

Uji Kecocokan Distribusi Binomial Negatif untuk Data Frekuensi Klaim Asuransi Kendaraan Bermotor di Indonesia

Mira Zakiah Rahmah*, Aceng Komarudin Mutaqin

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

*mira.zakiahrahmah@gmail.com, aceng.k.mutaqin@gmail.com

Abstract. The frequency of claims is the number of claims in an insurance policy block during a certain period of time or the number of claims made by the policyholder during the insurance period. In this research, we will discuss the negative binomial distribution modeling on the frequency data of motor vehicle insurance claims in Indonesia. The negative binomial discrete distribution is a Poisson-Gamma mixed distribution which contains two parameters, namely $r > 0$ and $\tau > 0$. Distribution suitability testing can be done using the fit test on the claim frequency data is the chi-square test. The data used is secondary data recorded by the company PT. XYZ in 2014, this data contains data on the frequency of claims for partial loss of policyholders for category 4 motor vehicle insurance products, namely non-bus and non-truck vehicles (all sum insured) region 1 (Sumatra and Islands surrounding). Based on the results of research to model the claim frequency data, it is stated that the negative binomial distribution with parameter estimates of $\hat{r} = 0,3827$ and $\hat{\tau} = 1,9011$ is suitable for modeling of motor vehicle insurance claim frequency data for category 4 region 1 in Indonesia.

Keywords: Claim frequency, negative binomial distribution, chi-square test and partial loss.

Abstrak. Frekuensi klaim merupakan jumlah klaim dalam satu blok polis asuransi selama periode waktu tertentu atau banyaknya klaim yang dilakukan oleh pemegang polis selama masa asuransinya. Dalam penelitian ini akan dibahas pemodelan distribusi binomial negatif pada data frekuensi klaim asuransi kendaraan bermotor di Indonesia. Distribusi diskrit binomial negatif merupakan distribusi campuran Poisson-Gamma yang mengandung dua parameter yaitu $r > 0$ dan $\tau > 0$. Pengujian kecocokan distribusi dapat dilakukan menggunakan uji kecocokan pada data frekuensi klaim adalah uji chi-kuadrat. Data yang digunakan adalah data sekunder hasil pencatatan perusahaan PT.XYZ tahun 2014, data tersebut berisi data frekuensi klaim *partial loss* pemegang polis untuk produk asuransi kendaraan bermotor kategori 4 yaitu kendaraan non bus dan non truk (semua uang pertanggungan) wilayah 1 (Sumatera dan Kepulauan di sekitarnya). Berdasarkan hasil penelitian untuk memodelkan data frekuensi klaim dinyatakan bahwa distribusi binomial negatif dengan taksiran parameter sebesar $\hat{r} = 0,3827$ dan $\hat{\tau} = 1,9011$ cocok untuk memodelkan data frekuensi klaim asuransi kendaraan bermotor kategori 4 wilayah 1 di Indonesia.

Kata Kunci: Frekuensi klaim, distribusi binomial negatif, chi-square dan *partial loss*.

1. Pendahuluan

Seiring perkembangan zaman yang semakin modern memberikan dampak positif bagi berkembangnya industri-industri di Indonesia salah satunya industri otomotif. Laju pertumbuhan paling menonjol dari industri otomotif adalah kendaraan bermotor. Seiring dengan beragam dan banyaknya kendaraan bermotor yang beredar telah menimbulkan semakin padatnya kondisi lalu lintas dan risiko yang harus dihadapi oleh manusia juga semakin kompleks. Risiko yang mungkin terjadi pada kendaraan bermotor seperti kecelakaan dan kehilangan kendaraan bermotor akibat berbagai sebab. Di Indonesia kecelakaan lalu lintas selalu terjadi dalam setiap tahunnya. Badan Pusat Statistik menyatakan bahwa pada tahun 2017 jumlah kecelakaan yang terjadi sebanyak 103.228 kasus kecelakaan dengan besar kerugian mencapai Rp. 215 miliar (www.bps.go.id). Maka diperlukan jaminan untuk mengendalikan masalah kerugian yang akan terjadi yaitu asuransi.

Asuransi atau pertanggungan menurut pasal 246 Kitab Undang-undang Hukum Dagang (KUHD) adalah suatu perjanjian dengan mana seorang penanggung mengikatkan diri kepada seorang tertanggung, dengan menerima suatu premi, untuk penggantian kepadanya karena sebuah kerusakan atau kehilangan keuntungan yang diharapkan yang mungkin akan dideritanya karena suatu peristiwa yang tidak tentu.

