

MODEL BERBASIS COPULA UNTUK ESTIMASI KERUGIAN AGREGAT DALAM ASURANSI KECELAKAAN LALU LINTAS

Ana Kristiana¹, Yogo Purwono²
Universitas Indonesia^{1,2}
anakristiana94@gmail.com¹

ABSTRAK

Penelitian ini bertujuan untuk mengetahui model berbasis copula untuk estimasi kerugian agregat dalam asuransi kecelakaan lalu lintas. Metode penelitian ini adalah deskriptif kuantitatif. Penentuan model terbaik dan akurasi model ditentukan berdasarkan *Akaike Information Criterion* (AIC), *Root Mean Square Error* (RMSE) terkecil, dan Uji Vuong. Hasil penelitian, terdapat hubungan kebergantungan atau korelasi antara frekuensi klaim dan *severity* klaim, sehingga digunakan copula untuk memodelkan frekuensi klaim yang berdistribusi poison dengan *severity* klaim yang berdistribusi gamma. Simpulan, model *copula clayton* merupakan model terbaik untuk memperkirakan kerugian agregat pada perusahaan asuransi PT XYZ dimasa yang akan datang.

Kata Kunci: Pertanggung, Kehilangan, Korelasi, Kopula

ABSTRACT

This study aims to determine the copula-based model for estimating aggregate losses in traffic accident insurance. This research method is descriptive quantitative. Determination of the best model and the accuracy of the model is determined based on the Akaike Information Criterion (AIC), the smallest Root Mean Square Error (RMSE), and the Vuong Test. The results of the study, there is a dependency relationship or correlation between the frequency of claims and the severity of claims, so that copula is used to model the frequency of claims with a poison distribution and claim severity with a gamma distribution. In conclusion, the clayton copula model is the best model for estimating aggregate losses at the insurance company PT XYZ in the future.

Keywords: Insurance, Loss, Correlation, Copula

PENDAHULUAN

Perkembangan inovasi dibidang transportasi yang cukup pesat dari masa ke masa membawa manfaat yang cukup besar bagi masyarakat untuk mempermudah aktivitas dan mobilitas untuk berpindah dari satu tempat ke tempat lain. Hal tersebut yang mendorong peningkatan pembelian dan penggunaan alat transportasi dari tahun ke tahun, termasuk di Indonesia, berdasarkan data Badan Pusat Statistik (BPS) selama tahun 2011–2018, terjadi peningkatan jumlah kendaraan bermotor yang beroperasi di Indonesia, dengan peningkatan rata-rata sebesar 8,03% per tahun. Melihat pertumbuhan penggunaan kendaraan bermotor di Indonesia yang terus mengalami peningkatan yang cukup signifikan, menyebabkan terjadinya peningkatan risiko kecelakaan yang akan

dialami oleh pengguna alat transportasi kendaraan bermotor di Indonesia. Berdasarkan data BPS tahun 2018, terjadi sekitar 109.215 kejadian kecelakaan lalu lintas di Indonesia dan jika dibandingkan dengan tahun sebelumnya jumlah kejadian kecelakaan di tahun 2018 mengalami peningkatan sebesar 4.888 kejadian kecelakaan atau sebesar 4,7%, yang mengakibatkan sebanyak 173.358 orang yang menjadi korban kecelakaan lalu lintas mulai dari luka ringan sampai dengan korban meninggal dunia. Melihat pertumbuhan jumlah kejadian dan korban kecelakaan lalu lintas yang meningkat tiap tahunnya, Pemerintah Indonesia berupaya menyediakan beberapa peraturan sebagai jaminan perlindungan terhadap korban kecelakaan lalu lintas di Indonesia, yang diatur dalam Undang-Undang (UU) Nomor 33 Tahun 1964 tentang Dana Pertanggungjawaban Wajib Kecelakaan Penumpang dan Nomor 34 Tahun 1964 tentang Dana Kecelakaan Lalu Lintas Jalan.

Jaminan perlindungan terhadap korban kecelakaan lalu lintas diberikan dalam bentuk asuransi kecelakaan lalu lintas oleh Pemerintah Indonesia bagi seluruh warga negaranya yang mengalami kecelakaan lalu lintas di wilayah Negara Kesatuan Republik Indonesia. Melalui asuransi kecelakaan lalu lintas, Pemerintah berupaya melindungi warga negaranya dari kerugian akibat kecelakaan, dimana setiap pengendara kendaraan dan pengguna alat transportasi umum terjamin dari biaya-biaya perawatan rumah sakit yang mungkin timbul sebagai akibat dari kecelakaan serta ahli warisnya dapat memperoleh manfaat santunan apabila korban kecelakaan lalu lintas meninggal dunia, yang besarnya sesuai dengan ketentuan dan Undang-undang yang berlaku.

