

# Estimasi *Aggregate Loss* Menggunakan Pendekatan Bayesian Metode MCMC Algoritma *Gibbs-Sampling* dengan *Software OpenBUGS*

Melati Sinta Nurdanita dan Azizah



Volume 13, Issue 3, Pages 300–310, Dec. 2025

Diterima 1 Agustus 2025, Direvisi 28 Oktober 2025, Disetujui 4 November 2025, Diterbitkan 1 Desember 2025

To Cite this Article : M. S. Nurdanita dan A. Azizah, "Estimasi *Aggregate Loss* Menggunakan Pendekatan Bayesian Metode MCMC Algoritma *Gibbs-Sampling* dengan *Software OpenBUGS*", *Euler J. Ilm. Mat. Sains dan Teknol.*, vol. 13, no. 3, pp. 300–310, 2025, <https://doi.org/10.37905/euler.v13i3.33769>

© 2025 by author(s)

## JOURNAL INFO • EULER : JURNAL ILMIAH MATEMATIKA, SAINS DAN TEKNOLOGI



🏠	Homepage	:	<a href="http://ejournal.ung.ac.id/index.php/euler/index">http://ejournal.ung.ac.id/index.php/euler/index</a>
📄	Journal Abbreviation	:	Euler J. Ilm. Mat. Sains dan Teknol.
📅	Frequency	:	Three times a year
🗣️	Publication Language	:	English (preferable), Indonesia
🔗	DOI	:	<a href="https://doi.org/10.37905/euler">https://doi.org/10.37905/euler</a>
📖	Online ISSN	:	2776-3706
📄	License	:	Creative Commons Attribution-NonCommercial 4.0 International License
🏢	Publisher	:	Department of Mathematics, Universitas Negeri Gorontalo
🌍	Country	:	Indonesia
🌐	OAI Address	:	<a href="http://ejournal.ung.ac.id/index.php/euler/oai">http://ejournal.ung.ac.id/index.php/euler/oai</a>
🔍	Google Scholar ID	:	QF_r-gAAAAJ
✉️	Email	:	<a href="mailto:euler@ung.ac.id">euler@ung.ac.id</a>

## JAMBURA JOURNAL • FIND OUR OTHER JOURNALS



Jambura Journal of Biomathematics



Jambura Journal of Mathematics



Jambura Journal of Mathematics Education



Jambura Journal of Probability and Statistics

# Estimasi Aggregate Loss Menggunakan Pendekatan Bayesian Metode MCMC Algoritma Gibbs-Sampling dengan Software OpenBUGS

Melati Sinta Nurdanita<sup>1</sup> dan Azizah<sup>1,\*</sup>

<sup>1</sup>Program Studi Matematika, Universitas Negeri Malang, Kota Malang 65145, Indonesia

## ARTICLE HISTORY

Diterima 1 Agustus 2025  
Direvisi 28 Oktober 2025  
Disetujui 4 November 2025  
Diterbitkan 1 Desember 2025

## KATA KUNCI

Aggregate Loss  
Bayesian  
Markov Chain Monte Carlo  
Gibbs-Sampling  
OpenBUGS

## KEYWORDS

Aggregate Loss  
Bayesian  
Markov Chain Monte Carlo  
Gibbs-Sampling  
OpenBUGS

**ABSTRAK.** Penelitian ini bertujuan untuk mengestimasi aggregate loss (total kerugian) pada data asuransi mobil penumpang pribadi menggunakan pendekatan Bayesian metode Markov Chain Monte Carlo (MCMC) algoritma Gibbs-Sampling dengan bantuan software OpenBUGS. Pendekatan dilakukan dengan memodelkan data frekuensi klaim menggunakan distribusi Geometrik dan Binomial Negatif, serta severitas klaim menggunakan distribusi Gamma dan Lognormal. Selanjutnya, ditentukan prior untuk masing-masing model serta perhitungan untuk fungsi likelihood, distribusi bersama, marginal, dan posterior. Karena distribusi posterior yang dihasilkan tidak dapat dihitung secara analitik, maka dilakukan simulasi melalui software OpenBUGS untuk menghitungnya. Simulasi juga digunakan dalam perhitungan prediktif posterior untuk mengestimasi aggregate loss di masa depan. Hasil penelitian menunjukkan bahwa pendekatan Bayesian dengan metode Markov Chain Monte Carlo menggunakan algoritma Gibbs-Sampling dan implementasinya melalui software OpenBUGS dapat digunakan untuk mengestimasi aggregate loss. Dari simulasi yang digunakan, diperoleh bahwa estimasi nilai aggregate loss asuransi mobil penumpang pribadi dipengaruhi oleh pemilihan model frekuensi dan severitas klaim. Model Binomial Negatif-Gamma menghasilkan estimasi prediktif posterior aggregate loss tertinggi sebesar \$75270,00, sedangkan model Geometrik-Lognormal memberikan estimasi terendah sebesar \$70500,0. Sementara itu, model dengan standar deviasi terkecil adalah model Binomial Negatif-Lognormal, yaitu \$62720,0. Penelitian ini memberikan implikasi bahwa pemilihan model dapat disesuaikan dengan target pasar yang ditetapkan oleh perusahaan. Hal ini bertujuan agar penentuan dana cadangan dan penetapan premi asuransi lebih efektif sehingga potensi kerugian yang ditimbulkan dapat diminimalkan.

**ABSTRACT.** This study aims to estimate aggregate loss (total loss) in private passenger car insurance data using the Bayesian approach of the Markov Chain Monte Carlo (MCMC) algorithm Gibbs-Sampling with the help of OpenBUGS software. The approach was carried out by modeling claim frequency data using Geometric and Negative Binomial distributions, and claim severity using Gamma and Lognormal distributions. Next, the prior for each model was determined, along with calculations for the likelihood function, joint distribution, marginal distribution, and posterior distribution. Since the resulting posterior distribution could not be calculated analytically, simulation was performed using OpenBUGS software to calculate it. Simulation was also used in predictive posterior calculations to estimate future aggregate losses. The results show that the Bayesian approach with the Markov Chain Monte Carlo method using the Gibbs-Sampling algorithm and its implementation through OpenBUGS software can be used to estimate aggregate loss. From the simulations used, it was found that the estimation of aggregate loss for private passenger car insurance is influenced by the selection of the frequency and severity of claims models. The Negative-Gamma Binomial model produced the highest posterior predictive estimate of aggregate loss at \$75270.0, while the Geometric-Lognormal model provided the lowest estimate at \$70500.0. Meanwhile, the model with the smallest standard deviation is the Negative Binomial-Lognormal model, which is \$62720.0. This study contributes to insurance risk modeling, particularly in determining reserve funds and setting insurance premiums tailored to the target market of insurance companies.



This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International License. **Editorial of EULER:** Department of Mathematics, Universitas Negeri Gorontalo, Jln. Prof. Dr. Ing. B. J. Habibie, Bone Bolango 96554, Indonesia.

## 1. Pendahuluan

Dalam kehidupan sehari-hari, manusia tidak akan lepas dari risiko, yaitu ketidakpastian yang dapat menyebabkan kerugian [1]. Risiko dapat terjadi pada individu maupun organisasi, seperti kecelakaan, penyakit, kehilangan pekerjaan, kerusakan atau kehilangan aset, fluktuasi harga pasar, risiko likuiditas, dan gangguan

pada proses internal yang dapat menimbulkan kerugian finansial [1]. Meskipun risiko tidak dapat sepenuhnya dihindari, kerugian yang ditimbulkan dapat diminimalkan melalui asuransi. Asuransi merupakan bentuk usaha untuk mengalihkan risiko dari individu atau organisasi kepada perusahaan asuransi [2]. Oleh sebab itu, ketika bertanggung mengajukan klaim atas risiko yang dijamin dalam polis, perusahaan asuransi berkewajiban untuk membayar

\*Penulis Korespondensi.

ganti rugi atas kerugian finansial yang dialami oleh tertanggung, yang disebut sebagai loss. Loss dibagi menjadi dua jenis, yaitu *individual loss* dan *aggregate loss*. *Individual loss* merupakan nilai kerugian finansial per klaim, sedangkan *aggregate loss* merupakan total kerugian finansial dari seluruh klaim yang diajukan [3]. Estimasi terhadap *aggregate loss* ini penting, karena estimasi tersebut memiliki pengaruh terhadap berbagai aspek operasional perusahaan, seperti penentuan premi, dana cadangan, dan strategi manajemen risiko [4]. Terdapat dua pendekatan utama dalam membangun model *aggregate loss*, yaitu *collective risk model* dan *individual risk model* [5]. *Collective risk model* sebagai pendekatan dalam per hitungan *aggregate loss* telah digunakan dalam penelitian berjudul Penerapan *Collective Risk Model* dalam Penentuan Premi Asuransi Bencana Alam [6]. Penggunaan model ini menunjukkan relevansi pendekatan *collective risk model* dalam konteks *aggregate loss*. Penelitian ini mengambil konteks serupa, namun dengan pendekatan yang berbeda dalam proses estimasinya, yaitu pendekatan Bayesian yang mengestimasi parameter melalui distribusi peluang, bukan nilai tunggal seperti pada pendekatan klasik. Sebelum melakukan estimasi, penting untuk memiliki pemahaman mengenai model dari *aggregate loss* itu sendiri.

