Epsilon: Jurnal Matematika Murni dan Terapan
http://ppjp.ulm.ac.id/journals/index.php/epsilon
p-ISSN: 1978-4422 | e-ISSN 2656-7660
Vol.18 No.2 (Desember 2024) Hal. 165-177

PENDEKATAN BAYESIAN MARKOV CHAIN MONTE CARLO (MCMC) METROPOLIS-HASTINGS PADA PEMODELAN KLAIM ASURANSI KESEHATAN

Aprida Siska Lestia[1*], Mochammad Idris[1]

[1]Program Studi Matematika Fakultas MIPA Universitas Lambung Mangkurat
Jl. A. Yani KM. 36, Banjarbaru 70714, Kalimantan Selatan
*Email: as_lestia@ulm.ac.id

ABSTRACT

The distribution of health insurance claims that tend to be asymmetric and have heavy tails, requires a method that can offer more flexible and robust solutions than traditional frequentist approaches, such as the Bayesian Method. The advantage of this method lies in its ability to comprehensively account for uncertainty in parameter estimation, thereby producing a posterior distribution that can capture complex pattern of claim data. This study aims to apply the Bayesian approach with the Markov Chain Monte Carlo (MCMC) Metropolis-Hastings method in modeling health insurance claims. The claim data used are divided into outpatient and inpatient claims, with the lognormal distribution fitting proven to be the most appropriate for both types of claims. Risk estimation through Value at Risk (VaR) and Conditional Tail Expectation (CTE) using the Bayesian approach showed more moderate results compared to empirical estimates, indicating that this approach can reduce risk overestimation.

Keywords: Health Insurance Claim, Bayesian, MCMC, Metropolis-Hastings

ABSTRAK

Distribusi klaim terkait asuransi kesehatan yang cenderung asimetris dan memiliki ekor berat, memerlukan sebuah metode yang bisa menawarkan solusi yang lebih fleksibel dan robust dibandingkan pendekatan frekuentis tradisional, misalnya Metode Bayesian. Keunggulan metode ini berupa kemampuannya dalam memperhitungkan ketidakpastian secara menyeluruh dalam estimasi parameter untuk kemudian menghasilkan distribusi posterior yang dapat menangkap pola kompleks pada data klaim. Penelitian ini bertujuan untuk menerapkan pendekatan Bayesian dengan metode Markov Chain Monte Carlo (MCMC) Metropolis-Hastings dalam memodelkan klaim asuransi kesehatan. Data klaim yang digunakan dibagi menjadi klaim rawat jalan dan rawat inap, dengan fitting distribusi lognormal yang terbukti paling sesuai untuk kedua jenis klaim. Estimasi risiko melalui Value at Risk (VaR) dan Conditional Tail Expectation (CTE) menggunakan pendekatan Bayesian menunjukkan hasil yang lebih moderat dibandingkan dengan estimasi empiris, mengindikasikan bahwa pendekatan ini dapat mengurangi overestimasi risiko.

Keywords: Klaim Asuransi Kesehatan, Bayesian, MCMC, Metropolis-Hastings

Received: 25 Oktober 2024, Accepted: 12 November 2024, Published: 14 November 2024

PENDAHULUAN

Asuransi kesehatan merupakan instrumen penting dalam mengelola risiko finansial yang disebabkan oleh biaya perawatan medis yang tidak terduga. Dengan semakin meningkatnya biaya kesehatan dan frekuensi klaim yang terjadi, perusahaan asuransi menghadapi tantangan besar dalam memastikan premi yang ditetapkan sesuai dengan risiko yang dihadapi. Menurut Fitriani & Gunardi (2020), hal ini penting untuk menjaga keseimbangan antara profitabilitas perusahaan dan kemampuan masyarakat untuk membayar premi yang terjangkau.

Salah satu tantangan utama dalam pemodelan klaim asuransi kesehatan adalah sifat asimetris dari distribusi klaim, di mana sebagian besar klaim berukuran kecil, namun terdapat beberapa klaim ekstrem dengan nilai yang sangat besar (outliers). Distribusi klaim yang memiliki ekor berat ini menuntut metode yang fleksibel dan robust untuk melakukan estimasi risiko secara akurat. Dalam konteks ini, pendekatan Bayesian menjadi relevan, karena metode ini memungkinkan penggabungan antara informasi sebelumnya (prior) dengan data observasi untuk menghasilkan estimasi parameter yang lebih akurat (Migon & Penna, 2006).

Pendekatan Bayesian menawarkan kerangka kerja yang fleksibel dalam memodelkan ketidakpastian dan variabilitas klaim asuransi, tidak seperti pendekatan frequentist tradisional. Hal ini sangat berguna dalam konteks klaim asuransi yang cenderung memiliki ketidakpastian tinggi. Salah satu metode yang digunakan untuk melakukan estimasi parameter dalam pendekatan Bayesian adalah melalui Markov Chain Monte Carlo (MCMC), yang digunakan untuk menghasilkan distribusi posterior dari parameter model klaim. Salah satu algoritma dalam MCMC adalah Metropolis-Hastings, yang secara efektif memungkinkan adanya simulasi distribusi posterior yang kompleks dan memiliki kemampuan untuk menangani berbagai bentuk distribusi posterior (Robert, 2015). Penggunaan algoritma Metropolis-Hastings dapat membantu dalam melakukan estimasi risiko dan penentuan premi yang lebih akurat, terutama ketika data memiliki klaim ekstrem. Parameter distribusi klaim dapat diestimasi dengan lebih baik melalui pengambilan sampel berulang dari distribusi posterior, yang mencerminkan ketidakpastian yang melekat pada parameter model klaim.