Secara garis besar asuransi dibagi menjadi dua macam, yaitu asuransi jiwa dan asuransi kerugian. Menurut Undang-Undang Republik Indonesia No. 2 Tahun 1992 asuransi jiwa memberikan jasa dalam penanggulangan risiko yang dikaitkan dengan hidup atau meninggalnya seseorang yang dipertanggungjawabkan, sedangkan asuransi kerugian memberikan jasa dalam penanggulangan risiko atas kerugian, kehilangan manfaat, dan tanggung jawab hukum kepada pihak ketiga yang timbul dari peristiwa yang tidak pasti, dimana salah satu produk dari asuransi kerugian adalah asuransi kendaraan bermotor.

Asuransi kendaraan bermotor banyak diminati konsumen karena memberikan pertanggungan atas kerugian atau berkurangnya nilai secara finansial atas obyek pertanggungan kendaraan bermotor. Sejalan dengan terus bertambahnya jumlah kendaraan yang ada, memiliki asuransi kendaraan bermotor merupakan suatu kewajiban bagi para pemilik kendaraan, apalagi nilai kecelakaan di Indonesia juga termasuk tinggi. Salah satu aktifitas penting pada asuransi kendaraan bermotor adalah menghitung besar premi yang akan dibebankan pada pemegang polis (Grize, 2014).

Secara statistik, besar premi dapat dihitung dengan terlebih dahulu memodelkan data besar klaim (*claim severity*) dan data frekuensi klaim (*claim frequency*) pemegang polis. Besar klaim dimodelkan oleh distribusi-distribusi kontinu seperti lognormal, gamma, eksponensial, Weibull dan Pareto. Sedangkan frekuensi klaim dimodelkan oleh distribusi-distribusi diskrit standar seperti distribusi Poisson, geometrik dan binomial negatif.

Pada kenyataannya, sering ditemukan data cacahan dengan suatu kondisi dimana varians yang lebih besar dari seharusnya, keadaan ini dikenal dengan overdispersi. Saat terjadi overdispersi, asumsi kesamaan rata-rata dan varians pada distribusi Poisson dilanggar, sehingga perlu dicari distribusi lain yang dapat digunakan untuk menganalisis data cacahan saat terjadi overdispersi. Salah satu distribusi yang dapat digunakan untuk mengatasi masalah overdispersi adalah distribusi binomial negatif yang merupakan distribusi campuran Poisson-Gamma (Tira, 2017). Berdasarkan uraian diatas akan digunakan distribusi binomial negatif untuk memodelkan data frekuensi klaim untuk data asuransi kendaraan bermotor di Indonesia.

2. Landasan Teori

Asuransi

Menurut Pasal I ayat 1 Undang-undang Republik Indonesia No. 40 Tahun 2014, menyebutkan bahwa Asuransi adalah perjanjian antara dua pihak, yaitu Perusahaan Asuransi dan Pemegang Polis yang menjadi dasar bagi penerimaan premi oleh perusahaan asuransi sebagai imbalan untuk memberikan penggantian kepada tertanggung atau pemegang polis karena kerugian, kerusakan, biaya yang timbul, kehilangan keuntungan atau tanggung jawab hukum kepada pihak ketiga yang mungkin diderita tertanggung atau pemegang polis karena terjadinya suatu peristiwa yang tidak pasti atau memberikan pembayaran yang didasarkan pada meninggalnya tertanggung atau pembayaran yang didasarkan pada hidupnya tertanggung dengan manfaat yang besarnya telah ditetapkan dan/atau didasarkan pada hasil pengelolaan dana (Rusman, 2018).

Asuransi Kendaraan Bermotor

Menurut otoritas jasa keuangan asuransi kendaraan bermotor adalah asuransi yang memberikan manfaat berupa pemberian ganti rugi atas kerugian dan atau kerusakan pada kendaraan bermotor yang disebabkan, antara lain tabrakan, terperosok, perbuatan jahat, kebakaran dan lain-lain.

Frekuensi Klaim

Frekuensi klaim merupakan jumlah klaim dalam satu blok polis asuransi selama periode waktu tertentu atau banyaknya klaim yang dilakukan oleh pemegang polis selama masa asuransinya. Misalkan tertanggung A mengajukan klaim sebanyak 3 kali selama masa asuransinya. Dengan demikian frekuensi klaim tertanggung A adalah 3. Beberapa distribusi diskrit yang biasa digunakan dalam memodelkan distribusi frekuensi klaim diantaranya adalah Poisson dan binomial negatif (Ayu, 2017).