Model *aggregate loss* adalah total kerugian yang diwakili oleh variabel acak berdasarkan polis asuransi [7]. *Aggregate loss* bergantung pada frekuensi klaim (banyaknya klaim) dan severitas klaim (besaran klaim) [8]. Karena klaim-klaim tersebut tidak dapat diprediksi secara pasti, perusahaan asuransi memerlukan data historis dan pendekatan statistik untuk mengestimasi besarnya nilai *aggregate loss* di masa depan. Secara umum, pendekatan statistik dibagi menjadi dua macam, yaitu frekuentis dan Bayesian. Pendekatan frekuentis didasarkan pada gagasan bahwa distribusi sampel dari suatu statistik uji tertentu dapat dibentuk melalui eksperimen dalam jumlah yang besar [9], sedangkan pendekatan Bayesian merupakan pendekatan statistik yang didasarkan pada teorema Bayes [10]. Pendekatan Bayesian berbeda dengan pendekatan frekuentis, karena pada pendekatan Bayesian diperlukan adanya distribusi prior dan distribusi posterior. Selain itu, parameter pada pendekatan frekuentis berupa besaran tetap yang nilainya tidak diketahui, sedangkan parameter pada pendekatan Bayesian berupa nilai yang keragamannya digambarkan melalui distribusi peluang (distribusi prior). Secara umum, pendekatan Bayesian lebih unggul daripada pendekatan frekuentis untuk mengestimasi *aggregate loss* ketika informasi dari data sebelumnya (distribusi prior) diketahui karena kemampuan pendekatan Bayesian dalam menggabungkan informasi dari data sebelumnya (distribusi prior) dengan data setelah dilakukannya pengamatan, dimana kemampuan ini tidak dimiliki oleh pendekatan frekuentis [11].

Pada penelitian yang dilakukan oleh Sunil Hariram Joshi dengan judul *A Bayesian Analysis of Aggregate Loss Models* [11], pendekatan Bayesian diterapkan untuk menganalisis *aggregate loss* dari suatu asuransi mobil. Penelitian tersebut menggunakan portofolio asuransi mobil dengan data yang relatif kecil dan tidak memiliki informasi mengenai data sebelumnya. Hal ini bertujuan untuk mengetahui ketepatan pendekatan Bayesian dibandingkan pendekatan frekuentis dalam memodelkan *aggregate loss* setelah informasi prior dihilangkan. Berdasarkan penelitian tersebut, diperoleh hasil bahwa pendekatan Bayesian lebih unggul daripada pendekatan frekuentis dalam memodelkan *aggregate loss*,

meskipun informasi awalnya tidak diketahui. Sejalan dengan penelitian tersebut, beberapa studi lainnya juga menunjukkan keunggulan pendekatan Bayesian dalam pemodelan risiko asuransi. Deng dan Amindazeh [12] membandingkan penggunaan pendekatan Bayesian dengan *Maximum Likelihood* (MLE) pada model kerugian asuransi komposit, dan menemukan bahwa pendekatan Bayesian memberikan akurasi yang lebih baik berdasarkan hasil simulasi. Selanjutnya, Tahajuddin dan Ismail [13] membandingkan model regresi *Zero-Inflated Poisson* (ZIP) dan *Zero-Inflated Negative Binomial* (ZINB) menggunakan pendekatan Bayesian dan frekuentis untuk frekuensi klaim asuransi kendaraan di Malaysia, dan menunjukkan bahwa pendekatan Bayesian memberikan hasil *goodness-of-fit* yang lebih baik dan kesalahan prediksi yang lebih kecil. Oleh karena itu, penelitian ini menggunakan pendekatan Bayesian untuk mengestimasi *aggregate loss*. Salah satu komponen utama pada pendekatan Bayesian adalah distribusi posterior. Distribusi posterior seringkali tidak dapat dihitung secara analitik, sehingga diperlukan metode numerik untuk menghitungnya. Salah satu metode numerik yang dapat digunakan adalah *Markov Chain Monte Carlo* (MCMC).

Metode MCMC adalah teknik simulasi berbasis rantai Markov yang digunakan untuk mengestimasi besaran yang nilainya tidak diketahui dari suatu distribusi peluang [14]. Metode ini menggabungkan dua konsep utama, yaitu *Markov Chain* (proses stokastik dimana keadaan berikutnya hanya bergantung pada keadaan saat ini) dan *Monte Carlo* (estimasi melalui pengambilan sampel). Metode MCMC memiliki beberapa algoritma, salah satunya adalah algoritma *Gibbs-Sampling*. Algoritma *Gibbs-Sampling* merupakan bentuk khusus dari algoritma *Metropolis-Hastings* yang menggunakan densitas proposal, yaitu distribusi target bersyarat penuh. Kelebihan algoritma ini adalah kemampuannya untuk diterapkan secara luas dalam analisis statistik model Bayesian yang kompleks, khususnya dalam konteks parameter multivariat, serta secara umum digunakan ketika simulasi dari distribusi bersyarat dapat dilakukan [15].

Pada penelitian yang berjudul Estimasi Parameter Model Volatilitas Stokastik dengan Metode Bayesian Rantai Markov Monte Carlo untuk Memprediksi Return Saham, dimana algoritma yang digunakan adalah *Gibbs-Sampling*, diperoleh hasil prediksi untuk *return* saham mendekati nilai sebenarnya [16]. Berdasarkan hasil tersebut, dapat disimpulkan bahwa pendekatan Bayesian menggunakan metode *Markov Chain Monte Carlo* dengan algoritma *Gibbs-Sampling* efektif digunakan dalam mengestimasi parameter model volatilitas stokastik. Di sisi lain, Goffard dan Laub menerapkan pendekatan *Approximate Bayesian Computation* (ABC) untuk memodelkan *aggregate loss* tanpa menggunakan fungsi *likelihood*. Meskipun pendekatan ini telah diterapkan, namun bidang ini relatif baru dan menjadi subjek penelitian intensif, sehingga terdapat banyak variasi yang sedang diteliti dan tidak ada konsensus yang jelas mengenai variasi mana dari ABC yang paling optimal dalam konteks estimasi *aggregate loss*. Selain itu, penelitian tersebut mendorong pembaca untuk mempertimbangkan alternatif lain seperti MCMC dalam konteks serupa [17]. Selanjutnya, pada penelitian yang berjudul *Bayesian Analysis of Aggregate Loss Models*, digunakan pendekatan Bayesian menggunakan distribusi Coxian dan *Reversible Jump Markov Chain Monte Carlo* [18]. Hasil penelitian tersebut menunjukkan bahwa metode yang digunakan dapat mengestimasi *aggregate loss* dengan lebih

akurat dibandingkan dengan metode parametrik, namun membutuhkan waktu komputasi yang lebih lama. Berdasarkan pertimbangan tersebut, penelitian ini bertujuan untuk mengestimasi *aggregate loss* menggunakan pendekatan Bayesian metode *Markov Chain Monte Carlo* algoritma *Gibbs-Sampling* dengan bantuan *software* OpenBUGS.

OpenBUGS dipilih karena *software* ini menggunakan algoritma *Gibbs-Sampling* dalam estimasi Bayesian. Selain itu, *software* ini memiliki beberapa kelebihan, diantaranya cukup fleksibel dalam memodelkan berbagai karakteristik data, mudah digunakan, serta hanya membutuhkan ruang penyimpanan yang relatif kecil dibandingkan *software* lainnya [19]. Pada halaman utama OpenBUGS tersedia berbagai macam menu, seperti Model yang berfungsi untuk memeriksa kesalahan spesifikasi model dan menentukan jumlah iterasi, serta *Inference* yang berfungsi untuk menghasilkan *output* dari estimasi yang diteliti. Di dalam menu model, pengguna dapat menentukan kode model, data, dan *inits* (nilai awal). Setelah menentukan kode-kode tersebut, hasil estimasi dapat dilihat di bagian menu *Inference* setelah proses simulasi dilakukan.

Beberapa penelitian sebelumnya telah menggunakan pendekatan Bayesian dalam konteks asuransi, namun menggunakan metode simulasi yang berbeda. Misalnya, penelitian berjudul *A Bayesian Analysis of Aggregate Loss Models* menggunakan algoritma *Metropolis-Hastings*, dan penelitian oleh Goffard dan Laub menggunakan *Approximate Bayesian Computation* (ABC) yang tidak memerlukan fungsi *likelihood*. Sementara itu, penelitian berjudul *Estimasi Parameter Model Volatilitas Stokastik dengan Metode Bayesian Rantai Markov Monte Carlo untuk Memprediksi Return Saham* telah menggunakan pendekatan Bayesian dengan algoritma *Gibbs-Sampling*, tetapi pada konteks data keuangan. Berbeda dengan penelitian-penelitian tersebut, penelitian ini menggunakan pendekatan Bayesian dengan algoritma *Gibbs-Sampling* untuk memodelkan *aggregate loss* dengan berbagai kombinasi model frekuensi-severitas pada data asuransi mobil penumpang pribadi. Penelitian ini penting guna memperluas pemahaman mengenai fleksibilitas penggunaan pendekatan Bayesian serta mendorong penerapan yang lebih luas dalam bidang aktuarial.

Kebaruan penelitian ini terletak pada penerapan pendekatan Bayesian menggunakan algoritma *Gibbs-Sampling* melalui OpenBUGS untuk memodelkan *aggregate loss* dengan beberapa kombinasi model frekuensi-severitas, yaitu Geometrik-Gamma, Geometrik-Lognormal, Binomial Negatif-Gamma, dan Binomial Negatif-Lognormal. Pendekatan ini dilakukan secara sistematis untuk menggambarkan karakteristik risiko pada data asuransi mobil penumpang pribadi. Dengan rancangan tersebut, penelitian ini diharapkan dapat memberikan kontribusi melalui penerapan pendekatan Bayesian di bidang aktuarial, khususnya dengan menerapkan penerapan yang mudah diikuti dan relevan untuk analisis risiko asuransi menggunakan data nyata.

## 2. Metode

### 2.1. Data

Pada penelitian ini, data yang digunakan berupa data sekunder pada tahun 2022 yang diperoleh dari *website* Kaggle. Terdapat total 32 data grup polis asuransi, masing-masing merupakan kombinasi usia pengguna mobil dengan jenis penggunaan. Setiap kombinasi tersebut, dilengkapi dengan data frekuensi klaim dan severitas klaim dari asuransi mobil penumpang pribadi. Ni-

lai pembayaran dinyatakan dalam mata uang poundsterling. Data ini disajikan pada Tabel 1, dengan frekuensi klaim dinotasikan dengan  $N$ , severitas klaim dengan  $X$ , dan  $i$  menotasikan grup polis asuransi ke- $i$ .