Penelitian terdahulu oleh Sukono *et al*. (2017), yang menerapkan metode Bayes untuk mengestimasi premi risko pada asuransi penyakit kritis. Hasil yang diperoleh menunjukkan bahwa pendekatan Bayesian dapat menghasilkan estimasi premi yang lebih moderat dibandingkan metode tradisional. Selain itu juga mengurangi risiko overestimasi yang sering terjadi pada pendekatan frekuentis. Penelitian lainnya terkait penggunaan pendekatan Bayesian dalam estimasi risiko dan perhitungan klaim asuransi kendaraan bermotor dilakukan oleh Sukono *et al*. (2018). Dalam penelitian ini, estimasi parameter distribusi frekuensi dan besar klaim dilakukan menggunakan metode Bayes, yang menunjukkan keunggulan pendekatan tersebut dalam menangani ketidakpastian pada estimasi klaim ekstrim. Perkiraan rata-rata dan variansi yang digunakan dalam penentuan premi lebih moderat dan sesuai, sehingga mengurangi risiko overestimasi yang dapat membebani tertanggung dengan premi yang terlalu tinggi.

Penelitian ini bertujuan untuk menerapkan pendekatan Bayesian dengan metode MCMC Metropolis-Hastings pada data klaim asuransi kesehatan untuk memperkirakan parameter distribusi klaim. Selain itu, juga akan dilakukan pengukuran risiko Selain itu, pengukuran risiko melalui Value at Risk (VaR) dan Conditional Tail Expectation (CTE). Ukuran risiko ini akan digunakan untuk memahami lebih dalam tentang risiko yang dihadapi oleh Perusahaan. Metode ini diharapkan dapat menghasilkan estimasi yang lebih andal dan akurat dalam menghadapi variabilitas dan ketidakpastian klaim asuransi kesehatan.

TINJAUAN PUSTAKA

1. Bayesian Estimation

Pendekatan Bayesian dalam estimasi parameter model parametrik melibatkan penggunaan data yang diamati untuk memperbarui keyakinan awal tentang parameter yang tidak diketahui melalui Teorema Bayes. Berbeda dengan pendekatan frekuentis yang mengasumsikan distribusi populasi tetap namun tidak diketahui, pendekatan Bayesian hanya mempertimbangkan data yang benar-benar diamati dan memandang distribusi populasi sebagai variabel yang dapat berubah. Dalam konteks aktuaria, pendekatan Bayesian digunakan untuk melakukan pembaharuan atas prediksi kerugian di masa yang akan dating. Misalkan X adalah variabel acak kerugian (banyak klaim, besar klaim, ataupun total klaim), di mana distribusi atas X bergantung pada suatu parameter θ. Parameter tersebut dapat bervariasi sehingga dianggap sebagai variabel acak yang juga memiliki distribusi sendiri (Tse, 2014).

Dalam estimasi Bayesian, parameter θ diperlakukan sebagai variabel acak Θ, dan estimasi posterior diperoleh dengan menggabungkan informasi prior dan likelihood dari data yang diamati. Langkah pertama dalam estimasi Bayesian adalah menentukan distribusi prior. Dalam (Klugman *et al.*, 2019), dinyatakan bahwa menggambarkan keyakinan awal kita tentang parameter berdasarkan pengetahuan sebelumnya atau informasi yang tersedia. Prior dapat bersifat informatif atau non-informatif. Prior informatif mencerminkan pengetahuan atau keyakinan sebelumnya tentang parameter, sedangkan prior non-informatif biasanya digunakan ketika hanya sedikit informasi awal yang tersedia (Diana & Soehardjoepri, 2016). Langkah selanjutnya adalah penentuan likelihood yang mencerminkan informasi yang disediakan oleh data yang diamati. Untuk suatu distribusi yang memiliki fungsi kepadatan peluang $f(x|\theta)$, fungsi likelihood yang terbentuk adalah

$$L(\theta|x) = \prod_{i=1}^{N} f(x|\theta) \tag{1}$$

Proses estimasi dilanjutkan dengan penentuan distribusi bersama dari fungsi likelihood dan distribusi prior, $f(\theta)$.

$$f(\theta, x) = L(\theta|x) \cdot f(\theta) \tag{2}$$

Dari sini diperoleh distribusi marginal atas X, yaitu

$$f(x) = \int_{\theta \epsilon \Omega} L(\theta|x) \cdot f(\theta)\, d\theta \tag{3}$$

Dengan Ω adalah *support* untuk Θ. Dalam (Tse, 2014), distribusi posterior yang menggabungkan prior dan likelihood serta distribusi marginal pada Persamaan (3) melalui Teorema Bayes, sehingga diperoleh

$$f(\theta|x) = \frac{L(\theta|x) \cdot f(\theta)}{\int_{\theta \epsilon \Omega} L(\theta|x) \cdot f(\theta)\, d\theta} \tag{4}$$

Menurut Markov (2013) distribusi posterior memuat banyak informasi yang lebih luas dibandingkan dengan estimasi titik yang biasa digunakan dalam analisis non-Bayesian tradisional. Untuk menghitung estimasi Bayesian, diperlukan distribusi posterior yang sering kali dihitung menggunakan algoritma numerik karena sifatnya yang kompleks, terutama Ketika distribusi prior dan likelihood tidak dalam bentuk yang sederhana.

