

# Pemodelan Klaim Asuransi Kendaraan Bermotor dengan Regresi ZAIG

YULIA RESTI<sup>1,2</sup>, NORISZURA ISMAIL<sup>2</sup> DAN SAIFUL HAFIZAH JAAMAN<sup>2</sup>

<sup>1</sup>Jurusan Matematika, FMIPA, Universitas Sriwijaya

Jl. Raya Palembang-Prabumulih Km.32 Inderalaya 30662, Sumsel, Indonesia

<sup>2</sup>Pusat Pengajian Sains Matematik, FST, Universiti Kebangsaan Malaysia

43600 UKM Bangi Selangor Darul Ehsan, Malaysia

Email: farras\_yani@yahoo.com

## ABSTRACT

In motor insurance pricing based on risk of policyholder, modeling claim is the most important step. The modeling includes two main models there are model which relates to event of claims and model the cost of claims submitted to insurance companies. Most studies modeling the cost of claims involving only the amount of claims which are positive, i.e. when an accident happens and then the policyholder filed a claim with the claims cost is greater than zero. In one period of insurance, there're policyholders who have not had an accident and there're policyholders who had an accident but does not have claim, in this case is said to the claims cost is zero. This paper investigate the implementation ZAIG (Zero Adjusted Inverse Gaussian) regression on the model of automobile insurance claims that involve the cost of claims that zero and positive use data supported by Insurance Services Malaysia (ISM) Berhad. By regression ZAIG note that both the event and the average of claim cost significantly affected by the premium.

keywords: insurance, premium, event of claim, cost of claim, ZAIG

## 1. PENDAHULUAN

Pemodelan klaim merupakan tahapan penting dalam menentukan premi suatu produk asuransi. Premi adalah harga suatu produk asuransi. Premi yang layak (*reasonable*) atau kadang juga disebut premi yang wajar (*fair*) adalah premi yang dihitung berdasarkan risiko pemegang polis. Pemegang polis yang mempunyai risiko tinggi akan membayar premi yang tinggi demikian pula pemegang polis yang mempunyai risiko rendah akan membayar premi yang rendah. Dengan melibatkan risiko pemegang polis, perusahaan asuransi tidak hanya dapat memperkirakan besarnya biaya klaim yang harus dipersiapkan ketika terjadi klaim melainkan juga dapat memperkirakan keuntungan yang akan diperoleh. Pada dasarnya pemodelan klaim pada asuransi kendaraan bermotor mencakup dua model utama yaitu pertama, model yang berkaitan dengan kejadian klaim (frekuensi) dan kedua, model besarnya biaya klaim yang diajukan ke perusahaan asuransi. Dalam literatur aktuaria, sejumlah model telah digunakan untuk memodelkan klaim asuransi kendaraan bermotor, baik frekuensi maupun model besarnya biaya klaim. Untuk model frekuensi regresi Poisson; Aitkin et al. 1990, Renshaw 1994, regresi quasi-Poisson; McCullagh & Nelder 1989, Brockman & Wright 1992, regresi binomial negatif; Cameron & Trivedi 1986, Lawless 1987, regresi generalized Poisson; Consul & Famoye 1992, Wang & Famoye 1997, Ismail & Jemain 2007. Untuk model besarnya biaya klaim regresi gamma; McCullagh & Nelder 1989, Brockman & Wright 1992, Renshaw 1994, Ismail & Jemain 2006, Venter 2007, Resti et al 2009a, 2009b, 2009c & 2010a, 2010b, regresi inverse Gaussian; Venter 2007, Resti et al 2010a, 2010b. Semua kajian tersebut memodelkan kejadian klaim tersendiri dan besarnya biaya klaim tersendiri dan hanya melibatkan besarnya biaya klaim yang positif. Biaya klaim dikatakan positif ketika suatu kecelakaan terjadi dan kemudian pemegang polis mengajukan klaim dengan biaya klaim lebih besar dari nol. Dalam satu periode asuransi, ada pemegang polis yang tidak mengalami kecelakaan dan ada pemegang polis yang mengalami kecelakaan namun tidak mengajukan klaim, dalam hal ini dikatakan besarnya biaya klaim adalah nol. Jorgensen dan de Souza (1994), Smyth dan Jorgensen (2002), Wuthrich (2003), dan Peter dan kawan-kawan (2008) menggunakan model Tweedie untuk mengakomodasi kedua biaya klaim ini-biaya klaim positif dan biaya klaim nol, sedangkan Heller dan kawan-kawan (2006) menggunakan model *Zero Adjusted Inverse Gaussian* (ZAIG). Pada kedua model ini, baik Tweedie maupun ZAIG, tidak