PT XYZ merupakan perusahaan yang ditunjuk oleh Pemerintah Indonesia untuk melakukan pengelolaan terhadap premi yang dibayarkan oleh pemilik kendaraan bermotor dan penumpang angkutan lalu lintas umum yang ada di Indonesia, untuk selanjutnya dibayarkan kepada korban kecelakaan lalu lintas di masa mendatang. Jenis manfaat yang diberikan kepada korban kecelakaan lalu lintas dibedakan menjadi beberapa jenis yaitu, santunan bagi ahli waris korban yang meninggal dunia, penggantian biaya rawatan rumah sakit bagi korban luka-luka, santunan bagi korban kecelakaan yang mengalami cacat, santunan penggantian biaya penguburan bagi korban kecelakaan yang tidak memiliki ahli waris, serta manfaat penggantian biaya ambulans dan P3K. Mengingat tingkat kecelakaan yang memiliki trend peningkatan yang terjadi pada beberapa tahun terakhir ini, penting bagi PT XYZ untuk melakukan estimasi atau

prediksi mengenai seberapa besar total kerugian klaim akibat kecelakaan lalu lintas yang ditanggung oleh PT XYZ di masa mendatang. Hal ini berguna agar Pemerintah Indoneia melalui PT XYZ dapat senantiasa memberikan perlindungan dasar akibat kecelakaan lalu lintas bagi masyarakat Indonesia.

PT XYZ merupakan perusahaan yang ditunjuk oleh Pemerintah Indonesia untuk melakukan pengelolaan terhadap premi yang dibayarkan oleh pemilik kendaraan bermotor dan penumpang angkutan lalu lintas umum yang ada di Indonesia, untuk selanjutnya dibayarkan kepada korban kecelakaan lalu lintas di masa mendatang. Jenis manfaat yang diberikan kepada korban kecelakaan lalu lintas dibedakan menjadi beberapa jenis yaitu, santunan bagi ahli waris korban yang meninggal dunia, penggantian biaya rawatan rumah sakit bagi korban luka-luka, santunan bagi korban kecelakaan yang mengalami cacat, santunan penggantian biaya penguburan bagi korban kecelakaan yang tidak memiliki ahli waris, serta manfaat penggantian biaya ambulans dan P3K. Mengingat tingkat kecelakaan yang memiliki trend peningkatan yang terjadi pada beberapa tahun terakhir ini, penting bagi PT XYZ untuk melakukan estimasi atau prediksi mengenai seberapa besar total kerugian klaim akibat kecelakaan lalu lintas yang ditanggung oleh PT XYZ di masa mendatang. Hal ini berguna agar Pemerintah Indoneia melalui PT XYZ dapat senantiasa memberikan perlindungan dasar akibat kecelakaan lalu lintas bagi masyarakat Indonesia.

Salah satu metode analisis statistika yang digunakan untuk menggabungkan dua distribusi data yang berbeda namun memiliki hubungan ketergantungan adalah metode copula. Metode copula merupakan suatu metode yang bersifat fleksibel karena tidak mensyaratkan adanya asumsi normalitas dan dapat menggabungkan beberapa fungsi distribusi marginal ke dalam fungsi distribusi gabungan dengan asumsi terdapat keterkaitan antara distribusi marginalnya. Oleh karena itu, metode copula digunakan untuk mengetahui keterkaitan antara *severity* klaim dan frekuensi klaim sehingga perusahaan asuransi dapat mengestimasi total klaim dan cadangan teknis yang harus disiapkan. Copula memiliki berbagai macam bentuk, namun copula yang paling populer untuk memodelkan hubungan ketergantungan antara frekuensi klaim dan *severity* klaim adalah copula dari keluarga Archimedean seperti copula Gumbel, Frank, dan Clayton.

Beberapa penelitian terdahulu tentang model copula telah dilakukan dan diterapkan pada bidang keuangan, asuransi, dan studi mengenai lingkungan seperti

bidang klimatologi. Penelitian terdahulu yang telah dilakukan antara lain oleh Azhar & Fauzi (2014), yang membahas tentang hubungan ketergantungan antara *severity* klaim dengan jenis klaim asuransi pada asuransi kendaraan bermotor di Malaysia, Czado et al., (2012) yang membahas tentang hubungan antara *severity* klaim dan frekuensi klaim dengan menggunakan model regresi gabungan berbasis Krämer et al., (2013) yang membahas mengenai model regresi gabungan untuk *severity* klaim dan frekuensi klaim dengan menggunakan bivariat copula yang bertujuan untuk mengakomodasi hubungan antara *severity* klaim dan frekuensi klaim, sehingga diperoleh estimasi terbaik mengenai total kerugian klaim asuransi mobil di Jerman. Berdasarkan latar belakang di atas maka dilakukan kajian untuk mengetahui model berbasis copula untuk estimasi kerugian agregat dalam asuransi kecelakaan lalu lintas.

KAJIAN TEORI

Distribusi Frekuensi Klaim

Distribusi frekuensi klaim merupakan distribusi yang menunjukkan jumlah atau frekuensi terjadinya suatu kerugian tanpa melihat nilai kerugiannya dalam periode waktu tertentu. Frekuensi klaim dimodelkan dengan menggunakan distribusi diskrit. Model distribusi diskrit yang paling sering digunakan untuk data frekuensi klaim, antara lain adalah distribusi poisson, uniform, dan binomial negatif (Muslich, 2007).

Distribusi Severity Klaim

Distribusi *severity* klaim merupakan distribusi yang memperlihatkan rata-rata besaran nilai kerugian dalam periode waktu tertentu. *Severity* klaim (besarnya klaim) dimodelkan dengan menggunakan distribusi kontinu, diantaranya adalah distribusi lognormal, eksponensial, pareto, invers Gaussian, Weibull, dan Gamma (Muslich, 2007).