2. Markov Chain Monte Carlo (MCMC) Metropolis-Hastings

Markov Chain Monte Carlo (MCMC) digunakan untuk memperkirakan distribusi posterior yang kompleks dengan menghasilkan sampel berulang dari distribusi tersebut. Metode ini faat karena memungkinkan estimasi parameter model, terutama ketika distribusi posterior tidak dapat dihitung secara analitik. Posterior digunakan untuk menginformasikan estimasi parameter, memberikan Batasan pada ketidakpastian, membuat prediksi, dan membandingkan model melalui integrasi atas posterior (Speagle, 2019).

Salah satu algoritma paling umum untuk menghasilkan sampel dari distribusi posterior adalah Metropolis-Hastings. Algoritma ini diterapkan dengan cara menghasilkan rantai Markov di mana setiap titik sampel dihasilkan berdasarkan distribusi probabilitas tertentu, dan Keputusan untuk menerima atau menolak sampel baru didasarkan pada rasio penerimaan (Robert & Casella, 2004). Langkah-langkah dasar algoritma Metropolis-Hastings dalam Gelman *et al*. (2013) yaitu:

(a). Menentukan fungsi target, yaitu distribusi posterior dari parameter yang akandiestimasi.
(b). Inisialisasi nilai awal parameter.
(c). Menggunakan fungsi proposal untuk menghasilkan kandidat sampel baru berdasarkan nilai parameter saat ini.
(d). Menghitung rasio penerimaan berdasarkan fungsi target (distribusi posterior) dan fungsi proposal.
(e). Menentukan apakah sampel baru diterima atau tidak. Jika diterima, perbarui nilai parameter; jika tidak, tetap gunakan nilai sebelumnya.
(f). Mengulangi proses ini untuk menghasilkan rantai Markov yang mencakup banyak iterasi.

Algoritma ini sangat fleksibel dan bisa diterapkan untuk berbagai masalah Bayesian di mana distribusi posterior tidak dapat dinyatakan secara eksplisit. Dalam konteks asuransi, misalnya, algoritma Metropolis-Hastings dapat digunakan untuk mengestimasi parameter model risiko.

3. Posterior Predictive Check (PPC)

Posterior Predictive Check (PPC) adalah Teknik evaluasi dalam inferensi Bayesian yang digunakan untuk memeriksa sejauh mana model cocok dengan data yang diamati. Dalam PPC, model Bayesian yang sudah dipasang (melalui posterior) digunakan untuk menghasilkan data baru secara simulatif, yang kemudian dibandingkan dengan data asli. Hal ini dilakukan untuk memeriksa apakah distribusi posterior yang diperoleh mampu merepresentasikan pola-pola utama dari data yang sebenarnya. Langkah-langkah PPC tersaji dalam Speagle (2019). Penggunaan PPC sangat penting untuk memvalidasi model Bayesian karena memungkinkan untuk melihat apakah model sesuai dengan data atau perlu diperbaiki. PPC sering digunakan bersamaan dengan visualisasi seperti trace plot atau density plot untuk memudahkan perbandingan antara distribusi data asli dan simulasi.

4. Ukuran Risiko

Dalam konteks asuransi, risiko diartikan sebagai peristiwa di masa depan yang memiliki kemungkinan menyebabkan kerugian finansial. Pengelolaan terhadap risiko tersebut oleh perusahaan asuransi sangat penting dilakukan, salah satunya melalui proses pengukuran nilai risiko menggunakan ukuran risiko. Dalam Tse (2014)ukuran risiko dinyatakan sebagai pemetaan fungsional yang mengubah distribusi kerugian menjadi nilai numerik.

Terdapat beberapa ukuran risiko dalam konteks aktuaria, di antaranya Value at Risk (VaR) dan Conditional Tail Expectation (CTE). VaR merupakan ukuran risiko yang berdasarkan nilai kuantil. Suatu ukuran risiko yang baik adalah yang koheren seperti dalam (Artzner P. *et al*., 1999). Dalam (Lestia, 2021), CTE terbukti koheren sedangkan VaR tidak koheren karena tidak memenuhi sifat *monotonicity* yang menyetakan bahwa jika $X \leq Y$, maka $\mathcal{H}(X) \leq \mathcal{H}(Y)$. Dalam (Klugman *et al*., 2019) didefinisikan,

Definisi 1

Misal X menyatakan variabel acak kerugian. VaR dari X pada Tingkat $100p\%$ adalah persentil ke-$100p$ dari distribusi X, yaitu

$$VaR(X) = \operatorname*{Inf}_{x \geq 0}[x|F(x) \geq p] \ , 0 < p < 1 \tag{5}$$

dengan

$$F_X(Q_\alpha) = Pr(X \leq Q_\alpha) = \alpha \tag{6}$$

Sedangkan dalam Lestia & Dumaria R.T. (2016), nilai CTE dinyatakan sebagai rata-rata nilai kerugian yang lebih dari VaR,

$$CTE(\mathrm{X}) = \mathrm{E}[\mathrm{X}|\mathrm{X} > Q_\alpha] = \mathrm{E}[\mathrm{X}|\mathrm{X} > VaR_\alpha\,(\mathrm{X})] \quad (7)$$

VaR membantu perusahaan asuransi dalam mengukur risiko yang berada dalam ambang batas tertentu dan penting untuk penentuan cadangan modal, premi, dan keputusan manajemen risiko. CTE sering digunakan untuk mengestimasi cadangan yang diperlukan untuk menghadapi kejadian risiko ekstrem yang mungkin terjadi di luar skenario yang sudah diperhitungkan oleh VaR (McNeil *et al*., 2015).