memodelkan kejadian klaim tersendiri dan besarnya biaya klaim tersendiri, melainkan sebagai satu model. Kelebihan model ZAIG dibandingkan dengan model tweedie ialah model peluang tidak terjadi klaim dapat dimodelkan sebagai fungsi peubah penjelas dari faktor klaim. Jadi model ZAIG merupakan model campuran antara model diskrit dan model kontinu. Model diskrit mewakili model kejadian klaim sedangkan model kontinu mewakili model besarnya biaya klaim. Tujuan artikel ini adalah mengkaji pemodelan klaim asuransi kendaraan bermotor dengan regresi ZAIG. Implementasi regresi ZAIG yang merupakan model campuran antara model diskrit dan model kontinu pada pemodelan klaim asuransi kendaraan bermotor menggunakan data polis dan data klaim yang diberikan oleh *Insurance Service Malaysia (ISM)* Berhad akan mengkaver semua risiko pemegang polis baik yang tidak mengalami kecelakaan, mengalami kecelakaan dan tidak mengajukan klaim maupun yang mengalami kecelakaan dan mengajukan klaim.

## 2. BAHAN DAN METODE

### 2.1 Ringkasan Data yang Digunakan

Kajian pada artikel ini menggunakan data polis dan data klaim asuransi kendaraan bermotor tahun 2000-2003 yang diperoleh dari ISM Berhad. Sampel data sebanyak 572,627 polis di mana banyaknya klaim 9.17%. Data ini diklasifikasi kepada faktor dan kelas risiko. Ada enam faktor risiko yang dipertimbangkan dalam kajian ini yaitu empat faktor binari dan dua faktor numerik. Empat faktor binari tersebut adalah jenis perlindungan (komprehensif dan bukan-komprehensif), produk kendaraan (lokal dan asing), kapasitas kendaraan (0-1000cc, 1001-1300cc, 1301-1500cc, 1500-1800cc dan 1800+cc) dan umur kendaraan (0-1 tahun, 2-3 tahun, 4-5 tahun, 6-7 tahun dan 8+ tahun), sedangkan faktor numerik adalah premi dan nilai kendaraan yang diasuransikan, sehingga total ada 100 kelas risiko yang dilibatkan dalam kajian ini.

### 2.2 Zero Adjusted Inverse Gaussian

Misalkan  $\pi_i$  dan  $y_i$  masing-masing mewakili peluang terjadinya klaim dan besarnya biaya klaim yang diajukan pemegang polis dalam kelas risiko ke- $i$ ,  $i = 1, 2, \dots, m$ , maka fungsi kepadatan peluang ZAIG (Stasinopoulos, 2006) didefinisikan sebagai

$$f(y_i | \mu_i, \sigma_i^2, \pi_i) = 1 - \pi_i \quad ; y_i = 0 \tag{1}$$

$$= \pi_i \frac{1}{\sqrt{2\pi y_i^3 \sigma_i^2}} \exp \left[ -\frac{1}{2\sigma_i^2 y_i^3 \pi_i} (y_i - \mu_i)^2 \right] \quad ; y_i > 0$$

dengan  $E(y_i) = \pi_i \mu_i$  dan  $Var(y_i) = \pi_i \mu_i^2 (1 - \pi_i + \mu_i \sigma_i^2)$ .

### 2.3 Fungsi Penghubung

Fungsi penghubung (*link function*) untuk masing-masing peluang terjadinya klaim  $\pi_i$ , rata-rata besarnya biaya klaim  $\mu_i$  dan dispersi rata-rata besarnya biaya klaim  $\sigma_i^2$  berturut-turut adalah log, log dan logit, yaitu

$$\log \left( \frac{\pi_i}{1 - \pi_i} \right) = \mathbf{X}_1' \beta_\pi + f_\pi(\mathbf{X}_2 \pi_2) \tag{2}$$

$$\log(\mu_i) = \mathbf{X}_1' \beta_\mu + f_\mu(\mathbf{X}_2 \mu_2) \tag{3}$$

$$\log(\sigma_i^2) = \mathbf{X}_1' \beta_\sigma + f_\sigma(\mathbf{X}_2 \sigma_2) \tag{4}$$

di mana  $\mathbf{X}_1$  merupakan vektor dari peubah penjelas,  $\beta$  merupakan vektor dari parameter dan  $f(\mathbf{X}_2)$  merupakan fungsi *smoothing spline*. Koefisien-koefisien dari parameter-parameter (2) hingga (4) dapat diperoleh menggunakan metode (*penalized*) maksimum likelihood.