Aggregate Loss

Aggregate Loss adalah total kerugian yang dialami oleh pemegang polis yang harus ditanggung oleh perusahaan asuransi dalam periode waktu tertentu. Untuk memperoleh *aggregate loss* dilakukan dengan mencatat masing-masing besar pembayaran klaim dan menjumlahkan semua klaim tersebut. Secara matematis

aggregate loss model dapat dinyatakan dengan persamaan sebagai berikut (Klugman et al., 2004):

$$S = \sum_{i=0}^N X_i = X_1 + X_2 + \dots + X_N \quad (1)$$

Dimana S adalah *aggregate loss*, X_i adalah besar klaim (*loss severity*) ke- i , dan N adalah banyaknya klaim (*loss frequency*) dalam satu periode. Asumsi–asumsi yang harus diperhatikan pada *aggregate loss* untuk model *collective risk* adalah diberikan $N = n$, peubah acak X_1, X_2, \dots, X_n merupakan peubah acak yang berdistribusi identik dan saling bebas. Diberikan $N = n$, distribusi bersama dari peubah acak X_1, X_2, \dots, X_n tidak bergantung pada nilai n . Distribusi dari peubah acak N tidak bergantung kepada nilai-nilai dari peubah acak X_1, X_2, \dots .

Ukuran Dependensi

Struktur dependensi menggambarkan suatu hubungan keeratan atau korelasi antara suatu variabel random terhadap variabel random lainnya. Salah satu ukuran yang biasa digunakan untuk menggambarkan struktur dependensi antara dua variabel random adalah korelasi *Kendall's Tau*. Korelasi *Kendall's Tau* digunakan sebagai alternatif pengukuran dependensi data selain korelasi linier, karena korelasi *Kendall's Tau* tidak terpengaruh oleh nilai outlier pada data dan antar variabel random tidak disyaratkan harus memiliki hubungan linier.

Misalkan (x_i, y_i) dan (x_j, y_j) merupakan dua observasi yang berasal dari variabel random kontinu (X, Y) . (x_i, y_i) dan (x_j, y_j) dikatakan konkordan jika $(x_i - x_j)(y_i - y_j) > 0$ dan diskordan apabila $(x_i - x_j)(y_i - y_j) < 0$. Misalkan $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ merupakan sampel random bivariat berukuran n dari vektor variabel random kontinu (X, Y) , maka terdapat $\binom{n}{2}$ pasang (x_i, y_i) dan (x_j, y_j) yang berbeda yang bisa diperoleh (baik konkordan maupun diskordan). Misal c merupakan jumlah pasangan konkordan dan d adalah jumlah pasangan yang diskordan, maka perhitungan korelasi *Kendall's Tau* untuk sampel didefinisikan sebagai berikut (Nelsen, 2006):

$$\tau = \frac{c-d}{c+d} = \frac{c-d}{\binom{n}{2}} \quad (2)$$

Pada pemodelan copula, pengukuran dependensi dengan korelasi *Kendall's Tau* dengan parameter copula θ didefinisikan pada Tabel 1, sebagai berikut:

Tabel 1.
Hubungan Parameter Copula θ dengan Korelasi *Kendall's Tau*

Copula	τ	Range τ
Gauss	$\tau = \frac{2}{\pi} \arcsin(\theta)$	$\tau \in \mathbb{R}$
Frank	$1 - \frac{4}{\theta} \left[1 - \frac{1}{\theta} \int_0^\theta \frac{t}{e^t - 1} dt \right]$	$\tau \in \mathbb{R} \setminus \{0\}$
Gumbel	$\tau = \frac{\theta - 1}{\theta}$	$\tau \in [0, \infty[$
Clayton	$\tau = \frac{\theta}{\theta + 2}$	$\tau \in]0, \infty[$

Generalized Linear Models (GLM)

Untuk memperkirakan total kerugian, perusahaan asuransi menggunakan variabel penjelas atau biasa disebut dengan variabel kovariat, seperti faktor-faktor yang mempengaruhi terjadinya klaim, sehingga dapat diperkirakan kerugian yang harus oleh suatu perusahaan asuransi. Pemodelan yang banyak digunakan untuk proses estimasi kerugian klaim adalah pendekatan *Generalized Linear Model* (GLM), yang merupakan perluasan dari model linear untuk membentuk suatu model analisis antara variabel independen atau variabel kovariat dengan variabel dependen atau variabel respon, dimana tidak harus disyaratkan adanya kenormalan dan kehomogenan data pada errornya. GLM memiliki karakteristik menurut Valentovicova (2015) pertama, distribusi untuk variabel dependen berasal dari keluarga eksponensial, dengan fungsi densitas sebagai berikut:

$$f_Y(y; \theta, \phi) \exp \left\{ \frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi) \right\} \quad (3)$$

Dengan nilai mean $E(Y_i) = \mu_i$. Kedua, komponen linier yang didefinisikan sebagai model linier sebagai berikut:

$$\eta_i = \mathbf{x}_i^T \boldsymbol{\beta} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} \quad (4)$$

Dimana η_i merupakan kombinasi linier antara koefisien regresi $\boldsymbol{\beta}$ dengan \mathbf{x}_i yang merupakan vektor kovariat. Fungsi penghubung atau fungsi *link* yang menghubungkan nilai $E(Y_i) = \mu_i$ dengan η_i . Fungsi *link* disebut kanonik apabila $g(\mu_i) = \eta_i$.