5. Uji Kolmogorov-Smirnov dan Kriterian Pemilihan Model

Uji Kolmogorov-Smirnov (K-S) adalah uji statistik non-parametrik yang digunakan untuk menguji apakah data yang diobservasi berasal dari distribusi yang diasumsikan, atau untuk melihat apakah dua distribusi data memiliki kesamaan. Uji K-S sering digunakan untuk memvalidasi apakah distribusi teoretis (seperti normal, gamma, atau lognormal) cocok untuk data yang diobservasi. Dalam uji satu sampel, uji ini membandingkan fungsi distribusi kumulatif (CDF) dari data sampel dengan CDF dari distribusi teoretis yang diasumsikan. Menurut Facchinett (2009), statistik uji D_n dihitung sebagai perbedaan maksimum antara CDF empiris dari sampel dengan distribusi teoretis atau antara dua CDF dari dua sampel. Nilai Statistik K-S mengukur perbedaan maksimum antara distribusi empiris dan distribusi teoretis. Selain itu, juga dapat digunakan nilai p-*value* di mana jika nilai p lebih kecil dari tingkat signifikansi yang ditetapkan (misalnya, 0.05), maka hipotesis nol yang menyatakan bahwa distribusi sampel sesuai dengan distribusi teoretis ditolak.

Dalam analisis statistik dan pemilihan model, memilih model terbaik sering kali menjadi tantangan, terutama ketika terdapat beberapa kandidat model yang perlu dibandingkan. Untuk tujuan ini, kriteria pemilihan model seperti *Akaike Information Criterion* (AIC), *Bayesian Information Criterion* (BIC), dan *Deviance Information Criterion* (DIC) digunakan untuk menilai dan membandingkan model berdasarkan keseimbangan antara kecocokan model terhadap data dan kompleksitas model. Dalam Rezaei & Malekjani (2021) dijelaskan sebagai berikut.

1) AIC merupakan kriteria untuk pemilihan model yang mengutamakan keseimbangan antara fit model terhadap data dan kompleksitas model. Semakin kecil nilai AIC, semakin baik model tersebut dalam mengimbangi akurasi dan kompleksitas.
2) BIC merupakan variasi dari AIC yang memberikan penalti lebih besar untuk model dengan banyak parameter, terutama dalam dataset yang besar. Model dengan nilai BIC terendah dianggap paling baik.
3) DIC merupakan kriteria pemilihan model yang dirancang khusus untuk model Bayesian hierarkis dan mempertimbangkan distribusi posterior dalam perhitungannya. Model dengan DIC terendah biasanya dianggap sebagai model terbaik.

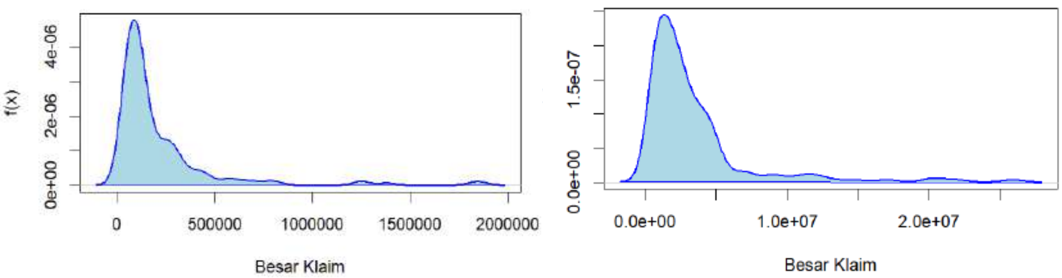
METODOLOGI

Penelitian ini menggunakan data besar klaim asuransi Kesehatan dari perusahaan PT. X di mana data terbagi menjadi dua kategori, yaitu besar klaim rawat jalan dan besar klaim rawat inap. Data untuk masing-masing kategori berukuran 200 pengamatan yang dicatat mulai Bulan Januari hingga Mei 2012. Prosedur analisis data yang dilakukan dengan tahapan sebagai berikut.

1) Membuat statistika deskriptif untuk melihat pola data.
2) Melakukan Fitting Distribusi Data Besar Klaim menggunakan uji Kolmogorov-Smirnov, termasuk melakukan estimasi parameter. Setelah itu, memilih distribusi terbaik menggunakan kriteria pemilihan *Akaike Information Criterion* (AIC) dan *Bayesian Information Criterion* (BIC).
3) Menggunakan distribusi yang terpilih untuk menentukan fungsi likelihood dan distribusi prior (Informatif dan Non-Informatif). Untuk distribusi informatif yang dipilih adalah Normal untuk parameter mean log atau Inverse-Gamma untuk variansi atau gamma untuk inverse variansi.
4) Menentukan Distribusi Posterior dengan Pendekatan Bayesian Menggunakan Metode MCMC dengan algoritma Metropolis-Hastings.
5) Mengevaluasi distribusi posterior menggunakan *Posterior Predictive Check* (PPC) dan menggunakan kriteria *Deviance Information Criterion* (DIC) untuk melihat penggunaan prior yang lebih baik.
6) Menentukan VaR dan CTE Empiris (Berdasarkan Data) dan Berdasarkan Estimasi Bayesian.

