

**SKRIPSI**

**ANALISIS PERBANDINGAN MODEL LINEAR  
TERGENERALISASI DAN MODEL ADITIF  
TERGENERALISASI UNTUK PENCADANGAN KLAIM  
ASURANSI DENGAN PENCILAN**



**BIQYTOFA**

**NPM: 6161901091**

**PROGRAM STUDI MATEMATIKA  
FAKULTAS TEKNOLOGI INFORMASI DAN SAINS  
UNIVERSITAS KATOLIK PARAHYANGAN  
2023**

**FINAL PROJECT**

**COMPARISON ANALYSIS OF GENERALIZED LINEAR  
MODEL AND GENERALIZED ADDITIVE MODEL FOR  
CLAIMS RESERVE WITH OUTLIER**



**BIQYTOFA**

**NPM: 6161901091**

**DEPARTMENT OF MATHEMATICS  
FACULTY OF INFORMATION TECHNOLOGY AND SCIENCES  
PARAHYANGAN CATHOLIC UNIVERSITY  
2023**

# LEMBAR PENGESAHAN

## ANALISIS PERBANDINGAN MODEL LINEAR TERGENERALISASI DAN MODEL ADITIF TERGENERALISASI UNTUK PENCADANGAN KLAIM ASURANSI DENGAN PENCILAN

Biqytofa

NPM: 6161901091

Bandung, 3 Agustus 2023

Menyetujui,

Pembimbing 1



Maria Anestasia, M.Si., M.Act.Sc.

Pembimbing 2



Rizky Reza Fauzi, D.Phil.Math.

Ketua Penguji



Felivia Kusnadi, M.Act.Sc.

Anggota Penguji



Dr. Ferry Jaya Permana

Mengetahui,

Ketua Program Studi



Dr. Livia Owen

## PERNYATAAN

Dengan ini saya yang bertandatangan di bawah ini menyatakan bahwa skripsi dengan judul:

### **ANALISIS PERBANDINGAN MODEL LINEAR TERGENERALISASI DAN MODEL ADITIF TERGENERALISASI UNTUK PENCADANGAN KLAIM ASURANSI DENGAN PENCILAN**

adalah benar-benar karya saya sendiri, dan saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan.

Atas pernyataan ini, saya siap menanggung segala risiko dan sanksi yang dijatuhkan kepada saya, apabila di kemudian hari ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya, atau jika ada tuntutan formal atau non-formal dari pihak lain berkaitan dengan keaslian karya saya ini.

Dinyatakan di Bandung,  
3 Agustus 2023



Biqytofa  
NPM: 6161901091

## ABSTRAK

Dengan tujuan menjaga kesejahteraan perusahaan asuransi dalam menjalani bisnisnya, perusahaan asuransi diharuskan untuk menganalisis cadangan klaim yang perlu disiapkan di masa mendatang, dengan cara mengestimasi nilai cadangan klaim. Pada makalah ini akan dibahas estimasi cadangan klaim menggunakan metode *Chain Ladder* sebagai metode non-parametrik dan teknik analisis data *Generalized Linear Model* sebagai metode parametrik. Selanjutnya, ditinjau keefektifan dalam pengerjaan metode pencadangan menggunakan kedua pendekatan tersebut dengan mempertimbangkan proses pengerjaan dan galat yang dihasilkan. Data yang digunakan adalah data asuransi *incurred but not reported* (IBNR), dengan rangka data segitiga *run-off* dalam menjalankan metode *Chain Ladder* dan *Generalized Linear Model* (GLM). Karena variansi data asuransi segitiga *run-off* jauh lebih besar dari rata-ratanya, maka data tersebut mempunyai sifat *overdispersed*, sehingga perlu digunakan pendekatan model *Over-Dispersed Poisson* (ODP) dalam melakukan metode *Generalized Linear Model*. Kedua metode yang telah dianalisis mempunyai kekurangan, di mana kedua metode tersebut terpengaruh secara signifikan dalam data yang terkontaminasi data pencilan. Oleh karena itu, akan digunakan perluasan dari model GLM, yaitu *Generalized Additive Model* (GAM). Dengan tujuan, fitur fungsi pemulusan (*smoothing function*) dalam GAM dapat mengatasi data yang terkontaminasi pencilan dan menghasilkan hasil prediksi nilai cadangan klaim yang lebih akurat untuk data yang terkontaminasi.