Copula

Menurut Czado et al., (2012), copula merupakan metode yang dapat menggabungkan beberapa distribusi marginal menjadi distribusi bersama yang bertujuan untuk mendeteksi dependensi antar variabel acak baik linier maupun tidak linier. Dalam penelitian ini, menggunakan copula bivariat untuk menggabungkan fungsi distribusi marginal dari frekuensi klaim dan *severity* klaim. Dalam Krämer et al., (2013) didefinisikan bahwa, copula bivariat $C: [0,1] \times [0,1] \rightarrow [0,1]$ atau copula berdimensi dua merupakan distribusi kumulatif bivariat pada $[0,1] \times [0,1]$ dengan distribusi uniform sebagai distribusi marginalnya. Dalam kasus copula bivariat, untuk setiap distribusi gabungan H_{XY} dari variabel bivariat (X, Y) dengan distribusi marginal $u = F_X$ dan $v = G_Y$ diperoleh copula dengan persamaan :

$$H_{XY}(x, y) = C(F_X(x), G_Y(y)) = C(u, v) \quad (5)$$

Copula yang digunakan berasal dari keluarga copula *Archimedean* yaitu copula *Clayton*, *Gumbel*, dan *Frank*, dengan alasan bahwa copula jenis ini mudah dikonstruksikan dan sering digunakan dalam penelitian di bidang asuransi. Secara umum, bentuk persamaan copula *Archimedean* sebagai berikut:

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)), 0 \leq u, v \leq 1 \quad (6)$$

dimana φ adalah fungsi pembangkit atau *generator* dari copula C dengan $0 \leq u, v \leq 1$ dan $\varphi(0) = \infty, \varphi(1) = 0$

Copula Clayton

Copula Clayton bivariat memiliki fungsi pembangkit $\varphi(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}$, dengan fungsi densitas bersyarat sebagai berikut:

$$f(y|x) = g(y) \frac{\theta^2 \frac{\Gamma(\frac{1}{\theta}+2)}{\Gamma(\frac{1}{\theta})(u^{-\theta-1})(v^{-\theta-1})(u^{-\theta}+v^{-\theta-1})^{-\frac{1}{\theta}-2}}}{(u^{-\theta}+v^{-\theta-1})^{-\frac{1}{\theta-1}-2} u^{-\theta-1}} \quad (7)$$

Copula Gumbel

Copula Gumbel bivariat memiliki fungsi pembangkit $\varphi(u, v) = \exp\left\{-[(-\ln(u))^\theta + (-\ln(v))^\theta]^{\frac{1}{\theta}}\right\}$, dengan fungsi densitas bersyarat sebagai berikut:

$$f(y|x) = g(y) \frac{c(u,v)u^{-1}v^{-1}(-\ln u)^\theta(-\ln v)^\theta [(-\ln u)^\theta + (-\ln v)^\theta]^{-2+\frac{1}{\theta}} [(-\ln u)^\theta + (-\ln v)^\theta + \theta + 1]}{u^{-1} \exp\left(-((-\ln u)^\theta + (-\ln v)^\theta)^{\frac{1}{\theta}}\right)} \quad (8)$$

Copula Frank

Copula Frank bivariat memiliki fungsi pembangkit $\varphi(u, v) = -\frac{1}{\theta} \ln\left(1 + \frac{(e^{\theta u} - 1)(e^{\theta v} - 1)}{(e^\theta - 1)}\right)$, dengan fungsi densitas bersyarat sebagai berikut:

$$f(y|x) = g(y) \frac{\frac{-\theta(\exp(-\theta)-1)(1+\exp(-\theta(u+v))-1)}{[(\exp(-\theta u)-1)(\exp(-\theta v)-1)+(\exp(-\theta)-1)]^2}}{\frac{\exp(-\theta) \exp(-\theta v-1)}{\exp(\theta(u+1))+\exp(\theta(v+1))-\exp(\theta)-\exp(\theta(u+v))}} \quad (9)$$

Pemilihan Model Copula Terbaik

Penentuan model copula yang tepat untuk mengestimasi besar frekuensi dan *severity* klaim kecelakaan lalu lintas pada perusahaan asuransi PT XYZ, dapat menggunakan perbandingan terkecil nilai perhitungan *Akaike Information Criterion* (AIC) dari masing-masing model copula. Dalam Krämer et al., (2013), persamaan AIC dituliskan sebagai berikut:

$$AIC := -2\ell(\hat{\nu}|\mathbf{x}, \mathbf{y}) + 2DoF(10)$$

dimana *DoF* (*Degrees of Freedom*) merupakan banyaknya parameter yang diestimasi dalam model. Dengan *DoF* untuk model gabungan adalah $p+q+2$, sedangkan untuk model independen adalah $p+q+1$.