HASIL DAN PEMBAHASAN

Untuk memahami karakteristik dasar dari data klaim asuransi kesehatan yang akan dianalisis menggunakan pendekatan Bayesian, terlebih dahulu akan dilihat pola data melalui Gambar 1. Rataan besar klaim rawat jalan dan rawat inap berturut-turut adalah 203.304,9 dan 3.343.066.



(a) (b)

Gambar 1. Grafik distribusi peluang pada data besar klaim (a) rawat jalan dan (b) rawat inap

Melalui visualisasi kedua grafik, menunjukkan bahwa data sangat condong ke kanan (*skewed to the right*). Hal ini menunjukkan mayoritas klaim berukuran kecil, namun terdapat klaim besar yang cukup signifikan. Selanjutnya dilakukan fitting distribusi pada kedua data besar klaim tersebut dengan memperhatikan karakteristik yang diperoleh. Beberapa distribusi yang dinilai cocok untuk data yang *skewed to the right* antara lain, gamma, Weibull, lognormal, pareto, dan loglogistic. Untuk memeriksa apakah data yang dimiliki cocok dengan distribusi teoritis tersebut, digunakan uji Kolmogorov-Smirnov (KS) dan hasil pengujian disajikan pada Tabel 1 berikut.

Tabel 1. Hasil pengujian kecocokan model dan estimasi parameter

Distribusi	Estimasi Parameter		Hasil Uji KS, nilai AIC dan BIC	
	Rawat Jalan	Rawat Inap	Rawat Jalan	Rawat Inap
Gamma ($\boldsymbol{shape} = \boldsymbol{\tau}, \boldsymbol{scale} = \boldsymbol{\beta}$)	$\hat{\tau} = 2.8963 \cdot 10^5$ $\hat{\beta} = 1.3457$	$\hat{\tau} = 3.7532 \cdot 10^6$ $\hat{\beta} = 1.1332$	$p - value = 0.72$ $AIC = 50561443$ $BIC = 50561449$	$p - value = 0.69$ $AIC = 616159994$ $BIC = 616160000$
Weibull ($\boldsymbol{shape} = \boldsymbol{\tau}, \boldsymbol{scale} = \boldsymbol{\beta}$)	$\hat{\tau} = 1.0523$ $\hat{\beta} = 2.1952 \cdot 10^5$	$\hat{\tau} = 1.1027$ $\hat{\beta} = 3.7011 \cdot 10^6$	$p - value = 0.15$ $AIC = 5292.955$ $BIC = 5299.552$	$p - value = 0.12$ $AIC = 6411.091$ $BIC = 6417.688$
lognormal ($\boldsymbol{location} = \boldsymbol{\mu}, \boldsymbol{scale} = \boldsymbol{\sigma}$)	$\hat{\mu} = 11.7876$ $\hat{\sigma} = 0.8711$	$\hat{\mu} = 14.6120$ $\hat{\sigma} = 0.8723$	$p - value = 0.09$ $AIC = 5231.410$ $BIC = 5238.007$	$p - value = 0.057$ $AIC = 6361.712$ $BIC = 6368.309$
Pareto ($\boldsymbol{shape} = \boldsymbol{\tau}, \boldsymbol{scale} = \boldsymbol{\beta}$)	$\hat{\tau} = 7.0274$ $\hat{\beta} = 1.2162 \cdot 10^6$	$\hat{\tau} = 11.0237$ $\hat{\beta} = 3.3449 \cdot 10^7$	$p - value = 0.17$ $AIC = 5285.584$ $BIC = 5292.181$	$p - value = 0.14$ $AIC = 6410.389$ $BIC = 6416.985$
loglogistik ($\boldsymbol{shape} = \boldsymbol{\tau}, \boldsymbol{scale} = \boldsymbol{\beta}$)	$\hat{\tau} = 0.1498$ $\hat{\beta} = 1.3159 \cdot 10^5$	$\hat{\tau} = 2.0049$ $\hat{\beta} = 2.1737 \cdot 10^6$	$p - value = 0.43$ $AIC = 6034.788$ $BIC = 6041.384$	$p - value = 0.04$ $AIC = 6365.610$ $BIC = 6372.207$

Berdasarkan nilai p-value yang mengukur signifikansi dari pengujian yang dilakukan. Dalam konteks uji KS, nilai $p - value > 0{,}05$ menunjukkan bahwa model yang diuji cocok dengan data. Berarti dari kelima distribusi yang diuji, hanya loglogistic yang tidak cocok dengan kedua data besar klaim. Kemudian nilai AIC dan BIC digunakan untuk membandingkan beberapa model yang teruji cocok. Dari Tabel 1, distribusi lognormal memiliki nilai AIC dan BIC yang lebih rendah dibandingkan distribusi gamma, Weibull, dan pareto. Selisih lebih dari 10 menunjukkan perbedaan yang signifikan. Dengan demikian, lognormal lebih sesuai untuk data klaim rawat jalan maupun rawat inap. Dalam (Klugman *et al.*, 2019) fungsi kepadatan peluang untuk distribusi lognormal adalah

$$f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left[-\frac{1}{2}\left(\frac{lnx - \mu}{\sigma}\right)^2\right], \quad x \geq 0$$