Tabel 1. Data (poundsterling)

$i$	$N$	$X$	$i$	$N$	$X$
1	21	250,48	17	151	153,62
2	40	274,78	18	479	201,67
3	23	244,52	19	381	238,21
4	5	797,80	20	166	256,21
5	63	213,71	21	245	208,59
6	171	298,60	22	970	202,80
7	92	298,13	23	719	236,06
8	44	362,23	24	304	352,49
9	140	250,57	25	266	207,57
10	343	248,56	26	859	202,67
11	318	297,90	27	504	253,63
12	129	342,31	28	162	340,56
13	123	229,09	29	260	192,00
14	448	228,48	30	578	196,33
15	361	293,87	31	312	259,79
16	169	367,46	32	96	342,58

### 2.2. Prosedur Penelitian

Pendekatan yang digunakan untuk mengestimasi *aggregate loss* dalam penelitian ini adalah pendekatan Bayesian. Berikut langkah-langkah yang dilakukan:

1. Menentukan statistika deskriptif pada data frekuensi dan severitas klaim.
2. Melakukan uji kecocokan distribusi untuk memodelkan frekuensi klaim dan severitas klaim.
3. Menentukan distribusi prior, fungsi *likelihood*, dan distribusi posterior.
4. Menggunakan metode MCMC dengan algoritma *Gibbs-Sampling* yang proses perhitungannya dibantu *software* OpenBUGS untuk mengestimasi distribusi posterior dan prediktif posterior.

Karena pada penelitian ini digunakan *software* OpenBUGS sebagai alat bantu perhitungan dalam estimasi *aggregate loss*, maka *probability mass function* (pmf) dan *probability density function* (pdf) dari setiap model distribusi disesuaikan dengan OpenBUGS.

### 2.3. Distribusi Frekuensi Klaim

Frekuensi klaim dimodelkan sebagai variabel acak diskrit tak-negatif yang dinotasikan dengan  $N$  [3]. Beberapa distribusi diskrit tak-negatif yang biasa digunakan untuk memodelkan frekuensi klaim adalah distribusi Geometrik dan Binomial Negatif.

#### 2.3.1. Distribusi Geometrik

Variabel acak  $N$  berdistribusi Geometrik dengan parameter  $p$ , dinotasikan  $N \sim GEO(p)$  dengan  $0 < p < 1$  [20]:

$$f(n) = (1 - p)^{n-1}p; \text{ untuk } n = 1, 2, \dots, \quad (1)$$

dengan nilai mean dari  $N \sim GEO(p)$ :

$$E(N) = \frac{1}{p}. \quad (2)$$

### 2.3.2. Distribusi Binomial Negatif

Distribusi Binomial Negatif digunakan untuk memodelkan variabel acak  $N$  dengan parameter  $p$  dan  $r$ , yang dinotasikan  $N \sim NB(p, r)$ , dengan  $0 < p < 1$ , dan  $r > 0$  [3]:

$$f(n) = \frac{(n+r-1)!}{n!(r-1)!} p^r (1-p)^n; \text{ untuk } n = 0, 1, 2, \dots, \quad (3)$$

dengan mean dari distribusi Binomial Negatif:

$$E(N) = \frac{r(1-p)}{p}. \quad (4)$$

### 2.4. Distribusi Severitas Klaim

Tidak seperti frekuensi klaim yang dimodelkan sebagai variabel acak bernilai bilangan bulat tak-negatif, severitas klaim dimodelkan sebagai variabel acak kontinu tak-negatif yang dinotasikan dengan  $X$  [1]. Beberapa distribusi kontinu tak-negatif yang dapat digunakan untuk memodelkan severitas klaim adalah distribusi Gamma dan Lognormal.

#### 2.4.1. Distribusi Gamma

Variabel acak  $X$  mengikuti distribusi Gamma dengan parameter bentuk (*shape parameter*)  $\kappa$  dan parameter laju (*rate parameter*)  $\theta$ , dinotasikan  $X \sim GAM(\kappa, \theta)$ , dengan  $\kappa > 0$ , dan  $\theta > 0$ :

$$f(x) = \frac{\theta^\kappa x^{\kappa-1} e^{-\theta x}}{\Gamma(\kappa)}; \text{ untuk } x > 0, \quad (5)$$

dengan mean dari  $X \sim GAM(\kappa, \theta)$ :

$$E(X) = \frac{\kappa}{\theta}. \quad (6)$$

#### 2.4.2. Distribusi Lognormal

Distribusi Lognormal dengan parameter lokasi (*location parameter*)  $\mu$  dan *precision* parameter  $\tau$  digunakan untuk memodelkan variabel acak  $X$ , yang dinotasikan  $X \sim LN(\mu, \tau)$ , dengan  $-\infty < \mu < \infty$ , dan  $\tau > 0$ :

$$f(x) = \sqrt{\frac{\tau}{2\pi}} \frac{1}{x} \exp\left(-\frac{\tau}{2}(\log x - \mu)^2\right) \text{ untuk } x > 0, \quad (7)$$

dengan nilai mean dari distribusi:

$$E(X) = \exp\left(\mu + \frac{1}{\tau}\right). \quad (8)$$

### 2.5. Uji Kecocokan Distribusi

Uji kecocokan distribusi merupakan metode statistik yang digunakan untuk menentukan apakah sampel acak  $x_1, x_2, \dots, x_n$  berasal dari suatu populasi dengan distribusi tertentu [2]. Uji kecocokan distribusi memiliki hipotesis sebagai berikut:

$H_0$  : Data berasal dari suatu populasi dengan distribusi tertentu

$H_1$  : Data tidak berasal dari suatu populasi dengan distribusi tertentu.

Penelitian ini menggunakan dua uji untuk menguji kecocokan distribusi, yaitu uji *Kolmogorov-Smirnov* untuk model frekuensi klaim dan uji *Chi-Square* untuk model severitas klaim.

#### 2.5.1. Uji Kolmogorov-Smirnov

Uji *Kolmogorov-Smirnov* dinotasikan sebagai  $D$ :

$$D = \text{Max}|F^*(x) - F(x)|, \quad (9)$$

dengan  $\text{Max}$  merupakan nilai maksimum,  $F^*(x)$  adalah fungsi distribusi kumulatif dari model, dan  $F(x)$  menyatakan fungsi distribusi kumulatif empiris dari data. Kriteria pengujianya adalah tolak  $H_0$  apabila  $p_{\text{value}} < \alpha$ , dimana  $\alpha$  merupakan taraf signifikansi [21]. Nilai  $\alpha$  yang digunakan dalam penelitian ini sebesar 0,05.

#### 2.6. Uji Chi-Square

Uji *Chi-Square* dinotasikan sebagai  $\chi^2$  dan dinyatakan sebagai berikut:

$$\chi^2 = \sum_{k=1}^b \frac{(n_k - np_k)^2}{np_k}, \quad (10)$$

dengan  $b$  adalah jumlah kategori dalam data,  $n_k$  merupakan frekuensi observasi (jumlah data nyata) pada kelas ke- $k$ , dan  $np_k$  menyatakan jumlah observasi yang diharapkan pada kategori ke- $k$ . Menggunakan taraf signifikansi  $\alpha = 0,05$ , gagal tolak  $H_0$  apabila  $p_{\text{value}} > \alpha$  [21].

##### 2.6.1. Model Aggregate Loss

*Aggregate loss* dinotasikan sebagai  $S$  diperoleh dari penjumlahan seluruh pembayaran klaim individu yaitu  $S = X_1 + X_2 + \dots + X_N$  dengan  $X$  menyatakan severitas klaim, dan  $N$  menyatakan frekuensi klaim [3]. Bentuk ini dikenal sebagai *collective risk model* dengan asumsi bahwa  $X_1 + X_2 + \dots + X_N$  adalah variabel acak berdistribusi independen dan identik serta tidak bergantung pada variabel  $N$ . Jika  $N$  adalah variabel acak bernilai bulat tak-negatif yang didistribusikan secara independen dari  $X_1, X_2, \dots, X_N$ , maka  $S$  dikatakan berdistribusi *compound*. Rumus mean dari  $S$  diformulasikan sebagai berikut [5]:

$$E(S) = E(N) \times E(X), \quad (11)$$

di mana  $E(N)$  adalah mean dari frekuensi klaim, dan  $E(X)$  adalah mean dari severitas klaim. Pada penelitian ini, masing-masing frekuensi klaim dan severitas klaim dimodelkan dengan dua distribusi. Distribusi Geometrik dan Negatif Binomial untuk memodelkan frekuensi klaim, serta distribusi Gamma dan Lognormal untuk memodelkan severitas klaim. Oleh karena itu, terdapat empat macam model *aggregate loss* yang dihasilkan, yaitu:

- a. Geometrik-Gamma
- b. Geometrik-Lognormal
- c. Negatif Binomial-Gamma
- d. Negatif Binomial-Lognormal

### 2.7. Pendekatan dan Estimasi Bayesian

Estimasi mean dari kerugian acak (*random loss*) yang merupakan fungsi dari  $\Theta$  dihitung menggunakan fungsi kepadatan peluang (*probability density function*, pdf) posterior dari  $\Theta$ . Estimasi ini disebut sebagai *estimasi Bayes*, yang sekaligus berfungsi sebagai prediktor terhadap kerugian di masa mendatang [3]. Komponen utama dalam pendekatan ini meliputi distribusi prior, fungsi likelihood, dan distribusi posterior.