**Kata-kata kunci:** *Generalized Linear Model; Generalized Additive Model; Chain Ladder; IBNR; Cadangan Klaim; Over-Dispersed Poisson; Smoothing Function.*

## ABSTRACT

In order to maintain the welfare of insurance companies in conducting their business, insurance companies are required to analyze the claim reserves that need to be prepared in the future by estimating the value of claim reserves. This paper will discuss the estimation of claim reserves using the Chain Ladder method as a non-parametric method and the Generalized Linear Model data analysis technique as a parametric method. Furthermore, the effectiveness of both reserve estimation methods will be evaluated by considering the implementation process and the errors produced. The data used in this study is incurred but not reported (IBNR) data, structured in a run-off triangle format for both the Chain Ladder and Generalized Linear Model (GLM) methods. Due to the significantly higher variance of insurance run-off triangle data compared to its mean, the data exhibits overdispersion. Thus, the Over-Dispersed Poisson (ODP) model approach is employed when applying the Generalized Linear Model. Both analyzed methods have limitations as they are significantly affected by outliers in the data. Therefore, an extension of the GLM method called the Generalized Additive Model (GAM) is used. The smoothing function feature in GAM aims to handle data contaminated with outliers and produce more accurate predictions of claim reserve values for contaminated data.

**Keywords:** *Generalized Linear Model; Generalized Additive Model; Chain Ladder; IBNR; Claim Reserves; Over-Dispersed Poisson; Smoothing Function.*

## KATA PENGANTAR

Penulis ingin mengucapkan terima kasih banyak kepada Tuhan yang maha esa atas berkat dan rahmat yang berkelanjutan, sehingga penulis dapat menyelesaikan skripsi ini. Tentunya, tanpa dukungan dan bantuan dari berbagai pihak, penulisan skripsi ini tidak dapat diselesaikan dengan kualitas yang baik. Oleh karena itu, penulis ingin memberikan ucapan terima kasih kepada:

1. Ayah, Bunda, Kaka tercinta yang selalu menjadi pemicu semangat dalam diri dan membantu dalam menyelesaikan kuliah ini.
2. Bapak Rizky Reza Fauzi, D.Phil.Math. dan Ibu Maria Anestasia, M.Si., M.Act.Sc. selaku pembimbing terbaik bagi penulis, pembimbing yang sangat peduli akan skripsi ini, dan bersedia membimbing dengan banyak keterbatasan penulis. Terlepas dari skripsi ini, pembimbing memberikan dampak positif yang besar dalam kehidupan penulis, sehingga penulis dapat menyelesaikan kuliah ini dengan membawa bekal yang baik untuk jenjang selanjutnya.
3. Bapak Dr. Ferry Jaya Permana dan Felivia Kusnadi, M.Act.Sc., ASAI. selaku penguji yang memastikan bahwa penulis mengerti seluruh isi skripsi ini, dan memberi masukan untuk menyempurnakan skripsi ini.
4. Seluruh dosen Universitas Katolik Parahyangan terutama dosen pada program studi matematika atas ilmu dan didikan yang telah diberikan dengan kualitas yang sangat baik.
5. Penghuni kosan The Maple Place Residence yang sangat memberikan support penuh dengan cara memberikan distraksi agar terus menunda pengerjaan skripsi ini demi menjunjung tinggi atas kebersamaan, keharmonisan, dan solidaritas.

*Last but not least, i wanna thank me. I wanna thank me for believing in me, i wanna thank me for doing all this hard work, i wanna thank me for having no days off, i wanna thank me for never quitting, i wanna thank me for always being a giver and trying to give more than i receive, i wanna thank me for just being me at all times.*