Dari sini didefinisikan fungsi likelihood untuk distribusi lognormal dengan parameter lokasi μ dan parameter skala σ.

$$L(\mu,\sigma^2|x) = \prod_{i=1}^{N} f(x|\mu,\sigma^2) = \prod_{i=1}^{N} \frac{1}{x_i\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(\log x_i - \mu)^2}{2\sigma^2}\right)$$

Dalam pendekatan Bayesian, harus ditentukan pemilihan distribusi prior baik informatif maupun non informatif. Untuk distribusi lognormal, salah satu distribusi prior informatif yang dapat digunakan adalah distribusi normal untuk meanlog μ dan distribusi inverse gamma untuk σ^2. Hasil estimasi parameter pada Tabel 1 dijadikan dasar dalam penentuan parameter pada distribusi prior. Selain itu, ditetapkan pula distribusi prior non-informatif dengan jenis distribusi yang sama namun dengan nilai parameter yang berbeda. Prior untuk μ diambili nilai varians yang sangat besar untuk menunjukkan bahwa tidak ada informasi tentang nilai μ tersebut. Sedangkan nilai parameter prior untuk σ^2 diambil nilai yang sangat kecil.

Tabel 2. Distribusi prior

Distribusi prior	**Informatif**		**Non-informatif**	
	Rawat Jalan	**Rawat Inap**	**Rawat Jalan**	**Rawat Inap**
$\boldsymbol{\mu \sim Normal(\mu_0, \sigma_0^2)}$	$\mu_0 = 11{,}79$, $\sigma_0^2 = 0{,}5^2$	$\mu_0 = 14{,}6120$, $\sigma_0^2 = 0{,}5^2$	$\mu_0 = 0$, $\sigma_0^2 = 10^6$	
$\boldsymbol{\tau = \frac{1}{\sigma^2} \sim Gamma(\alpha_0, \beta_0)}$	$\alpha_0 = 2$, $\beta_0 = 0{,}871$	$\alpha_0 = 2$, $\beta_0 = 0{,}8723$	$\alpha_0 = 0{,}001$, $\beta_0 = 0{,}001$	

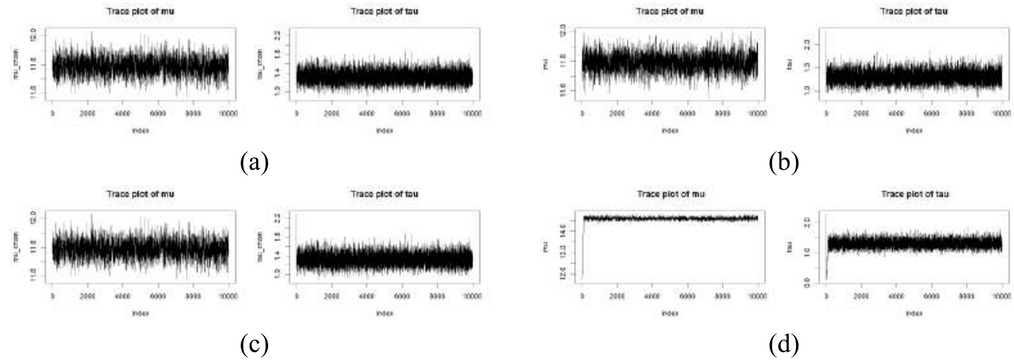
Selanjutnya, berdasarkan Persamaan (2) diperoleh

$$f(\mu,\tau,x) = L(\mu,\tau|x) \cdot f(\mu) \cdot f(\tau)$$

$$f(\mu,\tau,x) = \left[\prod_{i=1}^{N} \frac{1}{x_i\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(\log x_i - \mu)^2}{2\sigma^2}\right)\right] \times N(\mu|\mu_0,\sigma_0^2) \times Gamma(\tau|\alpha_0,\beta_0)$$

Solusi analitik untuk distribusi posterior dengan bentuk yang kompleks tidak akan mudah untuk ditentukan. Oleh karena itu, penentuan distribusi posterior marginal untuk μ dan $\tau = \frac{1}{\sigma^2}$ dilakukan menggunakan metode MCMC dengan algoritma Metropolis-Hastings. Dalam penelitian ini, prosedur Bayesian dilakukan menggunakan package MCMCpack dalam aplikasi R.