2.4 AIC dan SBC

Kriteria model pencocokan regresi terbaik yang digunakan dalam kajian ini adalah Akaike Information Criteria (AIC) dan Schwarz Bayesian Criteria (SBC). AIC dan SBC masing-masing didefinisikan sebagai

$$AIC = -2.\ln(L) + 2k \tag{5}$$

$$SBC = -2.\ln(L) + k.\ln(n) \tag{6}$$

di mana k adalah banyaknya parameter model dan n adalah banyaknya observasi.

3. HASIL DAN PEMBAHASAN

Tabel 1. Model Klaim ZAIG terbaik

	Parameter	Anggaran	Std. Error	Statistik- t	Nilai-p
Kejadian klaim	Pintasan	-5.24	1.91	-2.75	0.01
	x1 (bukan komprehensif)	1.08	0.88	1.24	0.22
	x2 (luar negara)	1.24	0.75	1.65	0.10
	x7 (0-1 tahun)	6.46	1.83	3.53	<0.01
	x8 (2-3 tahun)	4.79	1.85	2.59	0.01
	x9 (4-5 tahun)	4.99	1.45	3.44	<0.01
	x10 (6-7 tahun)	3.27	1.32	2.49	0.02
	x11 (premium)	-2.88	0.95	-3.05	<0.01
Rata-rata besarnya biaya klaim	Pintasan	0.88	0.21	4.23	<0.01
	x1 (bukan komprehensif)	1.05	0.47	2.23	0.03
	x2 (luar negara)	-0.36	0.17	-2.09	0.04
	x7 (0-1 tahun)	-0.37	0.29	-1.27	0.21
	x8 (2-3 tahun)	0.38	0.20	1.86	0.07
	x9 (4-5 tahun)	0.23	0.20	1.14	0.26
	x10 (6-7 tahun)	1.38	0.29	4.83	<0.01
	x11 (premium)	1.44	0.49	2.94	<0.01
	x12 (nilai kendaraan)	-0.02	0.01	-2.86	<0.01
Dispersi rata-rata besarnya biaya klaim	Pintasan	-1.04	0.22	-4.65	<0.01
	x1 (bukan komprehensif)	1.22	0.19	6.50	<0.01
	x2 (luar negara)	-0.67	0.19	-3.42	<0.01
	x3 (0-1000cc)	0.48	0.30	1.63	0.11
	x4 (1001-1300cc)	-0.76	0.34	-2.21	0.03
	x5 (1501-1800cc)	0.33	0.27	1.25	0.21
	x6 (1800+cc)	0.32	0.28	1.14	0.27
	AIC	491.70			
	SBC	554.23			

Model klaim ZAIG terbaik berdasarkan AIC dan SBC adalah

$$\begin{aligned} \log(\mu) &= \text{jenis perlindungan} + \text{buatan kendaraan} + \text{umur kendaraan} \\ &\quad + \text{smoothing spline (premi)} \\ \log(\sigma) &= \text{jenis perlindungan} + \text{buatan kendaraan} + \text{umur kendaraan} \\ &\quad + \text{smoothing spline (premium + nilai kendaraan)} \\ \log(\tau) &= \text{jenis perlindungan} + \text{buatan kendaraan} + \text{kapasitas kendaraan} \end{aligned}$$

Koefisien dari parameter model tersebut ditampilkan dalam tabel 1. Tabel tersebut menunjukkan bahwa kejadian klaim dipengaruhi oleh jenis perlindungan bukan komprehensif, produk kendaraan asing, umur kendaraan dan premi, namun yang signifikan hanya umur kendaraan (0-1 tahun dan 4-5 tahun) dan premi, manakala rata-rata besarnya biaya klaim dipengaruhi secara signifikan oleh umur kendaraan (6-7 tahun), premi, dan nilai kendaraan yang diasuransikan. Pada model dispersi rata-rata besarnya biaya klaim, parameter yang mempunyai risiko lebih berpengaruh adalah perlindungan bukan komprehensif dan produk kendaraan asing. Jadi pada model ini, baik kejadian klaim maupun rata-rata besarnya biaya klaim kedua-duanya dipengaruhi secara signifikan oleh premi, manakala kapasitas kendaraan (cc) mempunyai pengaruh yang tidak signifikan terhadap model besarnya biaya klaim.

#### 4. KESIMPULAN

Implementasi regresi ZAIG pada pemodelan klaim yang mengkaver kedua-dua jenis biaya klaim, diketahui bahwa baik kejadian klaim maupun rata-rata besarnya biaya klaim dipengaruhi secara signifikan oleh premi. Pada kejadian klaim, selain premi juga dipengaruhi secara signifikan oleh umur kendaraan (0-1 tahun dan 4-5 tahun). Pada rata-rata besarnya biaya klaim, selain premi juga dipengaruhi secara signifikan oleh nilai kendaraan dan umur kendaraan (hanya 6-7 tahun). Selain itu model dispersi rata-rata besarnya klaim dan model peluang terjadinya klaim dapat dimodelkan secara terpisah dari model rata-rata besarnya klaim. Jadi, tidak hanya model rata-rata besarnya klaim yang dapat dimodelkan sebagai fungsi peubah penjelas dari faktor klaim, tetapi juga model dispersinya dan model peluang terjadinya klaim.