Bandung, 3 Agustus 2023

Penulis

# DAFTAR ISI

<b>KATA PENGANTAR</b>	<b>viii</b>
<b>DAFTAR ISI</b>	<b>ix</b>
<b>DAFTAR GAMBAR</b>	<b>x</b>
<b>DAFTAR TABEL</b>	<b>xi</b>
<b>1 PENDAHULUAN</b>	<b>1</b>
1.1 Latar Belakang . . . . .	1
1.2 Rumusan Masalah . . . . .	2
1.3 Tujuan . . . . .	3
1.4 <i>State of the Art</i> . . . . .	3
<b>2 LANDASAN TEORI</b>	<b>4</b>
2.1 Cadangan Klaim . . . . .	4
2.2 <i>Chain Ladder</i> . . . . .	6
2.3 Regresi Linear . . . . .	10
2.4 <i>Generalized Linear Model</i> . . . . .	11
2.5 Model <i>Over Dispersed Poisson</i> . . . . .	13
2.6 Penaksiran <i>Likelihood</i> Maksimum . . . . .	13
<b>3 PENGAPLIKASIAN METODE PENCADANGAN KLAIM</b>	<b>15</b>
3.1 Karakteristik Data . . . . .	15
3.2 Estimasi Cadangan Klaim Menggunakan <i>Generalized Linear Model</i> (GLM) dengan Pendekatan ODP . . . . .	16
3.3 <i>Generalized Additive Model</i> (GAM) . . . . .	19
3.4 Pemulusan Menggunakan <i>Generalized Additive Model</i> . . . . .	19
3.5 Penaksiran Penalti <i>Likelihood</i> Maksimum dengan <i>Over-Dispersed Poisson</i> . . . . .	20
3.6 Prosedur <i>Bootstrap</i> Klasik . . . . .	21
<b>4 HASIL DAN ANALISIS PERBANDINGAN MODEL</b>	<b>23</b>
4.1 Estimasi Cadangan Klaim Menggunakan Metode <i>Chain Ladder</i> . . . . .	23
4.2 Data Cadangan Klaim dengan Pencilan . . . . .	24
4.3 Hasil Estimasi Cadangan Klaim Menggunakan GLM dan GAM . . . . .	25
4.4 Hasil Estimasi Nilai Cadangan Klaim dengan Letak Data Pencilan pada $C_{2011,10}$ . . . . .	28
4.5 Hasil Estimasi Nilai Cadangan Klaim dengan Letak Data Pencilan Pada $C_{2021,0}$ . . . . .	30
<b>5 KESIMPULAN DAN SARAN</b>	<b>32</b>
5.1 Kesimpulan . . . . .	32
5.2 Saran . . . . .	33
<b>DAFTAR REFERENSI</b>	<b>34</b>



**DAFTAR GAMBAR**

2.1 Fungsi penghubung ..... 12



## DAFTAR TABEL

2.1	Segitiga <i>run-off</i> inkremental . . . . .	6
2.2	Contoh segitiga <i>run-off</i> inkremental . . . . .	7
2.3	Contoh segitiga <i>run-off</i> kumulatif . . . . .	7
2.4	Faktor penundaan . . . . .	7
2.5	Metode <i>average</i> . . . . .	8
2.6	Metode <i>mean</i> . . . . .	8
2.7	Contoh segitiga <i>run-off</i> kumulatif penuh . . . . .	9
2.8	Contoh segitiga <i>run-off</i> penuh . . . . .	9
2.9	Perbandingan LM dan GLM . . . . .	11
2.10	Fungsi penghubung bawaan . . . . .	12
3.1	Segitiga <i>run-off</i> inkremental . . . . .	15
3.2	Segitiga <i>run-off</i> kumulatif . . . . .	16
3.3	<i>Run-off</i> data inkremental (GLM) . . . . .	18
4.1	Faktor penundaan menggunakan data asuransi kendaraan bermotor . . . . .	23
4.2	Estimasi cadangan klaim menggunakan <i>chain ladder</i> . . . . .	24
4.3	Segitiga <i>run-off</i> inkremental dengan data pencilan . . . . .	24
4.4	Estimasi parameter dengan model GLM . . . . .	25
4.5	Segitiga penuh model GLM dengan data asli . . . . .	26
4.6	Segitiga penuh model GLM dengan data pencilan . . . . .	26
4.7	Segitiga penuh model GAM dengan data asli . . . . .	27
4.8	Segitiga penuh model GAM dengan data pencilan . . . . .	27
4.9	Data asli . . . . .	28
4.10	Data terkontaminasi . . . . .	28
4.11	Segitiga penuh model GLM dengan data pencilan $C_{2011,10}$ . . . . .	29
4.12	Segitiga penuh model GAM dengan data pencilan $C_{2011,10}$ . . . . .	29
4.13	Data terkontaminasi dengan letak pencilan $C_{2011,10}$ . . . . .	30
4.14	Segitiga penuh model GLM dengan data pencilan $C_{2021,0}$ . . . . .	30
4.15	Segitiga penuh model GAM dengan data pencilan $C_{2021,0}$ . . . . .	31
4.16	Data terkontaminasi dengan letak pencilan $C_{2021,0}$ . . . . .	31