Menurut Daraghma (2010), pemilihan model terbaik juga dapat dilakukan dengan menggunakan uji statistik Vuong untuk menentukan model manakah yang memiliki *explanatory power* atau kekuatan penjelasan terhadap variabel dependen yang lebih baik. uji statistik Vuong dituliskan dengan persamaan sebagai berikut :

$$Z_{Vuong} = \frac{[\log(\sigma_w^2) - \log(\sigma_x^2)]}{\left[n^{0,5} \sum_1^n \left(\frac{e_{w,i}^2}{\sigma_w^2} - \frac{e_{x,i}^2}{\sigma_x^2} \right) \right]} \quad (11)$$

dengan $\sigma_x^2 = \sigma_y^2(1 - R_x^2)$ dan $\sigma_w^2 = \sigma_y^2(1 - R_w^2)$. Jika nilai Z_{vuong} bernilai positif atau $Z_{vuong} > 2$ maka model W lebih baik daripada model X. Namun jika Z_{vuong} bernilai negatif atau $Z_{vuong} < -2$ maka model X lebih baik dibandingkan dengan model W.

METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah data sekunder yang berasal dari data harian pembayaran klaim terhadap korban kecelakaan yang dilakukan oleh perusahaan asuransi PT XYZ periode Januari 2017 sampai dengan April 2020. Data harian pembayaran klaim korban kecelakaan meliputi informasi mengenai data diri korban kecelakaan seperti nama, alamat, jenis kelamin, usia, jenis kendaraan penyebab, dan jenis manfaat yang diterima oleh korban kecelakaan. Pengujian dan pemilihan model terbaik menggunakan data pembayaran klaim tahun 2017 sampai dengan tahun 2019 (*data in sample*), sedangkan untuk membandingkan hasil prediksi yang dihasilkan menggunakan data periode Januari 2020 sampai dengan April 2020 (*data out sample*).

Variabel dalam penelitian ini dibagi menjadi dua jenis, yaitu variabel bebas dan variabel tak bebas. Variabel tak bebas yang digunakan dalam melakukan estimasi *total aggregate loss* adalah data frekuensi dan *severity* klaim, sedangkan variabel yang merupakan kovariat atau faktor yang mempengaruhi variabel tak bebas dalam penelitian ini adalah variabel usia dan jenis kendaraan, sebagai berikut.

Tabel 2.
Variabel Penelitian

Variabel	Keterangan	Kategori
Frekuensi	Jumlah klaim kecelakaan lintas yang diajukan	-
<i>Severity</i>	Rata-rata pembayaran klaim pada korban kecelakaan lalu lintas per hari	-
Usia	Usia korban kecelakaan	1 : usia 0 -10 tahun
		2 : usia 11 - 25 tahun
		3 : usia 26 -55 tahun
		4 : > 56 tahun
Kendaraan	Jenis kendaraan penyebab terjadinya kecelakaan	1 : golongan angkutan umum (bus angkutan umum, kapal laut/penyebrangan, kereta api, mobil penumpang angkutan umum, dan pesawat udara)
		2 : golongan non angkutan umum (bulldozer, bus bukan angkutan umum, mobil penumpang bukan angkutan umum, mobil jenazah, mobil pemadam kebakaran, <i>pick up</i> /mobil barang, , sepeda motor truk, dan sejenisnya)

Langkah Analisis

Berikut adalah langkah-langkah untuk mengestimasi kerugian asuransi kecelakaan lalu lintas dengan regresi copula, a) melakukan identifikasi awal terhadap data penelitian dengan statistika deskriptif dari masing-masing data frekuensi klaim dan *severity* klaim; b) menentukan jenis distribusi yang sesuai dengan data frekuensi klaim dan *severity* klaim dengan uji inferensi *Goodness of Fit* melalui uji Kolmogorov Smirnov. Pada penelitian ini digunakan distribusi poisson untuk data frekuensi klaim dan distribusi gamma untuk data *severity* klaim; c) melakukan uji untuk korelasi *Kendall's Tau*, untuk menguji adanya hubungan/keterkaitan antara data frekuensi klaim dan *severity* klaim; d) melakukan permodelan terhadap data frekuensi klaim dan *severity* klaim, dengan variabel usia dan jenis kendaraan sebagai variabel kovariat menggunakan *Generalized Linear Model* (GLM) sebagai berikut:

$$X_i \sim \text{Gamma}(\mu_i, \delta) \text{ dengan } \ln(\mu_i) = \mathbf{r}_i \boldsymbol{\alpha}$$

$$Y_i \sim \text{ZTP}(\lambda_i, \delta) \text{ dengan } \ln(\lambda_i) = \ln(e_i) + \mathbf{r}_i^T \boldsymbol{\beta}$$

dimana \mathbf{r}_i' dan \mathbf{s}_i' adalah variabel kovariat dan e_i adalah waktu pengamatan; e) menggabungkan distribusi marginal dari masing-masing model yang dihasilkan dari data frekuensi klaim dan *severity* klaim menjadi satu model menggunakan regresi copula. Untuk persamaan copula *Archimedean* dengan melibatkan kovariat, maka model persamaan copula *Archimedean* bersyarat adalah sebagai berikut:

$$C(u, v|r) = \varphi^{-1}(\varphi(u|r) + \varphi(v|r))$$

dimana $u = F_X(x)$, $v = G_Y(y)$, dan r adalah kovariat yang mempengaruhi variabel random X dan Y . Sedangkan fungsi densitas copulanya dapat dituliskan sebagai berikut:

$$h(x, y|r, \alpha, \beta, \theta) = f(x)g(y)c\left((F(x|r, \alpha)), (G(y|r, \beta)); \theta\right)$$

dengan α dan β adalah parameter dari fungsi distribusi marginal dan θ adalah parameter dari copula. dilakukan estimasi parameter model berbasis copula menggunakan metode *Maximum Likelihood Estimation* (MLE), dengan parameter yang diestimasi adalah $\nu := (\alpha^T, \beta^T, \theta, \delta)^T \in \mathbb{R}^{p+q+2}$.

