleh nilai taksiran fungsi distribusi kumulatif dari distribusi Pareto sebagai distribusi campuran adalah:

$$F^*(x_i) = 1 - \left(\frac{13.880.000}{16.250 + 13.880.000} \right)^{2,6213} = 0,0031.$$

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

Tabel 3. Hasil Perhitungan untuk Uji Kolmogorov-Smirnov Data Besar Klaim

i	x_i	$F_n(x_i)$	$F^*(x_i)$	$ F_n(x_i) - F^*(x_i) $
1	16.250	0,0042	0,0031	0,0012
2	200.000	0,0084	0,0368	0,0284
3	325.000	0,0127	0,0589	0,0462

4	348.300	0,0169	0,0629	0,0460
5	348.530	0,0211	0,0629	0,0418
6	350.000	0,0253	0,0632	0,0379
7	400.000	0,0295	0,0718	0,0422
8	441.500	0,0338	0,0788	0,0450
9	500.000	0,0380	0,0886	0,0506
10	506.000	0,0422	0,0896	0,0474
⋮	⋮	⋮	⋮	⋮
237	13.000.0000	1	0,9978	0,0022

Dengan hasil perhitungan yang terdapat pada Tabel 3, dapat dihitung statistik uji Kolmogorov-Smirnov dengan menggunakan Persamaan (2.6) yaitu:

$$D = \max_{1 \leq i \leq n} |F_n(x_i) - F^*(x_i)| = 0,0506.$$

Dengan taraf nyata, $\alpha = 5\%$ maka nilai kritisnya adalah $\frac{1,36}{\sqrt{237}} = 0,0883$. Terlihat bahwa nilai statistik uji Kolmogorov-Smirnov di atas lebih kecil dibandingkan dengan nilai kritisnya, sehingga hipotesis nol diterima dan dapat disimpulkan bahwa data besar klaim tertanggung asuransi kendaraan bermotor di Indonesia kategori 7 wilayah 1 berasal dari populasi yang berdistribusi Pareto.

4. Kesimpulan

Berdasarkan pembahasan dalam penelitian ini, peneliti menyimpulkan bahwa hasil pemodelan distribusi Pareto untuk data besar klaim asuransi kendaraan bermotor di Indonesia dengan menggunakan uji kecocokan Kolmogorov-smirnov menunjukkan bahwa data besar klaim asuransi kendaraan bermotor kategori 7 wilayah 1 di Indonesia berasal dari populasi yang berdistribusi Pareto. Sehingga data asuransi kendaraan bermotor kategori 7 wilayah 1 di Indonesia dapat digunakan dalam penerapan perhitungan premi asuransi.

Acknowledge

Terima kasih banyak untuk semua pihak yang telah membantu dalam penelitian ini, sehingga penelitian ini dapat terselesaikan dengan baik.

Daftar Pustaka

- [1] Kementrian Keuangan. (2007). Peraturan Menteri Keuangan Nomor 74/PMK.010/2007 Tentang Penyelenggaraan Pertanggung Asuransi pada Lini Usaha asuransi Kendaraan Bermotor. Jakarta: Kemenkeu.
- [2] Kitab Undang-Undang Hukum Dagang (KUHD). Bab IX Tentang Asuransi atau Pertanggung pada Umumnya, Pasal 246.
- [3] Klugman, S. A., Panjer, H. H., & Willmot, G. E. (2012). Loss Models From Data To Decisions. New York: Willey Interscience.
- [4] Undang-Undang Republik Indonesia. (2014). UU RI. Nomor 40 Tahun 2014: Tentang Perasuransian. Jakarta.
- [5] Yulianto Anggi Priliani, Darwis Sutawanir. (2021). Penerapan Metode *K-Nearest Neighbors (kNN)* pada *Bearing*. Jurnal Riset Statistika, 1(1), 10-18.