### 2.7.1. Distribusi Prior

Distribusi prior merepresentasikan tingkat ketidakpastian mengenai nilai suatu parameter sebelum data diamati [11]. Misalkan  $X$  melambangkan variabel kerugian acak (*random loss variable*) yang berasal dari frekuensi klaim, severitas klaim, dan *aggregate loss*. Distribusi dari  $X$  bergantung pada parameter  $\theta$ , di mana parameter ini tidak bernilai tetap, tetapi bervariasi antar kelompok risiko. Oleh karena itu,  $\theta$  diperlakukan sebagai realisasi dari suatu variabel acak  $\Theta$ , yang memiliki distribusi statistik disebut *distribusi prior* [3]. Fungsi pdf prior dari  $\Theta$  dinotasikan dengan  $f(\theta)$ .

Parameter  $p$  pada distribusi Geometrik dan Negatif Binomial bernilai kontinu pada interval  $(0, 1)$ , sehingga distribusi prior yang digunakan untuk parameter tersebut adalah distribusi Beta. Fungsi pdf prior dari  $p$  diberikan sebagai berikut [20]:

$$f(p) = p^{a-1}(1-p)^{b-1} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}. \quad (12)$$

Parameter  $r$ ,  $\kappa$ ,  $\theta$ , dan  $\tau$  bernilai positif riil, sehingga distribusi Gamma dapat digunakan sebagai distribusi prior untuk parameter-parameter tersebut. Sementara itu, untuk parameter  $\mu$ , distribusi prior yang digunakan adalah distribusi Normal karena nilainya kontinu pada interval  $(-\infty, \infty)$ . Fungsi pdf prior dari  $\mu$  dinyatakan sebagai berikut:

$$f(\mu) = \sqrt{\frac{\beta}{2\pi}} \exp\left[-\frac{\beta}{2}(\mu - \alpha)^2\right]. \quad (13)$$

### 2.7.2. Fungsi Likelihood

Pdf bersyarat dari  $X$  terhadap parameter  $\theta$  dilambangkan dengan  $f(x|\theta)$ . Misalkan  $X = \{X_1, \dots, X_n\}$  merupakan sampel acak dari  $X$ , dan  $x = (x_1, \dots, x_n)$  adalah realisasinya. Maka pdf bersyarat dari  $X$  diformulasikan sebagai berikut:

$$f(x|\theta) = \prod_{i=1}^n f(x_i|\theta). \quad (14)$$

Fungsi  $f(x|\theta)$  disebut sebagai *fungsi likelihood* [3]. Karena terdapat empat jenis distribusi yang digunakan untuk memodelkan klaim (baik frekuensi maupun severitas klaim), maka diperoleh empat bentuk fungsi likelihood yang berbeda.

### 2.7.3. Distribusi Posterior

Untuk menjaga konsistensi logis, ketidakpastian mengenai nilai  $\theta$  harus diperbarui setiap kali informasi baru diperoleh. Setelah data diamati, distribusi prior diperbarui menjadi distribusi posterior [11]. Berdasarkan sampel data  $x$ , pdf bersyarat dari  $\Theta$  terhadap  $x$  dinyatakan sebagai  $f(\theta|x)$ , yang dirumuskan:

$$f(\theta|x) = \frac{f(x, \theta)}{f(x)}, \quad (15)$$

dengan  $f(x, \theta)$  merupakan pdf bersama antara fungsi likelihood  $X$  dan prior  $\Theta$ , sedangkan  $f(x)$  adalah pdf marginal dari  $X$ . Pdf bersama tersebut dapat dituliskan [5]:

$$f(x, \theta) = f(x|\theta)f(\theta), \quad (16)$$

sedangkan pdf marginal dari  $X$  untuk variabel acak kontinu  $\Theta$  dirumuskan [20]:

$$f(x) = \int_{-\infty}^{\infty} f(x, \theta) d\theta. \quad (17)$$

### 2.7.4. Distribusi Prediktif Posterior

Distribusi prediktif posterior adalah distribusi peluang bersyarat dari suatu pengamatan baru  $\tilde{S}$  yang diberikan data  $(x, n)$ , dinotasikan dengan  $f(\tilde{S}|x, n)$  [5]. Untuk memperoleh distribusi ini, dilakukan simulasi dari  $f(\tilde{S}|x, n)$ , di mana hasil simulasi digunakan untuk mendekati distribusinya. Sebagai contoh, algoritma simulasi untuk model Geometrik–Gamma dilakukan dengan langkah-langkah berikut [11]:

1. Ambil vektor parameter simulasi ke- $k$ , yaitu  $(p_k, \kappa_k, \theta_k)$  dari hasil MCMC.
2. Simulasikan satu nilai  $n$  dari distribusi Geometrik dengan parameter  $p_k$ .
3. Simulasikan sebanyak  $n$  nilai independen dari distribusi Gamma dengan parameter  $(\kappa_k, \theta_k)$ , misalkan nilai-nilai tersebut adalah  $x_1, \dots, x_n$ .
4. Jumlahkan nilai-nilai tersebut, dan ambil hasil penjumlahannya sebagai nilai simulasi ke- $k$  dari distribusi  $f(\tilde{S}|x, n)$ .

Setiap simulasi menghasilkan satu nilai  $\tilde{S}$ . Algoritma ini dapat diterapkan pada seluruh model *aggregate loss* untuk memperoleh estimasi dari distribusi  $f(\tilde{S}|x, n)$ .

### 2.8. Estimasi Parameter

Kompleksitas distribusi posterior menyebabkan nilainya sulit diperoleh secara analitik, sehingga digunakan pendekatan numerik seperti *Markov Chain Monte Carlo* (MCMC). Dalam metode ini dihasilkan barisan sampel acak yang saling berkorelasi, yaitu nilai ke- $j$  dari barisan  $\{\theta_j\}$  disampling dari distribusi peluang yang bergantung pada nilai sebelumnya  $\{\theta_{j-1}\}$ . Autokorelasi muncul akibat dinamika transisi dalam rantai Markov, yang menyebabkan rantai cenderung bertahan di suatu nilai selama beberapa iterasi [22]. Salah satu algoritma dalam metode MCMC adalah *Gibbs Sampling*, dengan langkah-langkah berikut [23]:

1. Tentukan nilai awal  $x^{(0)}$ .
2. Tentukan urutan koordinat yang akan diperbarui, misalnya untuk  $j = 1, \dots, p$ .
3. Pada setiap iterasi ke- $t$ , perbarui setiap komponen  $x_j$  berdasarkan distribusi bersyarat  $f(x_j|x_{-j})$ , sehingga diperoleh sampel baru  $x^{(t)}$ .
4. Ulangi proses ini hingga jumlah iterasi yang ditentukan tercapai.

### 2.9. Simulasi Menggunakan Perangkat Lunak OpenBUGS

Sebelum menjalankan *OpenBUGS*, diperlukan tiga berkas teks berisi kode model, data, dan nilai awal (*inits*). Langkah-langkah penggunaannya adalah sebagai berikut:

1. Menentukan kode model, data, dan inits.
2. Memeriksa spesifikasi model.  
Pemeriksaan dilakukan melalui *specification tool* untuk memastikan tidak ada kesalahan pada kode model dan data. Setelah keduanya valid, lakukan proses *compile*, kemudian periksa kode *inits*.
3. Menentukan parameter yang akan diestimasi.  
Pada penelitian ini, parameter yang diestimasi mencakup pa-

parameter dari distribusi Geometrik, Negatif Binomial, Gamma, dan Lognormal, serta estimasi *aggregate loss* masa depan melalui distribusi prediktif posterior.

#### 4. Menentukan jumlah iterasi dan tahap burn-in.

Iterasi awal dialokasikan sebagai *burn-in* untuk menghilangkan pengaruh nilai awal yang belum stabil. Setelah pengaturan selesai, *OpenBUGS* menjalankan proses simulasi.

#### 5. Menganalisis output hasil simulasi.

Dari hasil simulasi diperoleh statistik deskriptif serta distribusi dari masing-masing parameter dan distribusi prediktif posterior.

### 3. Hasil dan Pembahasan

#### 3.1. Distribusi Frekuensi Klaim

Statistika deskriptif digunakan untuk mengetahui karakteristik data secara umum. Statistika deskriptif untuk data frekuensi klaim dapat dilihat pada **Tabel 2**.

**Tabel 2.** Statistika deskriptif frekuensi klaim

Statistik	Nilai
Minimum	5
Maksimum	970
Rata-rata	279,438
Median	208
Varians	58.374,40

Berdasarkan hasil statistika deskriptif frekuensi klaim pada **Tabel 2**, diperoleh nilai median kurang dari nilai mean (rata-rata). Hal ini menunjukkan bahwa sebagian besar data berada di bawah rata-rata yang mengarah pada distribusi data condong ke kanan (*right-skewed*). Sementara itu, nilai varians melebihi nilai rata-rata mengidentifikasi bahwa data bersifat overdispersi, sehingga distribusi Poisson kurang tepat digunakan karena distribusi Poisson bersifat *equidispersed* [24]. Dengan mempertimbangkan karakteristik overdispersi tersebut, distribusi Binomial Negatif lebih sesuai digunakan untuk memodelkan frekuensi klaim, karena nilai varians lebih besar dari mean [11]. Selain itu, distribusi Geometrik merupakan kasus khusus dari distribusi Binomial Negatif ketika  $r$  yang menyatakan banyaknya sukses pertama sama dengan 1 [5]. Hal ini menjadikan distribusi Geometrik sebagai distribusi alternatif yang sesuai untuk memodelkan data frekuensi klaim. Selanjutnya, dilakukan uji kecocokan distribusi menggunakan uji *Kolmogorov-Smirnov* sesuai **pers. (9)**. Berdasarkan uji tersebut, diperoleh  $p_{value}$  distribusi Geometrik sebesar 0,8123 dan distribusi Binomial Negatif sebesar 0,10541 dengan  $\alpha = 0,05$ . Nilai  $p_{value}$  kedua distribusi tersebut lebih besar dari  $\alpha$  sehingga keputusan yang diambil adalah gagal tolak  $H_0$ . Hal ini mengidentifikasi bahwa distribusi Geometrik dan Binomial Negatif dapat digunakan untuk memodelkan data frekuensi klaim.