Distribusi Binomial Negatif

Peubah acak diskrit N dikatakan berdistribusi binomial negatif dengan parameter $r > 0$ dan $\tau > 0$ apabila fungsi peluangnya sebagai berikut:

$$P(N = k) = p_k = \binom{k+r-1}{k} \left(\frac{\tau}{1+\tau}\right)^r \left(\frac{1}{1+\tau}\right)^k; k = 0, 1, 2, \dots$$

$$= \frac{\Gamma(k+r)}{\Gamma(k+1)\Gamma(r)} \left(\frac{\tau}{1+\tau}\right)^r \left(\frac{1}{1+\tau}\right)^k, \text{ untuk } k = 0, 1, 2, \dots \quad (1)$$

Ekspektasi dan varians dari distribusi binomial negatif masing-masing adalah:

$$E(N) = \frac{r}{\tau},$$

$$Var(N) = \frac{r}{\tau} \left(1 + \frac{1}{\tau}\right).$$

Misalkan k_1, k_2, \dots, k_n adalah suatu sampel acak berukuran n dari distribusi binomial negatif dengan parameter r dan τ , dengan nilai dari sampel acak tersebut k_1, k_2, \dots, k_n . Penaksir kemungkinan maksimum untuk r dan τ dari distribusi binomial negatif tidak dapat diselesaikan secara analitik, namun dapat ditaksir dengan menggunakan metode numerik. Salah satu metode numerik tersebut adalah Newton Rapshon. Taksiran parameter r dan τ adalah solusi dari 2 persamaan berikut:

$$n \ln \left(\frac{1+\tau}{\tau}\right) = \sum_{i=1}^n \sum_{j=0}^{k_i-1} \frac{1}{r+j}, \quad (2)$$

$$\tau = \frac{r}{\bar{k}}. \quad (3)$$

Perangkat lunak Matlab menyediakan fungsi untuk menghitung taksiran kemungkinan maksimum dari parameter distribusi binomial negatif di atas.

Uji Kecocokan Distribusi

Uji kecocokan distribusi adalah suatu pengujian hipotesis statistik yang digunakan untuk mengetahui apakah x_1, x_2, \dots, x_n adalah nilai dari sampel acak X_1, X_2, \dots, X_n yang berasal dari distribusi dengan fungsi distribusi $F(\cdot)$. Pada pengujian kecocokan distribusi dapat digunakan hipotesis sebagai berikut:

H_0 : x_1, x_2, \dots, x_n merupakan nilai dari sampel acak yang berdistribusi $F(\cdot)$.

H_1 : x_1, x_2, \dots, x_n merupakan nilai dari sampel acak yang berdistribusi bukan $F(\cdot)$.

Terdapat sejumlah alternatif untuk mengukur kecocokan distribusi, salah satunya adalah uji kecocokan chi-kuadrat yang dapat digunakan untuk menguji kecocokan distribusi untuk data diskrit dan kontinu. Statistik uji untuk menguji kecocokan chi-kuadrat untuk kasus data diskrit, yaitu:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \quad (4)$$

Dimana O_i adalah banyaknya pengamatan untuk kategori i , E_i adalah nilai harapan untuk kategori i . Untuk mendapatkan nilai E_i dapat dihitung dengan persamaan berikut:

$$E_i = np_x, x = 0, 1, 2, \dots \quad (5)$$

Nilai kritis dihitung dari distribusi chi-kuadrat dengan derajat bebas $k - p - 1$, dimana k adalah banyaknya kategori dan p adalah banyaknya parameter dari distribusi diskrit tersebut. Dengan demikian kriteria pengujianya adalah tolak hipotesis nol jika statistik uji chi-kuadrat lebih besar dari nilai kuantil distribusi chi-kuadrat pada taraf nyata α dan derajat bebas $k - p - 1$ atau $\chi^2 \geq \chi^2_{(k-p-1)(1-\alpha)}$. Jika ada kategori yang nilai harapannya (E_i) kurang dari 5, maka kategori-kategori tersebut digabungkan untuk menghasilkan kategori baru yang mempunyai nilai harapan (E_i) lebih besar sama dengan 5.