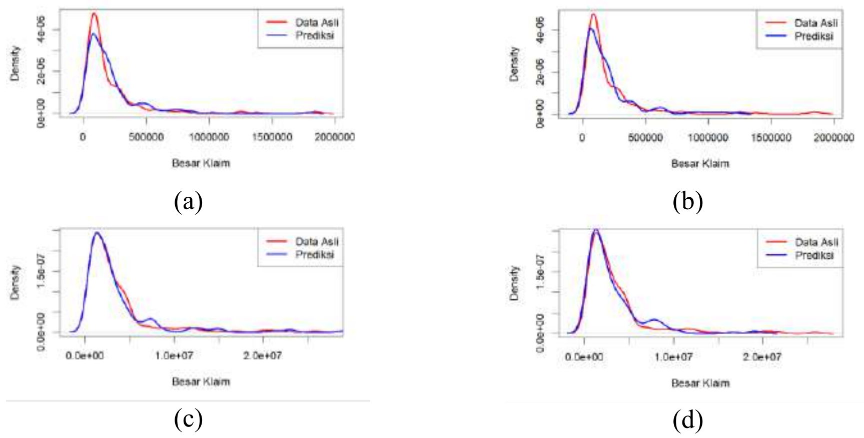
Gambar 2 menunjukkan *trace plot* untuk parameter distribusi posterior berdasarkan model Bayesian untuk data rawat jalan dan rawat inap dengan menggunakan prior informatif dan non-informatif. *Trace plot* digunakan untuk mengevaluasi konvergensi algoritma MCMC dalam sampling parameter posterior. Konvergensi yang baik ditandai dengan fluktuasi acak di sekitar nilai yang stabil tanpa tren jelas.



Gambar 2. *Trace plot* parameter μ dan τ (a) Rawat jalan informatif, (b) Rawat jalan non-informatif, (c) Rawat inap informatif, dan (d) Rawat inap non-informatif

Semua model terlihat telah mencapai konvergensi, meskipun pada rawat inap non-informatif ada indikasi kemungkinan autokorelasi atau lambatnya proses konvergensi pada parameter μ. Dalam artian, perlu lebih banyak iterasi untuk mencapai konvergensi yang baik.

Selanjutnya akan dilihat seberapa baik model dengan prior yang berbeda dapat merepresentasikan distribusi data klaim melalui perbandingan antara sebaran data asli dan prediksi (hasil dari model Bayesian) untuk data rawat jalan dan rawat inap.



Gambar 3. Perbandingan sebaran antara data asli dan prediksi (a) Rawat jalan informatif, (b) Rawat jalan non-informatif, (c) Rawat inap informatif, dan (d) Rawat inap non-informatif

Secara umum, model dengan prior informatif (panel a dan c) memberikan hasil prediksi yang lebih mendekati data asli, terutama pada puncak distribusi. Untuk rawat jalan, baik prior informatif maupun non-informatif (panel a dan b) menghasilkan sebaran prediksi yang serupa, dengan sedikit perbedaan di beberapa bagian distribusi. Namun, untuk rawat inap, prior non-informatif (panel d) menunjukkan perbedaan yang lebih signifikan di bagian ekor distribusi dibandingkan dengan prior informatif (panel c), meskipun secara keseluruhan hasilnya tetap dapat diterima.

Untuk membandingkan distribusi posterior yang diperoleh melalui model Bayesian dengan prior yang berbeda namun pada data yang sama, digunakan nilai DIC.

Tabel 3. Nilai Deviance Information Criterion (DIC)

Kategori klaim	DIC dengan Prior Informatif	DIC dengan Prior Non-informatif
Rawat jalan	5231,384	5231,437
Rawat inap	6362,153	6366,481

Perbedaan nilai DIC untuk model rawat jalan antara prior informatif dan non-informatif sangat kecil, yaitu kurang dari 1. Hal ini menunjukkan bahwa performa saat menggunakan prior informatif maupun non-informatif sangat mirip dalam menyesuaikan data. Sedangkan pada model rawat inap perbedaan DIC sebesar 4,328, menunjukkan bahwa model dengan prior informatif sedikit lebih baik dalam menyesuaikan data rawat inap dibandingkan dengan prior non-informatif. Meskipun perbedaan ini tidak terlalu signifikan karena masih berada di bawah 10.

Tabel 1 di bawah ini menyajikan perbandingan hasil estimasi VaR dan CTE untuk kedua jenis data pada tingkat kepercayaan 95% dan 99%. Nilai VaR dan CTE dihitung berdasarkan tiga pendekatan, yakni pendekatan empiris (berdasarkan data asli), model dengan prior informatif, dan prior non-informatif. VaR digunakan untuk mengukur risiko maksimum kerugian pada tingkat kepercayaan tertentu, sedangkan CTE memberikan informasi tambahan tentang ekspektasi kerugian yang terjadi melebihi VaR, terutama di bagian ekor distribusi. Perbandingan ini untuk untuk mengevaluasi perbedaan antara pendekatan empiris dan pendekatan Bayesian, serta melihat pengaruh prior terhadap estimasi risiko.

Tabel 4. Nilai VaR dan CTE

Data	*Prior*	*VaR*		*CTE*	
		95%	99%	95%	99%
Rawat Jalan	**Empiris**	602.639,1	1.383.207,3	1.124.231	1.845.119
	Informatif	554.992,6	1.008.401	852.576,2	1.417.948
	Non Informatif	548.203,8	1.007.310	870.540,4	1.475.238
Rawat Inap	**Empiris**	10.949.415.4	20.702.411.6	16.925.505	23.912.356
	Informatif	9.017.904	16.413.225	13.868.888	23.116.176
	Non Informatif	9.422.393	17.586.809	14.595.559	24.463.448

Hasil VaR dan CTE menunjukkan bahwa pendekatan empiris memberikan estimasi risiko yang lebih tinggi dibandingkan dengan model Bayesian. Namun, tidak ada perbedaan signifikan antara estimasi risiko baik menggunakan prior informatif maupun non-informatif pada data rawat jalan. Sedangkan pada data rawat inap, prior non-informatif memberikan prediksi risiko yang lebih tinggi, terutama pada tingkat kepercayaan yang lebih tinggi. Berarti prior non-informatif dapat digunakan jika kita lebih khawatir tentang risiko ekstrem. Namun, jika Anda memiliki informasi sebelumnya yang andal, prior informatif tetap merupakan pilihan yang baik.