Metode MLE digunakan untuk mengestimasi nilai ν , dalam fungsi *likelihood*

sebagai berikut:

$$\ell(v|\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \ln(f_{X,Y}(x_i, y_i|v))$$

Sehingga diperoleh estimasi parameter dari regresi berbasis copula dengan metode MLE sebagai berikut:

$$\hat{v} = \underset{v}{\operatorname{arg\,max}} \ell(v|\mathbf{x}, \mathbf{y})$$

Menentukan model copula terbaik yang digunakan untuk mengestimasi besarnya total kerugian dengan membandingkan nilai AIC dan Uji Vuong dari masing-masing model copula. Melakukan estimasi total kerugian klaim berdasarkan model regresi berbasis copula yang tepat.

HASIL PENELITIAN

Karakteristik Data

Statistika deskriptif data harian frekuensi dan *severity* klaim asuransi kecelakaan lalu lintas yang dibayarkan oleh perusahaan asuransi PT XYZ periode 2017–2019 dirangkum pada Tabel 3. Berdasarkan Tabel 3 diketahui bahwa frekuensi pengajuan klaim tertinggi terjadi pada tahun 2019 dengan total klaim yang diajukan sepanjang tahun 2019 adalah sebanyak 203.136 klaim kecelakaan lalu lintas. Sedangkan untuk rata-rata *severity* klaim terbesar terjadi pada tahun 2018 sebesar Rp 23.629.211,66, dengan rata-rata pembayaran tertinggi dalam satu hari di tahun 2018 adalah sebesar Rp 50.848.514,29.

Tabel 3.
Statistika Deskriptif Data Frekuensi dan *Severity* Klaim 2017-2019

Ukuran Pemusatan Data	2017	2018	2019
Frekuensi Klaim			
Mean	437	549	677
Minimum	1	1	1
Maksimum	1.736	1.423	1.213
Total	134.655	183.291	203.136
<i>Severity</i> Klaim			
Mean	19.473.600,81	23.629.211,66	19.425.448,14
Minimum	1.986.420	6.091.782,52	1.169.400
Maksimum	50.819.950	50.848.514,29	50.064.266,07
Total	5.997.869.048	7.892.156.694	5.827.634.443

Selanjutnya untuk mengetahui jenis distribusi pada data frekuensi pengajuan klaim dan severity klaim, dilakukan uji *goodness of fit* seperti yang ditunjukkan pada Tabel 4. Berdasarkan Tabel 4 diperoleh kesimpulan bahwa jenis distribusi yang sesuai menggambarkan karakteristik data frekuensi klaim adalah distribusi Poisson sedangkan distribusi yang sesuai menggambarkan karakteristik data *severity* klaim adalah distribusi Gamma, ditunjukkan dengan nilai *p-value* > taraf signifikansi (α) sebesar 5%. Sedangkan hasil pengukuran korelasi *Kendall's Tau* antara frekuensi dengan *severity* klaim adalah sebesar 0,7911045, yang menandakan bahwa terdapat hubungan ketergantungan atau korelasi antara frekuensi klaim dengan *severity* klaim.

Tabel 4.
Uji Kecocokan Distribusi Kolmogorov-Smirnov Data Frekuensi dan Severity Klaim

Jenis Distribusi	<i>P-Value</i>	Keputusan
Frekuensi Klaim		
Distribusi Poisson	0,51855	Data Berdistribusi Poisson
Severity Klaim		
Distribusi Gamma	0,22076	Data Berdistribusi Gamma

Generalized Linear Model (GLM) Frekuensi dan Severity Klaim

Karena ditemukan terdapat hubungan ketergantungan atau korelasi antara frekuensi klaim dan *severity* klaim, maka digunakan regresi copula untuk dapat memodelkan data frekuensi klaim dengan *severity* klaim. Namun sebelum dilakukan permodelan terhadap data frekuensi klaim dan *severity* klaim, terlebih dahulu dilakukan permodelan terhadap masing-masing data frekuensi klaim dan *severity* klaim secara terpisah dengan variabel (kelompok usia dan jenis kendaraan) yang diduga mempengaruhi variabel frekuensi klaim dan *severity* klaim. Untuk melakukan permodelan tersebut digunakan model marginal *Generalized Linear Model (GLM)*, sehingga diperoleh hasil estimasi parameter seperti terlihat pada Tabel 5.

Tabel 5.
Estimasi Parameter GLM Frekuensi dan Severity Klaim

Model GLM	Variabel	Estimasi	<i>P-Value</i>	Keputusan
Frekuensi Klaim Distribusi Poisson	Intercept	-4,562880	$< 2,2 \times 10^{-16}$	Signifikan
	Usia	0,552589	$< 2,2 \times 10^{-16}$	Signifikan
	Kendaraan	4,638735	$< 2,2 \times 10^{-16}$	Signifikan
Severity Klaim Distribusi Gamma	Intercept	19,21299	$< 2,2 \times 10^{-16}$	Signifikan
	Usia	-0,25782	$8,52 \times 10^{-6}$	Signifikan
	Kendaraan	-0,82382	$4,06 \times 10^{-10}$	Signifikan

Berdasarkan model GLM yang terbentuk dari data frekuensi klaim dan *severity* klaim, diketahui bahwa variabel kovariat (usia dan jenis kendaraan) signifikan berpengaruh pada fekuensi klaim dan *severity* klaim.