#### 3.2. Distribusi Severitas Klaim

Karakteristik data severitas klaim secara umum dapat digambarkan melalui statistika deskriptif. Statistika deskriptif untuk data ini dapat dilihat pada **Tabel 3**

Dari **Tabel 3**, diketahui nilai rata-rata lebih besar dari nilai median yang menunjukkan bahwa distribusi data cenderung tidak simetris dan condong ke kanan (*right-skewed*), sehingga distribusi Normal tidak cocok digunakan, karena data berdistribusi Normal memiliki bentuk pdf yang simetris [25]. Nilai varians yang

**Tabel 3.** Statistika deskriptif severitas klaim

Statistik	Nilai
Minimum	153,62
Maksimum	797,80
Rata-rata	276,352
Median	250,525
Varians	12.198,30

tinggi mengidentifikasi keragaman penyebaran yang tinggi pada data. Berdasarkan karakteristik tersebut, diduga data severitas klaim berdistribusi Gamma dan Lognormal [26]. Untuk memastikan kebenaran dugaan tersebut, dilakukan uji kecocokan distribusi menggunakan uji *Chi-Square*. Menggunakan **pers. (10)**, diperoleh  $p_{value}$  dari masing-masing distribusi Gamma dan Lognormal secara berurutan sebesar 0,09573 dan 0,41378. Berdasarkan hasil tersebut, nilai  $p_{value}$  lebih besar dari  $\alpha = 0,05$ , sehingga keputusan yang diambil adalah gagal tolak  $H_0$ . Dengan demikian, severitas klaim dapat dimodelkan dengan distribusi Gamma dan Lognormal.

#### 3.3. Distribusi Prior

Data dalam penelitian ini tidak memiliki informasi prior, sehingga digunakan *weakly informative prior*. Pemilihan *weakly informative prior* dalam penelitian ini dilakukan untuk menghindari hasil distribusi posterior yang ekstrem atau tidak masuk akal [27]. Berdasarkan pertimbangan tersebut, digunakan prior Beta (2, 2) prior Gamma (2, 0,1), dan prior Normal (0, 2,5) [28]. Standar deviasi dari masing-masing prior tersebut secara berurutan adalah 0,22360, 14,14213, dan 1,58113.

#### 3.4. Fungsi Likelihood

Menggunakan rumus fungsi *likelihood* pada **pers. (14)**, pmf pada **pers. (1)** dan **pers. (3)**, serta pdf pada **pers. (5)** dan **pers. (7)**, diperoleh fungsi *likelihood* dari variabel acak berdistribusi Geometrik, Negatif Binomial, Gamma, dan Lognormal yang disajikan pada **Tabel 4**.

Hasil fungsi *likelihood* pada **Tabel 4** menunjukkan bentuk matematis yang digunakan untuk menghitung kemungkinan munculnya data klaim asuransi yang diamati, jika parameter tertentu digunakan dalam suatu model. Melalui fungsi *likelihood* yang diperoleh ini, parameter distribusi diasumsikan acak dan akan diestimasi menggunakan pendekatan Bayesian. Estimasi parameter yang diperoleh diharapkan dapat mencerminkan karakteristik data secara akurat, baik data frekuensi klaim maupun severitas klaim.

#### 3.5. Distribusi Posterior

Sebelum menentukan distribusi posterior, terlebih dahulu dilakukan perhitungan untuk memperoleh pdf distribusi bersama dan distribusi marginal. Pdf distribusi bersama diperoleh dengan mengalikan fungsi *likelihood* pada **Tabel 4** dengan pdf priornya, sesuai **pers. (16)**. Berdasarkan fungsi *likelihood* pada **Tabel 4** serta pdf prior Beta pada **pers. (12)**, prior Gamma pada **pers. (6)**, dan prior Normal pada **pers. (13)**, hasil dari pdf distribusi bersama dapat dilihat pada **Tabel 5**.

**Tabel 5** menunjukkan pdf distribusi bersama untuk masing-masing model. Pdf ini menggabungkan data klaim yang diamati dengan informasi prior. Selanjutnya, pdf distribusi margi-

**Tabel 4.** Fungsi likelihood

Distribusi	Fungsi Likelihood
Geometrik	$f(n   p) = p^{32} (1 - p)^{(\sum_{i=1}^{32} n_i) - 32}$
Binomial Negatif	$f(n   p, r) = p^{32r} \left(\frac{1}{\Gamma(r)}\right)^{32} (1 - p)^{\sum_{i=1}^{32} n_i} \prod_{i=1}^{32} \frac{\Gamma(n_i + r)}{\Gamma(n_i + 1)}$
Gamma	$f(x   \kappa, \theta) = \frac{\theta^{32\kappa}}{[\Gamma(\kappa)]^{32}} \exp\left(-\theta \sum_{i=1}^{32} x_i\right) \prod_{i=1}^{32} x_i^{\kappa-1}$
Lognormal	$f(x   \mu, \tau) = \left(\frac{\tau}{2\pi}\right)^{16} \left(\prod_{i=1}^{32} \frac{1}{x_i}\right) \exp\left(-\frac{\tau}{2} \sum_{i=1}^{32} (\log x_i - \mu)^2\right)$

**Tabel 5.** Distribusi bersama

Distribusi	Pdf Distribusi Bersama
Geometrik dengan Prior Beta	$f(n, p) = \frac{\Gamma(4)}{\Gamma(2)^2} p^{33} (1 - p)^{(\sum_{i=1}^{32} n_i) - 31}$
Binomial Negatif dengan Prior Beta dan Gamma	$f(n, p, r) = p^{32r+1} (1 - p)^{(\sum_{i=1}^{32} n_i) + 1} \left(\prod_{i=1}^{32} \frac{\Gamma(n_i + r)}{\Gamma(n_i + 1)}\right) \frac{\Gamma(4)0.01re^{-0.1r}}{(\Gamma(2))^3(\Gamma(r))^{32}}$
Gamma dengan Prior Gamma	$f(x, \kappa, \theta) = \frac{\theta^{32\kappa+1} e^{-\theta(\sum_{i=1}^{32} x_i + 0.1)} 0.0001\kappa e^{-0.01\kappa}}{[\Gamma(\kappa)]^{32} (\Gamma(2))^2} \prod_{i=1}^{32} x_i^{\kappa-1}$
Lognormal dengan Prior Normal dan Gamma	$f(x, \mu, \tau) = \frac{(2.5)^{1/2} 0.01}{2\pi^{33/2} \Gamma(2)} \tau^{17} \left(\prod_{i=1}^{32} \frac{1}{x_i}\right) \exp\left(-\frac{\tau}{2} \sum_{i=1}^{32} (\log x_i - \mu)^2 - 1.25\mu^2 - 0.1\tau\right)$

**Tabel 6.** Distribusi marginal

Distribusi	Pdf Distribusi Marginal
Geometrik	$f(n) = \frac{\Gamma(4)}{\Gamma(2)^2} \frac{\Gamma(34)\Gamma\left(\left(\sum_{i=1}^{32} n_i\right) - 30\right)}{\Gamma\left(\left(\sum_{i=1}^{32} n_i\right) + 4\right)}$
Binomial Negatif	$f(n) = \frac{\Gamma\left(\sum_{i=1}^{32} n_i + 2\right) \Gamma(4) 0.01}{\left(\prod_{i=1}^{32} \Gamma(n_i + 1)\right) (\Gamma(2))^3} \int_0^\infty \frac{\Gamma(32r + 2) \left(\prod_{i=1}^{32} \Gamma(n_i + r)\right) r e^{-0.1r}}{\Gamma\left(32r + 4 + \sum_{i=1}^{32} n_i\right) (\Gamma(r))^{32}} dr$
Gamma	$f(x) = \frac{0.0001}{(\Gamma(2))^2} \int_0^\infty \frac{\Gamma(32\kappa + 2)}{\left(\sum_{i=1}^{32} x_i + 0.1\right)^{32\kappa+2}} \frac{\kappa e^{-0.1\kappa}}{(\Gamma(\kappa))^{32}} \prod_{i=1}^{32} x_i^{\kappa-1} d\kappa$
Lognormal	$f(x) = \frac{(2.5)^{1/2} 0.01}{2\pi^{33/2} \Gamma(2)} \left(\prod_{i=1}^{32} \frac{1}{x_i}\right) \int_0^\infty \int_{-\infty}^\infty \tau^{17} \exp\left(-\frac{\tau}{2} \sum_{i=1}^{32} (\log x_i - \mu)^2 - 1.25\mu^2 - 0.1\tau\right) d\mu d\tau$

nal diperoleh berdasarkan pdf distribusi bersama menggunakan pers. (17), yang disajikan pada Tabel 6.

Distribusi marginal pada Tabel 6 diperoleh dengan mengintegrasikan parameter tertentu dari pdf distribusi bersama, sehingga parameter tersebut tereliminasi dari model. Berdasarkan pdf distribusi bersama pada Tabel 5, dan distribusi marginal pada Tabel 6, dihasilkan distribusi posterior menggunakan pers. (15) yang disajikan pada Tabel 7.

Distribusi posterior pada Tabel 7 menunjukkan pembaruan informasi terhadap parameter distribusi setelah mempertimbangkan data klaim yang diamati. Distribusi ini diperoleh dengan menggabungkan distribusi prior dengan data klaim yang diamati melalui fungsi likelihood, sehingga akan diperoleh estimasi parameter yang lebih mencerminkan karakteristik empiris data klaim asuransi, baik frekuensi klaim maupun severitas klaim. Selanjutnya, karena distribusi posterior yang dihasilkan tidak

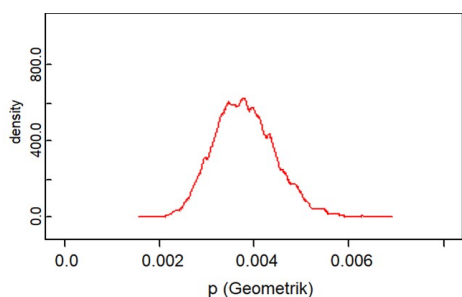
dapat dihitung secara analitik, maka digunakan metode numerik untuk menghitungnya. Pada penelitian ini, metode numerik yang digunakan adalah metode *Markov Chain Monte Carlo* (MCMC) dengan algoritma *Gibbs-Sampling* yang proses perhitungannya dibantu *software* OpenBUGS.