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Di masa modern ini, peran kendaraan bermotor cukup vital sebagai alat transportasi darat yang dibutuhkan untuk aktivitas manusia. Berdasarkan Badan Pusat Statistik,<sup>1</sup> jumlah kendaraan bermotor di Indonesia ada sebanyak 141.992.573 pada tahun 2021 dan 136.137.451 pada tahun 2020, yang mengartikan bahwa jumlah kendaraan bermotor di Indonesia mengalami peningkatan dengan jumlah yang tidak sedikit. Selain itu, jumlah kecelakaan lalu lintas darat di Indonesia mencapai 103.645 kasus pada tahun 2021 menunjukkan kenaikan sebesar 3,62% ketimbang tahun sebelumnya yang sebanyak 100.028 kasus. Hal ini penting untuk dijadikan perhatian penduduk Indonesia untuk mempertimbangkan penggunaan asuransi.

Asuransi merupakan sebuah alat perlindungan bagi pihak tertanggung apabila mengalami risiko di masa yang akan datang, dengan pihak tertanggung akan membayar premi pada setiap periodenya sebagai bentuk imbalan jasa untuk mendapatkan klaim atau ganti rugi dari pihak penanggung. Besar klaim dan besar premi yang dibayarkan ditentukan setelah calon tertanggung melewati proses identifikasi dan seleksi risiko atau *underwriting*. Setiap tertanggung tentunya memiliki hasil *underwriting* atau risiko yang berbeda-beda, yang menyebabkan perusahaan asuransi mendapatkan pengajuan klaim dengan besaran yang variatif, dari klaim yang besarnya kecil hingga klaim yang sangat besar. Maka dari itu, perusahaan asuransi harus memiliki cadangan dana untuk dapat menyelesaikan klaim kepada pihak tertanggung.

Untuk menjaga kesejahteraan perusahaan asuransi dalam menjalankan bisnisnya, perusahaan asuransi diharuskan untuk mengoptimalkan cadangan dana dengan cara memprediksi besar klaim yang dapat terjadi di masa mendatang. Terdapat dua jenis metode yang dapat digunakan, yaitu metode parametrik yang memanfaatkan sifat-sifat distribusi variabel acak seperti yang dibahas oleh Mahardika [1], dan metode nonparametrik yang tidak melibatkan sifat-sifat distribusi variabel acak seperti yang dibahas oleh Jonathan [2]. Kedua jenis metode ini mempunyai kelebihan dan kekurangan masing masing, seperti metode parametrik yang mempunyai tingkat kesulitan yang relatif tinggi dan menggunakan waktu yang cukup lama untuk diterapkan, sementara metode nonparametrik pada umumnya relatif lebih sederhana untuk diterapkan.

Metode pencadangan nonparametrik yang umumnya digunakan perusahaan asuransi untuk memprediksi klaim dengan data yang terjadi tetapi belum dilaporkan (*Incurred But Not Reported*) (IBNR) dan dilaporkan tetapi belum terselesaikan (*Reported But Not Settled*) (RBNS), salah satunya adalah metode pencadangan *chain ladder*. Metode pencadangan *chain ladder* (CL) terbilang cukup

---

<sup>1</sup><https://www.bps.go.id/indicator/17/57/1/jumlah-kendaraan-bermotor.html>

sederhana karena tidak melibatkan sifat-sifat distribusi variabel acak, dan karena kesederhanaannya, analisis yang bisa dilakukan kurang mendalam dan tidak menyediakan fitur galat secara transparan. Selain itu, metode CL juga sensitif terhadap data ekstrem atau mempunyai perbedaan yang jauh terhadap karakteristik data observasi lainnya, atau dapat disebut juga sebagai data *outlier*. Oleh karena itu, metode ini dapat dilanjutkan dengan menggunakan alternatif lainnya yaitu model parametrik GLM.

Model parametrik GLM ini melibatkan sifat-sifat distribusi dan menghabiskan waktu yang cukup banyak. Namun, model ini dapat menyediakan galat yang berguna bagi aktuaris untuk menentukan apakah cadangan tambahan harus diadakan untuk meminimalisasi risiko bagi perusahaan. Dalam rangka pengerjaan CL, data yang digunakan adalah segitiga *run-off*, yang biasanya mempunyai variansi jauh lebih besar dibandingkan rata-ratanya. Oleh karena itu, digunakan pendekatan model *over-dispersed poisson* (ODP). Model ini mempunyai kekurangan yang serupa dengan metode CL, yaitu belum dapat menangani kasus yang melibatkan data pencilan.