3. Hasil Penelitian dan Pembahasan

Data yang digunakan adalah data sekunder hasil pencatatan yang diperoleh dari perusahaan asuransi umum PT. XYZ tahun 2014. Data tersebut berisi data frekuensi klaim pemegang polis untuk produk asuransi kendaraan bermotor *comprehensive* di perusahaan asuransi PT. XYZ. Data yang akan dipakai untuk keperluan aplikasi adalah data frekuensi klaim dan data besar klaim pemegang polis asuransi kendaraan bermotor untuk kategori 4 yaitu kendaraan non bus dan non truk (semua uang pertanggungan) di wilayah 1 (Sumatera dan Kepulauan di sekitarnya).

Uji Kecocokan Distribusi Binomial Negatif

Hipotesis uji kecocokan distribusi binomial negatif adalah:

H_0 : Data frekuensi klaim asuransi kendaraan bermotor kategori 4 wilayah 1 di Indonesia berasal dari populasi yang berdistribusi binomial negatif.

H_1 : Data frekuensi klaim asuransi kendaraan bermotor kategori 4 wilayah 1 di Indonesia bukan berasal dari populasi yang berdistribusi binomial negatif.

Langkah selanjutnya adalah menghitung taksiran parameter distribusi binomial negatif dengan bantuan perangkat lunak Matlab R2017b. Hasil taksiran parameter distribusi binomial negatifnya adalah $\hat{r} = 0,3827$ dan $\hat{t} = 1,9011$.

Tabel 1. Nilai-Nilai yang Dibutuhkan untuk Perhitungan Statistik Uji

Frekuensi Klaim (k)	Banyaknya Tertanggung (n_k)	Peluang Terjadinya Klaim (p_k)	Nilai Harapan Terjadinya Klaim (np_k)	$\frac{(n_k - np_k)^2}{np_k}$
(1)	(2)	(3)	(4)	(5)
0	520	0,8507	519,7777	0,0001
1	67	0,1122	68,5542	0,0352

2	18	0,0267	16,3137	0,1743
≥ 3	6	0,0104	6,3554	0,0198
Jumlah	611	1	611	0,2294

Nilai statistik uji Chi-kuadratnya ada dalam Tabel 1 kolom (5) baris terakhir, yaitu 0,2294. Dengan taraf nyata 5%, nilai kuantil distribusi Chi-kuadrat dengan derajat bebas $1(4 - 2 - 1)$, adalah 3,84. Terlihat bahwa nilai statistik ujinya lebih kecil dibandingkan dengan kuantilnya ($0,2294 < 3,84$). Dengan demikian hipotesis nol diterima dan disimpulkan bahwa data frekuensi klaim asuransi kendaraan bermotor kategori 4 wilayah 1 di Indonesia berasal dari populasi yang berdistribusi binomial negatif.

4. Kesimpulan

Dalam makalah ini terbukti jika distribusi binomial negatif dengan taksiran parameter distribusi binomial negatifnya adalah $\hat{r} = 0,3827$ dan $\hat{t} = 1,9011$ cocok digunakan untuk memodelkan data frekuensi klaim asuransi kendaraan bermotor kategori 4 wilayah 1 PT. XYZ tahun 2014 di Indonesia.

5. Saran

Disarankan kepada peneliti lain untuk menggunakan distribusi diskrit lain untuk dapat memodelkan data frekuensi klaim kendaraan bermotor di Indonesia.

Daftar Pustaka

- [1] Grize, Y-L. (2014). *Applications of Statistics in The Field of General Insurance: An Overview*. International Statistical Review, 1-21
- [2] Rusman, I.Y. S. (2018). *Pengantar Asuransi*. Jakarta: ACA Asuransi.
- [3] Undang-Undang Republik Indonesia. (2014). UU RI. Nomor 40 Tahun 2014: *Tentang Perasuransian*. Jakarta.
- [4] Nazmi, N. (2019). *Pemodelan Distribusi Binomial Negatif Poisson-Lindley Diboboti Pada Data Frekuensi Klaim Asuransi Kendaraan Bermotor Di Indonesia*. Skripsi S1 Jurusan Statistika Universitas Islam Bandung.
- [5] Tira, dkk. (2017). *Distribusi Binomial Negatif-Lindley pada Data Frekuensi Klaim Asuransi Kendaraan Bermotor di Indonesia*. Vol. 3, No. 02, ISSN :2460-2456.