### 3.6. Ekspektasi Aggregate Loss

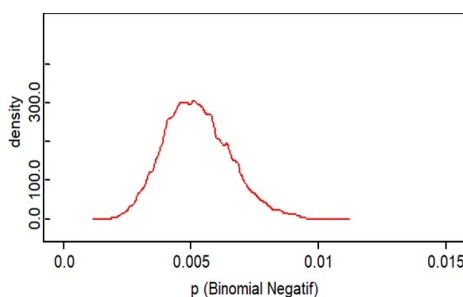
Simulasi posterior dilakukan menggunakan *software* OpenBUGS, yang menerapkan metode MCMC dengan algoritma *Gibbs-Sampling*. Dasar penerapan metode ini telah dijelaskan pada bagian 2.8. Terdapat total 11000 iterasi yang dijalankan, dengan 1000 iterasi awal dialokasikan sebagai tahap *burn-in* untuk mengeliminasi nilai awal (*initial values*) yang belum stabil. Selanjutnya, 10000 iterasi digunakan sebagai sampel aktual untuk membentuk estimasi distribusi posterior dari parameter-parameter distribusi. Menggunakan langkah-langkah penggunaa-

**Tabel 7.** Distribusi posterior

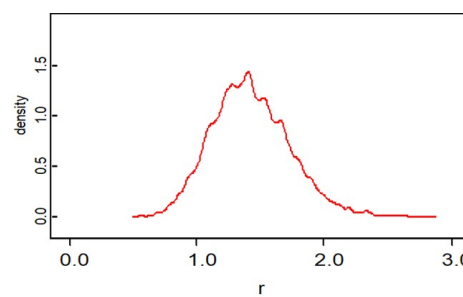
Distribusi	Pdf Distribusi Marginal
Geometrik	$f(p   n) = \frac{p^{33} (1 - p)^{(\sum_{i=1}^{32} n_i) - 31} \Gamma(\sum_{i=1}^{32} n_i + 4)}{\Gamma(34) \Gamma(\sum_{i=1}^{32} n_i - 30)}$
Binomial Negatif	$f(p, r   n) = \frac{p^{32r+1} (1 - p)^{(\sum_{i=1}^{32} n_i) + 1} \left( \prod_{i=1}^{32} \frac{\Gamma(n_i + r)}{\Gamma(n_i + 1)} \right) \frac{r e^{-0.1r}}{(\Gamma(r))^{32}}}{\frac{\Gamma(\sum_{i=1}^{32} n_i + 2)}{\prod_{i=1}^{32} \Gamma(n_i + 1)} \int_0^\infty \frac{\Gamma(32r + 2) \left( \prod_{i=1}^{32} \Gamma(n_i + r) \right) \frac{r e^{-0.1r}}{(\Gamma(r))^{32}}}{\Gamma(32r + 4 + \sum_{i=1}^{32} n_i)} dr}$
Gamma	$f(\kappa, \theta   x) = \frac{\theta^{32\kappa+1} \exp\left[-\theta \left(\sum_{i=1}^{32} x_i + 0.1\right)\right] \kappa e^{-0.01\kappa} \prod_{i=1}^{32} x_i^{\kappa-1} / (\Gamma(\kappa))^{32}}{\int_0^\infty \frac{\Gamma(32\kappa + 2) \kappa e^{-0.1\kappa} \prod_{i=1}^{32} x_i^{\kappa-1}}{\left(\sum_{i=1}^{32} x_i + 0.1\right)^{32\kappa+2} (\Gamma(\kappa))^{32}} d\kappa}$
Lognormal	$f(\mu, \tau   x) = \frac{\tau^{17} \exp\left(-\frac{\tau}{2} \sum_{i=1}^{32} (\log x_i - \mu)^2 - 1.25 \mu^2 - 0.1 \tau\right)}{\int_0^\infty \int_{-\infty}^\infty \tau^{17} \exp\left(-\frac{\tau}{2} \sum_{i=1}^{32} (\log x_i - \mu)^2 - 1.25 \mu^2 - 0.1 \tau\right) d\mu d\tau}$



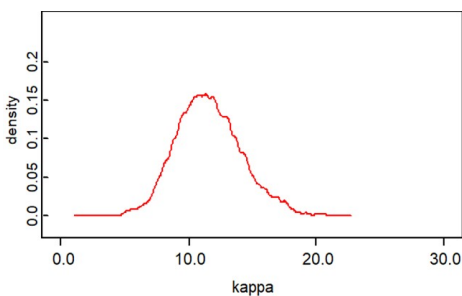
(a) p (geometrik)



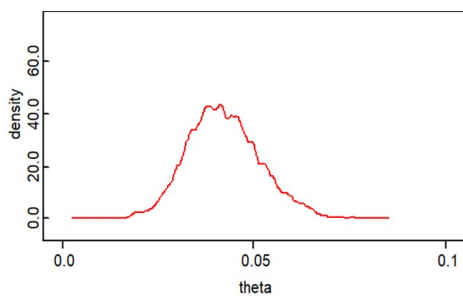
(b) p (binomial negatif)



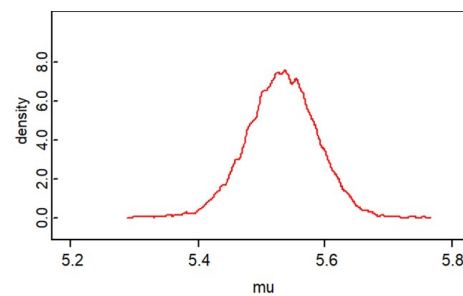
(c) r



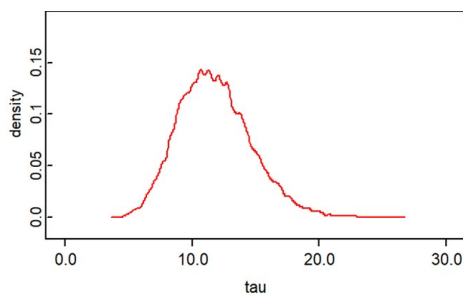
(d) κ



(e) θ



(f) μ



(g) τ

**Gambar 1.** Distribusi posterior parameter

an software OpenBUGS pada bagian 2.9, diperoleh **Tabel 8** dan **Gambar 1** yang menyajikan hasil simulasi posterior parameter menggunakan software OpenBUGS.

**Tabel 8.** Posterior parameter

Parameter	Mean	Standar Deviasi	95% Credible Interval
$p$ (Geometrik)	0,003805	0,000656	(0,002648, 0,00517)
$p$ (Binomial Negatif)	0,005277	0,001321	(0,002951, 0,008132)
$r$	1,424	0,3035	(0,8832, 2,076)
$\kappa$	11,63	2,565	(7,021, 17,12)
$\theta$	0,04229	0,009487	(0,02523, 0,06255)
$\mu$	5,533	0,05443	(5,423, 5,637)
$\tau$	11,83	2,85	(6,851, 17,93)

Dari hasil simulasi pada **Tabel 8**, dapat dilihat bahwa setiap distribusi posterior dari parameter memiliki standar deviasi yang relatif rendah. Dalam semua distribusi, standar deviasi dari posterior parameter secara signifikan lebih kecil dibandingkan standar deviasi dari priornya, yang sejalan dengan hasil pada penelitian sebelumnya oleh Sunil Hariram Joshi [11]. Standar deviasi posterior parameter  $p$  untuk distribusi Geometrik dan Binomial Negatif kurang dari 0,22360, posterior parameter  $r$ ,  $\kappa$ ,  $\theta$  dan  $\tau$  kurang dari 14,14213, serta posterior parameter  $\mu$  kurang dari 1,58113. Hal ini menunjukkan peningkatan yang signifikan dalam kepastian keyakinan posterior dibandingkan dengan keyakinan sebelumnya (prior), yang juga ditunjukkan dari *credible interval* 95% yang sempit, sehingga perusahaan asuransi dapat menetapkan premi dan dana cadangan dengan lebih akurat.

Selanjutnya, **Gambar 1** memperlihatkan distribusi posterior parameter frekuensi dan severitas klaim. Distribusi posterior parameter frekuensi klaim ditunjukkan pada **Gambar 1a**, **Gambar 1b**, dan **Gambar 1c**. **Gambar 1a** dan **1b** menampilkan distribusi posterior parameter  $p$  pada model Geometrik dan Binomial Negatif, sedangkan **Gambar 1c** menampilkan parameter  $r$ . Ketiga distribusi tersebut tampak rapat di sekitar nilai mean-nya, menunjukkan ketidakpastian estimasi yang rendah serta variasi banyaknya klaim yang relatif kecil. Sementara itu, pada severitas klaim, **Gambar 1d** dan **1e** menunjukkan distribusi posterior parameter  $\kappa$  dan  $\theta$  pada model Gamma, sedangkan **Gambar 1f** dan **1g** menampilkan posterior parameter  $\mu$  dan  $\tau$  pada model Lognormal. Parameter  $\theta$  dan  $\mu$  terlihat terkonsentrasi di sekitar nilai mean-nya, menandakan estimasi parameter relatif stabil, sedangkan distribusi posterior parameter  $\kappa$  dan  $\tau$  memiliki sebaran yang lebih lebar, mencerminkan ketidakpastian yang lebih besar dalam estimasi kedua parameter tersebut. Ketidakpastian pada parameter  $\kappa$  dan  $\tau$  dapat meningkatkan risiko terjadinya *underpricing* atau *overpricing*, sehingga perusahaan asuransi perlu berhati-hati dalam menetapkan premi agar sesuai dengan risiko sebenarnya.