Dalam realitanya, data pencilan atau *outlier* mungkin untuk terjadi. Sebagai contoh pada asuransi kendaraan bermotor, klaim yang luar biasa besar dapat melonjak setelah hari raya lebaran atau imlek. Dalam praktek aktuaria, fenomena ini dapat mengakibatkan efek negatif terhadap akurasi estimasi dan inferensi cadangan klaim secara keseluruhan. Oleh karena itu, penting bagi aktuaris untuk mempersiapkan model yang dapat menanggulangi kasus tersebut. Menurut Le Chang dan Guangyuan Gao [3], beberapa solusi untuk kasus tersebut, misalnya dengan melakukan *exponential smoothing*, menggunakan kurva Hoerl, dan menggunakan *Generalized Additive Model* (GAM).

Model GAM merupakan model yang digunakan untuk menangani model linear berganda yang mempunyai variabel non-linear di dalam persamaan awal. Untuk menangani kasus tersebut, model ini menggunakan prosedur fungsi pemulusan (*smooth function*) yang berguna untuk menangani data yang terdeteksi data pencilan seperti yang telah dibahas oleh Le Chang dan Guangyuan Gao [3].

Pada skripsi ini dianalisis performansi metode CL, GLM, dan GAM dalam mengestimasi besar nilai cadangan klaim dengan data yang normal dan data yang terdeteksi adanya pencilan. Pengaruh dari letak pencilan pada segitiga *run-off* pun akan dianalisis dan dibandingkan berdasarkan nilai galatnya.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang ada, pada makalah ini akan difokuskan untuk metode pencadangan klaim dari metode CL, GLM, dan GAM dengan data yang terkontaminasi pencilan. Jadi, rumusan masalah yang terbentuk adalah:

1. Bagaimana cara mengestimasi nilai cadangan klaim dengan metode pencadangan *generalized linear model* dan *generalized additive model* menggunakan pendekatan *over-dispersed poisson* (ODP)?
2. Bagaimana memodifikasi *generalized linear model* dan *generalized additive model* dengan kerangka pengerjaan *chain ladder* menggunakan data segitiga *run-off*?

3. Bagaimana analisis perbandingan antara *generalized linear model* dengan *generalized additive model* dalam mengestimasi nilai cadangan klaim yang mengandung data pencilan?
4. Bagaimana analisis perbandingan antara *generalized linear model* dengan *generalized additive model* dalam mengestimasi nilai cadangan klaim berdasarkan letak pencilannya?

### 1.3 Tujuan

Berdasarkan rumusan masalah yang telah dituliskan sebelumnya, maka tujuan dari makalah ini adalah:

1. Mengetahui bagaimana prosedur dan langkah-langkah pengerjaan dalam mengestimasi nilai cadangan klaim menggunakan *generalized linear model* dan *generalized additive model* dengan pendekatan *over-dispersed poisson* (ODP).
2. Mengetahui bagaimana memodifikasi teknik *generalized linear model* dalam mengestimasi besar nilai cadangan klaim menggunakan rangka pengerjaan metode *chain ladder* untuk data segitiga *run-off*.
3. Mengetahui analisis perbandingan *generalized linear model* dan *generalized additive model* dalam mengestimasi besar nilai cadangan klaim ketika adanya data pencilan.
4. Mengetahui analisis perbandingan *generalized linear model* dan *generalized additive model* dalam mengestimasi besar nilai cadangan klaim berdasarkan letak pencilannya.

### 1.4 State of the Art

Analisis terhadap metode-metode pencadangan klaim pada skripsi ini menggunakan data segitiga *run-off* dengan rangka pengerjaan *chain ladder*. Basis pengerjaan skripsi ini didasari dengan pengerjaan rangka *chain ladder* (non-parametrik) yang diambil dari skripsi Jonathan [2]. Untuk menerapkan metode CL terhadap model parametrik seperti GLM dan GAM, skripsi ini menggunakan artikel Kremer [4] dalam mengembangkan metode *chain ladder* secara stokastik. Untuk pengerjaan dan analisis model parametrik GLM dan GAM, skripsi ini menerapkan artikel Le Chang dan Guanyuan Gao [3] dalam menganalisis data asli maupun data pencilan. Setelah itu, pengembangan dilanjutkan dengan menganalisis performa model dengan keberadaan letak pencilan yang berbeda menggunakan galat yang telah diproduksi masing-masing.