Estimasi Parameter Copula

Untuk melakukan analisis regresi berbasis copula, dilakukan dengan mengkombinasikan distribusi marginal dari model GLM yang sudah terbentuk sebelumnya dari data frekuensi klaim dan *severity* klaim. Tahapan selanjutnya adalah melakukan estimasi parameter copula dengan metode MLE antara dua distribusi marginal *severity* dan frekuensi klaim meliputi *copula Gauss*, *copula Clayton*, *copula Gumbel*, dan *copula Frank* seperti yang terlihat pada Tabel 6.

Tabel 6.
Estimasi Parameter Copula

Parameter	Gauss	Clayton	Gumbel	Frank	
Severity Klaim α	Intercept	1,73651	0,982936	0,89273	0,013231
	Usia	1,45362	1,032154	1,045452	1,022761
	Kendaraan	-1,32416	-1,012632	-1,234152	-1,009941
Frekuensi Klaim β	Intercept	1,892011	0,5632521	0,8982731	0,9212832
	Usia	1,0434239	1,042341	1,0124351	1,0476869
	Kendaraan	-5,7468298	-0,782619	-0,546418	-0,983279
δ	0,491224	0,4926724	0,5013344	0,498644	
θ	0	0,4564531	1,537251	2,218706	
Loglikelihood	-3.323,182	-3.109,1	-3.321,190	-3.463,182	
AIC	6.921,213	5.557,192	6.789,324	6.936,364	

Berdasarkan hasil estimasi parameter yang disajikan pada Tabel 6, diketahui bahwa model copula *Clayton* memiliki nilai AIC terkecil dibandingkan dengan model lainnya. Untuk lebih memperkuat bukti bahwa memang *copula clayton* adalah model terbaik yang dapat digunakan untuk memprediksi tingkat frekuensi dan *severity* pada asuransi kecelakaan lalu lintas PT XYZ, dilakukan Uji Vuong dengan hasil seperti yang terlihat pada Tabel 7. Berdasarkan uji vuong, model terbaik untuk mengestimasi frekuensi dan *severity* klaim asuransi kecelakaan lalu lintas PT XYZ adalah model copula *clayton*.

Tabel 7.
Uji Vuong

Model 1	Model 2			
	Gauss	Clayton	Gumbel	Frank
Gauss		-6,326039 (clayton)	82,658905 (gauss)	6,756308 (gauss)

Clayton	6,326039 (clayton)		382,658905 (clayton)	5,756308 (clayton)
Gumbel	-82,658905 (gauss)	-382,658905 (clayton)		-421,042218 (frank)
Frank	-6,756308 (gauss)	-5,756308 (clayton)	421,042218 (frank)	

Berdasarkan hasil pemodelan copula *clayton*, *gumbel*, dan *frank* yang telah diperoleh, kemudian dilakukan prediksi terhadap frekuensi dan *severity* klaim kecelakaan lalu lintas pada perusahaan asuransi PT XYZ selama periode Januari sampai dengan April 2020. Hasil prediksi frekuensi dan *severity* klaim kecelakaan beserta nilai RMSE yang dihasilkan dari perbandingan dengan data actual (data *out sample*) dapat dilihat pada Tabel 8.

Tabel 8.
Prediksi Frekuensi dan *Sevferity* Klaim Periode Januari-April 2020

Hari	Frekuensi			<i>Severity</i>		
	Clayton	Gumbel	Frank	Clayton	Gumbel	Frank
1	296	296	293	23.597.378	23.597.378	23.838.989
2	574	574	571	14.652.929	14.652.929	14.729.914
3	1038	1038	1035	13.687.870	13.687.870	13.727.545
4	716	716	713	14.255.910	14.255.910	14.315.893
5	846	846	843	13.618.861	13.618.861	13.667.327
6	1048	1048	1045	10.118.531	10.118.531	10.147.579
7	856	856	853	12.650.115	12.650.115	12.694.605
8	1	2	5	49.999.947	24.999.974	9.999.969
9	749	749	746	15.278.060	15.278.060	15.339.500
10	988	988	985	12.309.974	12.309.974	12.347.466
...
101	272	270	267	13.598.971	13.699.704	13.853.633
RMSE	0,07807281	1,138206976	1,6168078	0,930542047	0,956430529	0,96808019

Berdasarkan hasil prediksi frekuensi dan *severity* klaim serta nilai RMSE yang dihasilkan diperoleh bahwa copula *clayton* menghasilkan nilai RMSE terkecil dibandingkan dengan model copula lainnya, sehingga dapat disimpulkan bahwa model copula *clayton* merupakan model copula terbaik untuk mengestimasi frekuensi dan *severity* klaim asuransi kecelakaan lalu lintas pada PT XYZ.