Berdasarkan hasil estimasi posterior dari parameter pada **Tabel 8**, menggunakan **pers. (2)**, **pers. (4)**, **pers. (6)**, dan **pers. (8)**, diperoleh nilai ekspektasi dari frekuensi klaim yang berdistribusi Geometrik sebesar \$262,81208, Binomial Negatif sebesar \$268,42629, serta severitas klaim yang berdistribusi Gamma sebesar \$275,00591, dan Lognormal sebesar \$299,48421. Dari hasil tersebut, nilai ekspektasi dari *aggregate loss* diperoleh menggunakan rumus pada **pers. (11)**, yang disajikan pada **Tabel 9**.

**Tabel 9** menunjukkan nilai espektasi *aggregate loss* dari kombinasi model frekuensi klaim dan severitas klaim. Hasil estimasi menunjukkan bahwa model Binomial Negatif-Lognormal

**Tabel 9.** Ekspektasi *Aggregate Loss* (Poundsterling)

Model <i>Aggregate Loss</i>	Ekspektasi <i>Aggregate Loss</i>
Geometrik-Gamma	72274,87521
Geometrik-Lognormal	78708,06815
Binomial Negatif-Gamma	73818,81614
Binomial Negatif-Lognormal	80389,43540

memberikan nilai ekspektasi *aggregate loss* tertinggi sebesar \$80389,43540, sedangkan model Geometrik-Gamma menghasilkan ekspektasi terendah sebesar \$72274,87521.

### 3.7. Distribusi Prediktif Posterior

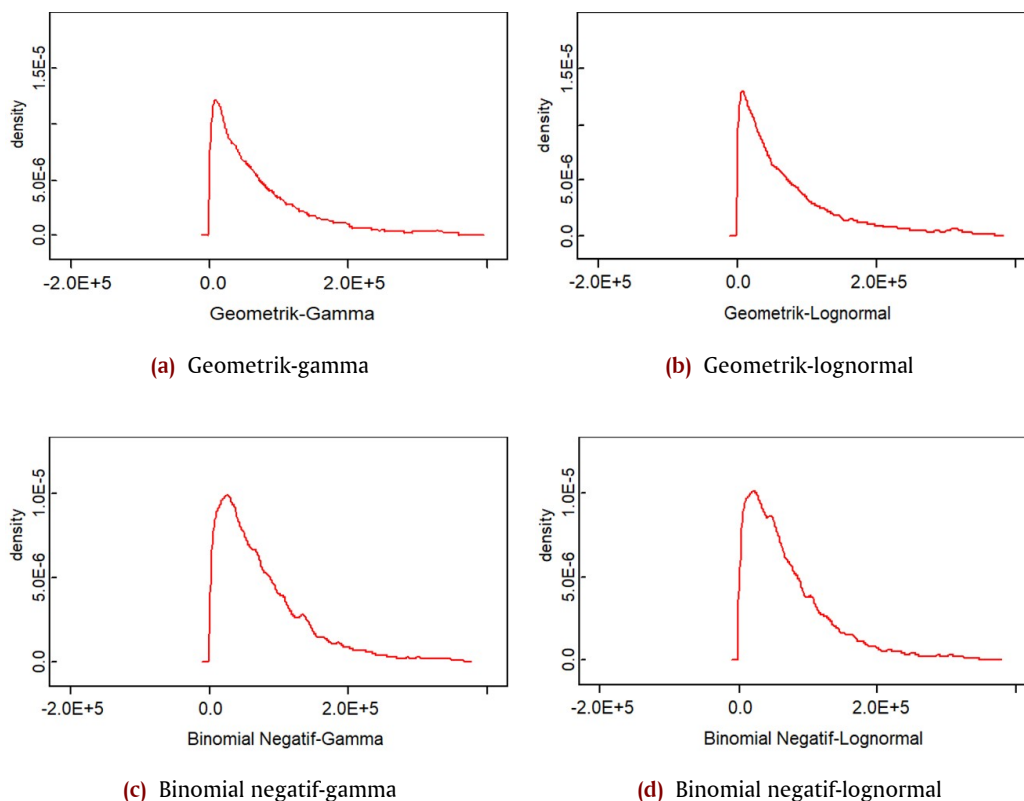
Selanjutnya untuk mengestimasi *aggregate loss* di masa depan melalui distribusi prediktif posterior, dilakukan simulasi pada OpenBUGS dengan algoritma pada bagian 2.7.4, yang hasilnya dapat dilihat pada **Gambar 2** dan **Tabel 10**.

**Tabel 10.** Prediktif posterior (Poundsterling)

Model	Mean	Standar Deviasi	95% Credible Interval
Geometrik-Gamma	73880,0	71770,0	(1984,0, 281900,0)
Geometrik-Lognormal	70500,0	69050,0	(1956,0, 268600,0)
Binomial Negatif-Gamma	75270,0	63960,0	(4365,0, 248600,0)
Binomial Negatif-Lognormal	72240,0	62720,0	(3582,0, 247400,0)

**Gambar 2** menunjukkan distribusi prediktif posterior *aggregate loss* untuk empat kombinasi model frekuensi dan severitas klaim. **Gambar 2a** menampilkan model Geometrik-Gamma, dimana puncak distribusi berada pada nilai kerugian rendah dengan ekor kanan yang panjang, menandakan adanya peluang kecil untuk kerugian ekstrem. Model Geometrik-Lognormal pada **Gambar 2b** memiliki karakteristik serupa, namun dengan ekor kanan yang lebih berat, mengindikasikan peluang terjadinya kerugian ekstrem yang lebih besar dibandingkan model Geometrik-Gamma. **Gambar 2c** memperlihatkan model Binomial Negatif-Gamma dengan puncak yang lebih lebar, menunjukkan variasi prediksi *aggregate loss* yang lebih tinggi. Sementara itu, model Binomial Negatif-Lognormal pada **Gambar 2d** menampilkan puncak yang juga lebar dengan ekor kanan yang lebih berat, menunjukkan terjadinya peluang kerugian ekstrem yang lebih besar. Secara keseluruhan, dapat dilihat bahwa keempat model distribusi memiliki puncak dengan kisaran kerugian yang rendah. Hal ini berarti estimasi *aggregate loss* yang diprediksi oleh mayoritas simulasi tidak terlalu besar. Namun, setiap model distribusi memiliki ekor panjang ke kanan yang menandakan bahwa meskipun mayoritas simulasi tidak terlalu besar, terdapat peluang signifikan untuk kerugian yang besar akibat adanya kejadian ekstrem. Oleh karenanya, perusahaan asuransi perlu mempertimbangkan adanya ekor distribusi ini dengan menetapkan premi yang cukup untuk menutupi kemungkinan kerugian ekstrem serta menyiapkan dana cadangan yang memadai untuk menjaga solvabilitas perusahaan.

Hasil simulasi pada **Tabel 10** menunjukkan nilai mean, standar deviasi, dan *credible interval* 95% dari distribusi prediktif posterior berdasarkan kombinasi model frekuensi dan severitas klaim. Nilai mean ini menggambarkan nilai rata-rata *aggregate loss* yang diprediksi di masa depan. Model Binomial Negatif-Gamma menghasilkan estimasi mean dari prediktif posterior tertinggi sebesar \$75270,0, sedangkan model Geometrik-Lognormal memberikan estimasi terendah sebesar \$70500,0. Sementara itu, model Bino-



Gambar 2. Distribusi prediktif posterior

mial Negatif-Lognormal memiliki standar deviasi terkecil, yaitu \$62720,0 yang mengindikasikan bahwa estimasi yang dihasilkan oleh model ini relatif lebih stabil dikarenakan variabilitasnya yang lebih rendah. Dalam konteks asuransi, pemilihan model *aggregate loss* ini disesuaikan dengan target pasar yang ditetapkan oleh perusahaan. Misalnya, apabila perusahaan menargetkan nasabah (tertanggung) dari segmen dengan risiko rendah, maka model Geometrik-Lognormal lebih sesuai digunakan untuk memodelkan *aggregate loss*. Hal ini bertujuan agar premi yang ditetapkan tetap terjangkau tanpa mengorbankan margin risiko perusahaan. Sebaliknya, jika target pasar perusahaan adalah segmen dengan risiko tinggi, maka model Binomial Negatif-Gamma perlu dipertimbangkan agar dana cadangan dapat dialokasikan dengan efektif.

#### 4. Kesimpulan

Penelitian ini menunjukkan bahwa pendekatan Bayesian dengan metode *Markov Chain Monte Carlo* (MCMC) menggunakan algoritma *Gibbs Sampling* serta implementasinya melalui perangkat lunak *OpenBUGS* dapat digunakan secara efektif untuk mengestimasi *aggregate loss*. Berdasarkan hasil simulasi, diperoleh bahwa estimasi nilai *aggregate loss* pada asuransi mobil penumpang pribadi dipengaruhi oleh pemilihan model frekuensi dan severitas klaim. Model *Negatif Binomial-Gamma* menghasilkan estimasi nilai prediktif posterior *aggregate loss* tertinggi sebesar £75,270, sedangkan model *Geometrik-Lognormal* memberikan estimasi terendah sebesar £70,500. Sementara itu, model dengan standar deviasi terkecil adalah *Negatif Binomial-Lognormal*, yaitu sebesar £62,720. Hasil penelitian ini mendukung penerapan pendekatan Bayesian dalam estimasi *aggregate loss* dan memberikan

implikasi praktis bahwa pemilihan model dapat disesuaikan dengan karakteristik serta target pasar perusahaan asuransi. Dengan demikian, penentuan dana cadangan dan penetapan premi dapat dilakukan secara lebih efisien, sehingga potensi kerugian dapat diminimalkan. Selain itu, penelitian ini dapat menjadi dasar bagi studi lanjutan yang bertujuan memperluas cakupan kajian, khususnya melalui validasi prediktif secara kuantitatif menggunakan metode seperti *Posterior Predictive Check* (PPC) untuk menilai ketepatan dan keakuratan model yang digunakan.