#### Ucapan Terima Kasih

Ucapan terima kasih disampaikan kepada Insurances Malaysia (ISM) Berhard atas bantuan datanya, juga kepada Program Sains Aktuari dan Program Statistik, Pusat Pengajian Sains Matematik, Fakultas Sains dan Teknologi, Universiti Kebangsaan Malaysia serta Jurusan Matematika, FMIPA, Universitas Sriwijaya, Sumsel, Indonesia atas dukungannya.

#### Daftar Pustaka

- Aitkin, M., Anderson, D., Francis, B. & Hinde, J., 1990. *Statistical modelling in GLIM*. New York, Oxford University Press.
- Brockman, M.H. & Wright, T.S., 1992. Statistical motor rating: making effective use of your data. *Journal of the Institute of Actuaries* 119(3): 457-543.
- Cheong, P.W., Jemain, A.A., and Ismail, N. 2008. Practice and pricing in Non-Life Insurance: The Malaysian Experience. *Journal of Quality Measurement and Analysis* 4(1): 11-24.
- Coutts, S.M., 1984. Motor insurance rating, an actuarial approach. *Journal of the Institute of Actuaries* 111: 87-148.
- Heller, G.Z., Stasinopoulos, D.M., and Rigby, R.A. 2006. The zero adjusted inverse Gaussian distribution as a model for insurance data. *Proceeding of the International Workshop on Statistical Modelling*, Galway, 226-233.
- Ismail, N. & Jemain, A.A. 2006. A comparison of risk classification methods for claim severity data. *Journal of Modern Applied Statistical Methods* 5(2): 513-528.
- Ismail, N. & Jemain, A.A. 2007a. Handling overdispersion with Negative Binomial and Generalized Poisson regression models. *Casualty Actuarial Society Forum*. Winter: 103-158.
- Ismail, N. & Jemain, A.A. 2007b. Rating factors identification using claim frequency approach: the Malaysian Experience. *ICFAI Journal of Applied Economics* 6(2): 60-77.

- Jorgensen, B., de Souza, M.C.P. 1994. Fitting Tweedie's compound Poisson model to insurance claims data. *Scandinavian Actuarial Journal*, 69-93.
- Lemaire, J., 1985. *Automobile insurance actuarial models*. Boston, Kluwer Academic Publishers Group.
- Peters, G.W., Shevchenko, P.V. dan Wuthrich, M.V. 2008. Model risk in claim reserving within Tweedie's compound Poisson model. *International ASTIN Colloquium 38th, Manchester, UK*.
- Renshaw, A.E., 1994. Modelling the claims process in the presence of covariates. *ASTIN Bulletin* 24(2): 265-285.
- Resti, Y., Ismail, N., and Jaaman, S.H. 2009a. Modelling Dependence of Claim Amount between Different Claim Types using Copula. *Proceeding of the 5<sup>th</sup> International Conference on Mathematics, Statistics, and Their Application (ICMSA)*. 9-11 Juni. Bukit Tinggi, 629-632.
- Resti, Y., Ismail, N., and Jaaman, S.H. 2009b. Modelling Risk Premium of Motor Insurance Using Copula. *Prosiding Seminar Kerjasama ke-2 Universiti Kebangsaan Malaysia-Universitas-Indonesia*. 22-23 Juni. Fakultas Sains dan Teknologi, UKM, Bangi, 644-650.
- Resti, Y., Ismail, N., and Jaaman, S.H. 2009c. Application of t-Copula in Assessing Malaysian Motor Insurance Pricing. *Prosiding Konferensi Matematika se-Asia (AMC) ke-5*. 22-26 Juni. Kuala Lumpur, 421-425.
- Resti, Y., Ismail, N., and Jaaman, S.H. 2010a. Handling the Dependence of Claim Severities with Copula Models. Jurnal telah dikirimkan ke *Journal of Mathematics and Statistics*.
- Resti, Y., Ismail, N., and Jaaman, S.H. 2010b. Application of Copulas for Motor Insurance Claims. Jurnal telah dikirimkan ke *Sains Malaysiana*.
- Smyth, G.K., Jorgensen, B. 2002. Fitting Tweedie's compound Poisson model to insurance claims data: dispersion modelling. *ASTIN Bulletin*, 32(1), 143-157.
- Wuthrich, M.V. 2003. Claim reserving using Tweedie's compound Poisson model. *ASTIN Bulletin*. 33(2): 331-346.