PEMBAHASAN

Adapun penelitian terdahulu yang relevan dengan penelitian ini, penelitian yang dilakukan oleh Nugrahani et al., (2016) menunjukkan bahwa berdasarkan hasil estimasi data tiruan terdapat dua alternatif model distribusi yaitu normal dan gamma

yang mana membutuhkan seratus kali pengulangan siklus untuk membentuk distribusi terbaik dan hasilnya membuktikan bahwa dengan pengulangan sepuluh kali pada siklus ini didapatkan model distribusi terbaik adalah normal dengan peluang 86%.

Penelitian lain yang dilakukan oleh Prihandoko & Syuhada (2019) menunjukkan bahwa dengan menggunakan model kebergantungan copula, model klaim ACA dan metode Value-at-Risk maka didapatkan prediksi VaR untuk risiko agregat. Hasil Prediksinya adalah sebesar 13.0171 dengan model kebergantungan terbaik menggunakan Copula Gumbel. Sedangkan dalam penelitian ini menganalisis estimasi kerugian agregat dalam asuransi kecelakaan lalu lintas menggunakan model berbasis copula yang dimana hasilnya terdapat hubungan kebergantungan atau korelasi antara frekuensi klaim dan *severity* klaim, sehingga digunakan copula untuk memodelkan frekuensi klaim yang berdistribusi poison dengan *severity* klaim yang berdistribusi gamma. Dalam hal ini, model copula adalah model terbaik dalam memperkirakan kerugian agregat dalam suatu perusahaan.

SIMPULAN

Berdasarkan hasil analisis diperoleh kesimpulan bahwa terdapat hubungan kebergantungan atau korelasi antara frekuensi klaim dan *severity* klaim, sehingga digunakan copula untuk memodelkan frekuensi klaim yang berdistribusi poison dengan *severity* klaim yang berdistribusi gamma. Berdasarkan estimasi parameter model copula yang dihasilkan diperoleh bahwa *copula clayton* menjadi model terbaik untuk mengestimasi nilai frekuensi dan *severity* klaim asuransi kecelakaan lalu lintas PT XYZ. Diharapkan pada penelitian selanjutnya dapat menambahkan jenis copula lainnya yang dapat dibandingkan dan penambahan faktor risiko seperti lokasi kecelakaan sebagai kovariat yang berpengaruh terhadap banyaknya klaim asuransi kecelakaan lalu lintas.

DAFTAR PUSTAKA

- Azhar, A. A., & Fauzi, R. R. (2014). Pengaruh Indeks Good Corporate Governance terhadap Kinerja Keuangan Perusahaan di Iicg dari Tahun 2007-2010. *Jurnal Online Mahasiswa Fakultas Ekonomi Universitas Riau*, 1(2), 1-18. <https://www.neliti.com/id/publications/33430/pengaruh-indeks-good-corporate-governance-terhadap-kinerja-keuangan-perusahaan-d#cite>
- Czado, C., Kastenmeier, R., Brechmann, E. C., & Min, A. (2012). A Mixed Copula Model for Insurance Claims and Claim Sizes. *Scandinavian Actuarial Journal*, 4,

- 278–305. <https://doi.org/10.1080/03461238.2010.546147>
- Daraghma, Z. M. A. (2010). The Relative and Incremental Information Content of Earnings and Operating Cash Flows: Empirical Evidence from Middle East, the Case of Palestine. *European Journal of Economics, Finance, and Administrative Sciences*, 22, 123–135. https://www.academia.edu/18672774/The_Relative_and_Incremental_Informati_on_Content_of_Accounting_Earnings_versus_the_Cash_Recovery_Rate_?auto=download
- Klugman, S. A., Panjer, H., & Willmot, G. (2004). *Loss Models from Data to Decisions* (Second). New Jersey: John Wiley and Sons
- Krämer, N., Brechmann, E. C., Silvestrini, D., & Czado, C. (2013). Total Loss Estimation Using Copula-Based Regression Models. *Insurance: Mathematics and Economics*, 53(3), 829–839. <https://doi.org/10.1016/j.insmatheco.2013.09.003>
- Muslich, M. (2007). *Manajemen Risiko Operasional*. Jakarta: PT Bumi Aksara
- Nelsen, R. B. (2006). *An Introduction to Copula* (Second). New York: Springer Science and Business, Inc
- Nugrahani, I. S., Linawati, L., & Sasongko, L. R. (2016). Simulasi untuk Menentukan Model Distribusi Total Kerugian Agregat (Studi Kasus Data Klaim Polis Asuransi Kesehatan Manfaat Rawat Inap). *Prosiding Seminar Nasional Matematika dan Pendidikan Matematika*, 904–915. Surakarta
- Presiden, RI. (1964). *UU No 33 Tahun 1964 Jo PP No 17 Tahun 1965 tentang Dana Pertanggung Jawaban Wajib Kecelakaan Penumpang Umum*. https://www.dpr.go.id/dokjdih/document/uu/UU_1964_33.pdf
- Prihandoko, D. I., & Syuhada, K. I. A. (2019). Prediksi Ukuran Risiko Agregat Klaim Berbasis Copula pada Model Autoregressive Conditional Amount (ACA). *Jurnal Matematika Integratif*, 15(2), 129–137. <http://dx.doi.org/10.24198/jmi.v15.n2.23443.131>
- Valentovicova, K. (2015). *Claim Reserving with Copulae for Multiple Lines of Business*. Prague: Charles University