**Kontribusi Penulis.** Melati Sinta Nurdanita: Konseptualisasi, metodologi, perangkat lunak, analisis formal, investigasi, penulisan–persiapan draf asli, visualisasi. Azizah: Konseptualisasi, metodologi, validasi, kurasi data, penulisan–tinjauan dan penyuntingan, supervisi. Semua penulis telah membaca dan menyetujui versi manuskrip yang diterbitkan.

**Ucapan Terima Kasih.** Penulis mengucapkan terima kasih kepada semua pihak yang telah berkontribusi dalam penelitian dan penyusunan artikel ini. Penulis juga mengucapkan terima kasih kepada editor dan reviewer atas segala masukan serta dukungannya dalam menyempurnakan tulisan ini.

**Pembiayaan.** Penelitian ini tidak menerima pendanaan eksternal.

**Konflik Kepentingan.** Para penulis menyatakan tidak ada konflik kepentingan yang terkait dengan artikel ini.

**Ketersediaan Data.** Data tersedia dan dapat diakses di website *Kaggle*.

## Referensi

- [1] C. B. Al Husaini, "Pemahaman Resiko Dan Manajemen Resiko," *Jurnal Nuansa: Publikasi Ilmu Manajemen dan Ekonomi Syariah*, vol. 1, no. 3, pp. 318–325, 2023, doi: [10.61132/nuansa.v1i3%20September.272](https://doi.org/10.61132/nuansa.v1i3%20September.272).
- [2] D. Rudianto and A. H. Dewangga, "Risiko Keuangan Pengaruhnya Terhadap Tingkat Profitabilitas Pada Perusahaan Asuransi Umum di BEI," *Sikap*, vol. 6, no. 1, p. 64, 2021, doi: [10.61132/nuansa.v1i3.272](https://doi.org/10.61132/nuansa.v1i3.272).
- [3] Yiu-Kuen Tse, *Nonlife Actuarial Models Theory, Methods and Evaluation*. New York: Cambridge, 2009.
- [4] S. Chen, Z. Wang, and M. Kelly, "Aggregate Loss Model With Poisson-Tweedie Frequency," *Big Data and Information Analytics*, vol. 6, no. 0, pp. 56–73, 2021, doi: [10.3934/bdia.2021005](https://doi.org/10.3934/bdia.2021005).
- [5] H. H. P. G. E. W. Stuart A. Klugman, *Loss Models*, 4th ed. Hoboken, NJ, USA: Wiley, 2012.
- [6] F. I. Yusuf, P. Z. R. Adi, and T. E. T. Saragih, "Penerapan Collective Risk Model dalam Penentuan Premi Asuransi Bencana Alam," *Euler: Jurnal Ilmiah Matematika, Sains dan Teknologi*, vol. 12, no. 2, pp. 213–219, Dec. 2024, doi: [10.37905/euler.v12i2.28632](https://doi.org/10.37905/euler.v12i2.28632).
- [7] S. Rahmawati, A. Z. Adib, and V. Rusyn, "Aggregate Loss Models to Calculate Risk Measures," *Operations Research: International Conference Series*, vol. 5, no. 2, pp. 38–45, 2024, doi: [10.47194/orics.v5i2.314](https://doi.org/10.47194/orics.v5i2.314).
- [8] A. K. Mutaqin and K. Sa'diah, "The Determination of the Aggregate Loss Distribution Through the Numerical Inverse of the Characteristic Function Using the Trapezoidal Quadrature Rule," *Desimal: Jurnal Matematika*, vol. 4, pp. 339–348, 2021, doi: [10.24042/djm](https://doi.org/10.24042/djm).
- [9] J. Pek and T. Van Zandt, "Frequentist and Bayesian Approaches to Data Analysis: Evaluation and Estimation," *Sage Journals*, vol. 19, no. 1, pp. 21–35, Mar. 2020, doi: [10.1177/1475725719874542](https://doi.org/10.1177/1475725719874542).
- [10] F. Yanuar, R. Febriyuni, and I. R. HG, "Bayesian Generalized Self Method to Estimate Scale Parameter of Invers Rayleigh Distribution," *Cauchy: Jurnal Matematika Murni dan Aplikasi*, vol. 6, no. 4, pp. 270–278, May 2021, doi: [10.18860/ca.v6i4.11482](https://doi.org/10.18860/ca.v6i4.11482).
- [11] S. Joshi, "A Bayesian Analysis of Aggregate Loss Models," Curtin University, 2023.
- [12] M. Deng and M. S. Aminzadeh, "Bayesian Inference for the Loss Models via Mixture Priors," *Risks*, vol. 11, no. 9, Sep. 2023, doi: [10.3390/risks11090156](https://doi.org/10.3390/risks11090156).
- [13] R. R. M. Tajuddin and N. Ismail, "Frequentist And Bayesian Zero-Inflated Regression Models On Insurance Claim Frequency: A Comparison Study Using Malaysian Motor Insurance Data," *Malaysian Journal of Science*, vol. 41, no. 2, pp. 16–29, Jun. 2022, doi: [10.22452/MJS.VOL41NO2.2](https://doi.org/10.22452/MJS.VOL41NO2.2).
- [14] G. L. Jones and Q. Qin, "Markov Chain Monte Carlo in Practice," *Annual Review of Statistics and Its Application*, vol. 55, p. 50, 2025, doi: [10.1146/annurev-statistics-040220](https://doi.org/10.1146/annurev-statistics-040220).
- [15] Q. Qin and G. L. Jones, "Convergence Rates of Two-Component MCMC Samplers," *Bernoulli*, vol. 28, no. 2, pp. 859–885, May 2022, doi: [10.3150/21-BEJ1369](https://doi.org/10.3150/21-BEJ1369).
- [16] R. P. Desiresta, F. Firdaniza, and K. Parmikanti, "Estimasi Parameter Model Volatilitas Stokastik dengan Metode Bayesian Rantai Markov Monte Carlo untuk Memprediksi Return Saham," *Jurnal Matematika Integratif*, vol. 17, no. 2, p. 73, Jan. 2022, doi: [10.24198/jmi.v17.n2.34805.73-83](https://doi.org/10.24198/jmi.v17.n2.34805.73-83).
- [17] P. J. L. Pierre-Olivier Goffard, "Approximate Bayesian Computations to Fit and Compare," *Insurance: Mathematics and Economics*, vol. 100, pp. 350–371, 2021, doi: [10.1016/j.insmatheco.2021.06.002](https://doi.org/10.1016/j.insmatheco.2021.06.002).
- [18] M. C. Ausín, J. M. Vilar, R. Cao, and C. González-Fragueiro, "Bayesian Analysis of Aggregate Loss Models," *Math Finance*, vol. 21, no. 2, pp. 257–279, Apr. 2011, doi: [10.1111/j.1467-9965.2010.00428.x](https://doi.org/10.1111/j.1467-9965.2010.00428.x).
- [19] E. Baek and J. M. Ferron, "Bayesian Analysis for Multiple-Baseline Studies Where the Variance Differs Across Cases in OpenBUGS," *Developmental Neurorehabilitation*, vol. 24, no. 2, pp. 130–143, 2021, doi: [10.1080/17518423.2020.1858455](https://doi.org/10.1080/17518423.2020.1858455).
- [20] Susiswo, *Pengantar Statistika Matematis*, 1st ed. Malang, Indonesia: Universitas Negeri Malang, 2017.
- [21] T. Rahmawati and D. Susanti, "Determining Pure Premium of Motor Vehicle Insurance with Generalized Linear Models (GLM)," *International Journal of Quantitative Research and Modeling*, vol. 4, no. 4, pp. 207–214, 2023, doi: [10.46336/ijqrm.v4i4.492](https://doi.org/10.46336/ijqrm.v4i4.492).
- [22] S. La Cagnina, C. Grunwald, T. Janßen, K. Kröninger, and S. Schumann, "Phase Space Sampling with Markov Chain Monte Carlo Methods," preprint, Dec. 2024, doi: [10.48550/arXiv.2412.12963](https://doi.org/10.48550/arXiv.2412.12963).
- [23] D. Luengo, L. Martino, M. Bugallo, V. Elvira, and S. Suärkkä, "A Survey of Monte Carlo Methods for Parameter Estimation," Jul. 2021, doi: [10.1186/s13634-020-00675-6](https://doi.org/10.1186/s13634-020-00675-6).
- [24] Ö. İ. Güneri and B. Durmuş, "Models for Overdispersion Count Data with Generalized Distribution: An Application to Parasites Intensity," *Journal of New Theory*, pp. 48–61, 2021, doi: [10.20527/epsilon.v18i2.13872](https://doi.org/10.20527/epsilon.v18i2.13872).
- [25] E. Gómez-Déniz, E. Calderín-Ojeda, and H. W. Gómez, "Symmetric and Asymmetric Distributions: Theoretical Developments and Applications III," MDPI, Oct. 2022, doi: [10.3390/sym14102143](https://doi.org/10.3390/sym14102143).
- [26] A. S. Lestia and M. Idris, "Pendekatan Markov Chain Monte Carlo (MCMC) Metropolis-Hastings Pada Pemodelan Klaim Asuransi Kesehatan," *Epsilon: Jurnal Matematika Murni dan Terapan*, vol. 18, no. 2, pp. 165–177, 2024, doi: [10.20527/epsilon.v18i2.13872](https://doi.org/10.20527/epsilon.v18i2.13872).
- [27] A. Gelman *et al.*, *Bayesian Data Analysis*, 3rd ed. New York, NY, USA: Taylor & Francis, 2013.
- [28] G. Jehli, "Setting Reasonable Priors for Computational Modeling," 2024. [Online]. Available: <https://www.gingjehli.com/single-post/choosing-effective-samplers-and-setting-priors/>. [Accessed: 08-Jul-2025].