KESIMPULAN

Pendekatan Bayesian yang diterapkan melalui metode MCMC Metropolis-Hastings dalam pemodelan klaim asuransi kesehatan telah menunjukkan hasil yang signifikan dalam mengestimasi parameter distribusi klaim dengan lebih akurat. Distribusi lognormal terbukti lebih sesuai untuk data klaim rawat jalan dan rawat inap, memberikan hasil yang lebih baik dalam mengatasi asimetri data. Pada pendekatan Bayesian, baik dengan prior informatif maupun non-informatif, hasil VaR dan CTE cenderung lebih rendah, terutama di tingkat kepercayaan yang lebih tinggi seperti 99%. Ini menunjukkan bahwa model Bayesian, dengan mempertimbangkan informasi tambahan melalui prior, menghasilkan estimasi risiko yang lebih terkendali dan tidak berlebihan dibandingkan dengan pendekatan empiris yang cenderung memberikan estimasi risiko lebih besar, khususnya di ekor distribusi. Hal ini mengindikasikan bahwa pendekatan Bayesian dapat mengurangi ketidakpastian dan menghindari *overestimation* risiko, yang bisa lebih bermanfaat dalam pengambilan keputusan terkait manajemen risiko.

DAFTAR PUSTAKA

Artzner P., Delbaen, F., Eber, J. M., & Heath D. (1999). Coherent measures of risk. *Math. Finance*. Vol. 9(3), 203–228. https://doi.org/https://doi.org/10.1111/1467-9965.00068

Diana, E. N., & Soehardjoepri. (2016). Pendekatan Metode Bayesian untuk Kajian Estimasi Parameter Distribusi Log-Normal untuk Non-Informatif Prior. *Jurnal Sains Dan Seni ITS*. Vol. 5(2), A14–A16. https://doi.org/https://dx.doi.org/10.12962/j23373520.v5i2.16468

Facchinett, S. (2009). A Procedure to Find Exact Critical Values of Kolmogorov-Smirnov Test. *Italian Journal of Applied Statistics*. Vol. 21, 337–359.

Fitriani, R., & Gunardi. (2020). Implementasi Metode Bayes pada perhitungan Premi Asuransi Kendaraan Bermotor. *Journal of Fundamental Mathematics and Applications (JFMA)*. Vol. 3(2), 112–123. https://doi.org/10.14710/jfma.v3i2.8257

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis* (3rd Edition). Chapman and Hall Press.

Klugman, S. A., Panjer, H. H., & Willmot, G. E. (2019). *Loss Models from Data to Decisions* (5th ed.). John Wiley and Sons, Inc. www.wiley.com

Lestia, A. S. (2021). *Model Risiko* (Vol. 1). CV. Banyubening Cipta Sejahtera.

Lestia, A. S., & Tampubolon, D. R. (2016). Capital-Based Risk Measures For Some Probability Models. *Proceedings of International Conference on NAMES 2015, Faculty of Mathematics and Natural Sciences*. 217–224.

Markov, U. E. (2013). Principal applications of Bayesian methods in actuarial science. *North American Actuarial Journal*. Vol. 5(4), 53–57. https://doi.org/10.1080/10920277.2001.10596011

McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques, and Tools (Revised Edition)*. Princeton University Press.

Migon, H. S., & Penna, E. M. O. (2006). Bayesian Analysis of a Health Insurance Model. *Journal of Actuarial Practice*. Vol. 13, 60–80.

Rezaei, M., & Malekjani, M. (2021). Comparison between different methods of model selection in cosmology. *The European Physical Journal Plus*. Vol. 136(2), 219. https://doi.org/10.1140/epjp/s13360-021-01200-w

Robert, C. P. (2015). The Metropolis–Hastings Algorithm. In *Wiley StatsRef: Statistics Reference Online*. 1–15. Wiley. https://doi.org/10.1002/9781118445112.stat07834

Robert, C. P., & Casella, G. (2004). *Monte Carlo Statistical Methods* (2nd Edition). Springer.

Speagle, J. S. (2019). A Conceptual Introduction to Markov Chain Monte Carlo Methods. *Harvard & Smithsonian Center for Astrophysics.*, *arXiv-1909.*

Sukono, Riaman, Lesmana, E., Wulandari, R., Napitupulu, H., & Supian, S. (2018). Model estimation of claim risk and premium for motor vehicle insurance by using Bayesian method. *IOP Conference Series: Materials Science and Engineering*. Vol. 300(1). https://doi.org/10.1088/1757-899X/300/1/012027

Sukono, Suyudi, M., Islamiyati, F., & Supian, S. (2017). Estimation model of life insurance claims risk for cancer patients by using Bayesian method. *IOP Conference Series: Materials Science and Engineering*. Vol. 166(1). https://doi.org/10.1088/1757-899X/166/1/012022

Tse, Y. K. (2014). *Nonlife Actuarial Models Theory, Methods and Evaluation*. United States of America by Cambridge University Press.