

SKRIPSI

**PEMODELAN BESAR KLAIM ASURANSI KESEHATAN
DENGAN DISTRIBUSI SKEW-NORMAL,
ALPHA-SKEW-NORMAL, DAN ALPHA-SKEW-LAPLACE**



Steven Sergio

NPM: 6161801001

**PROGRAM STUDI MATEMATIKA
FAKULTAS TEKNOLOGI INFORMASI DAN SAINS
UNIVERSITAS KATOLIK PARAHYANGAN
2022**

FINAL PROJECT

**MODELLING SEVERITY OF HEALTH INSURANCE CLAIMS
USING SKEW-NORMAL, ALPHA-SKEW-NORMAL, AND
ALPHA-SKEW-LAPLACE DISTRIBUTION**



Steven Sergio

NPM: 6161801001

**DEPARTMENT OF MATHEMATICS
FACULTY OF INFORMATION TECHNOLOGY AND SCIENCES
PARAHYANGAN CATHOLIC UNIVERSITY
2022**

LEMBAR PENGESAHAN

PEMODELAN BESAR KLAIM ASURANSI KESEHATAN DENGAN DISTRIBUSI SKEW-NORMAL, ALPHA-SKEW-NORMAL, DAN ALPHA-SKEW-LAPLACE

Steven Sergio

NPM: 6161801001

Bandung, 28 Januari 2022

Menyetujui,

Pembimbing 1



Iwan Sugiarto, M.Si.

Pembimbing 2



Rizky Reza Fauzi, D.Phil.Math.

Ketua Tim Penguji



Farah Kristiani, Ph.D.

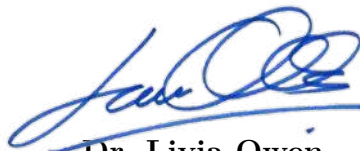
Anggota Tim Penguji



Benny Yong, Ph.D.

Mengetahui,

Ketua Program Studi



Dr. Livia Owen

PERNYATAAN

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

PEMODELAN BESAR KLAIM ASURANSI KESEHATAN DENGAN DISTRIBUSI SKEW-NORMAL, ALPHA-SKEW-NORMAL, DAN ALPHA-SKEW-LAPLACE

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,
Tanggal 28 Januari 2022



Steven Sergio
NPM: 6161801001

ABSTRAK

Kegiatan pengajuan klaim dalam asuransi kesehatan bergantung pada besar klaim asuransi untuk mengganti kerugian yang timbul atas menurunnya kondisi kesehatan tertanggung. Kegiatan tersebut tentu saja membentuk suatu data besar klaim asuransi yang perlu dimodelkan dengan suatu distribusi tertentu. Hal ini diperlukan agar perusahaan asuransi mampu memprediksi klaim yang terjadi di masa depan sekaligus mempersiapkan cadangan asuransi yang cukup. Namun, kadang kala terjadi suatu peristiwa luar biasa yang mempengaruhi kemiringan, modalitas, dan tebal ekor data besar klaim asuransi. Oleh karena itu, dalam skripsi ini digunakan distribusi Skew-normal, Alpha-skew-normal, dan Alpha-skew-Laplace, yang termasuk dalam kelompok distribusi *flexible skewed* guna memodelkan suatu data besar klaim asuransi yang telah terdampak peristiwa luar biasa tersebut. Untuk mencapai tujuan dari skripsi, yaitu mengetahui distribusi yang terbaik dalam memodelkan data besar klaim asuransi, langkah-langkah yang dilakukan adalah menggali informasi terkait distribusi Skew-normal, Alpha-skew-normal, dan Alpha-skew-Laplace, mengestimasi nilai parameter penaksir dari distribusi menggunakan metode estimasi *maximum likelihood*, serta menentukan distribusi terbaik untuk memodelkan data besar klaim asuransi. Penentuan distribusi terbaik untuk memodelkan data besar klaim asuransi dilakukan dengan terlebih dahulu melaksanakan uji Kolmogorov-Smirnov dan uji Anderson-Darling agar diketahui kecocokan dari kandidat distribusi dalam memodelkan data besar klaim asuransi. Selanjutnya, nilai kriteria informasi Akaike dan Bayesian digunakan untuk menentukan distribusi yang terbaik dalam memodelkan data besar klaim asuransi dari kandidat distribusi yang dipandang cocok untuk memodelkan data besar klaim asuransi. Berdasarkan penelitian yang telah dilaksanakan, distribusi Alpha-skew-Laplace adalah yang terbaik dalam memodelkan data besar klaim asuransi kesehatan, diikuti dengan distribusi Skew-normal dan distribusi Alpha-skew-normal.

Kata-kata kunci: besar klaim asuransi, modalitas, distribusi Skew-normal, distribusi Alpha-skew-normal, distribusi Alpha-skew-Laplace, metode estimasi *maximum likelihood*, uji Kolmogorov-Smirnov, uji Anderson-Darling, nilai kriteria informasi Akaike, nilai kriteria informasi Bayesian.

ABSTRACT

Claim experience in health insurance depends on amount of claim paid by the insurance company to cover the loss of the insured due to peril, in this case, it is deterioration of insured's health condition. Sequence of claim experience will then form an insurance claim amount data and insurance company will establish such mathematical model based on insurance claim amount data. Mathematical modelling toward amount of insurance claim data is necessary to be performed since insurance company needs to predict claims that possibly happened in the future and also prepare sufficient amount of insurance reserving. However, in some point of time, there are several unpredictable events, such as pandemic or catastrophe affecting skewness, modality, and tail of the data. To perform a mathematical modelling towards affected insurance claim amount data, Skew-normal, Alpha-skew-normal, and Alpha-skew-Laplace Distribution will be used. First step to be done is discover information needed from each distribution involved in this final project, then estimate the value of distribution parameters with maximum likelihood estimation method. Next step is to apply Kolmogorov-Smirnov and Anderson-Darling test to find suitable distributions to fit insurance claim amount data. There is some probability of appearance of more than one distribution among suitable distributions, hence the selection of best distribution to fit insurance claim amount data will be conducted by calculation of Akaike and Bayesian Information Criteria value. According to research conducted, it is known that Alpha-skew-Laplace distribution is the best distribution to model the insurance claim amount data and it is followed by Skew-normal distribution and Alpha-skew-normal distribution, respectively.

Keywords: claim severity, modality, Skew-normal distribution, Alpha-skew-normal distribution, Alpha-skew-Laplace distribution, maximum likelihood estimation, Kolmogorov-Smirnov test, Anderson-Darling test, Akaike Information Criteria, Bayesian Information Criteria.

Untuk yang tercinta, Papi dan Mami.

KATA PENGANTAR

Puji syukur ke hadirat Tuhan Yesus Kristus karena berkat dan anugerah-Nya, penulis berhasil menyelesaikan skripsi ini. Skripsi berjudul “Pemodelan Besar Klaim Asuransi Kesehatan dengan Distribusi Skew-Normal, Alpha-skew-normal, dan Alpha-skew-Laplace” ini disusun sebagai salah satu syarat untuk menyelesaikan studi sarjana di Program Studi Matematika, Fakultas Teknologi Informasi dan Sains, Universitas Katolik Parahyangan, Bandung.

Keberhasilan penyusunan skripsi ini tidak terlepas dari dukungan dan dorongan dari berbagai pihak. Untuk itu, penulis ingin mengapresiasi dan mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Kedua orang tua penulis yang senantiasa mendoakan, menasihati, dan memberikan dukungan penuh kepada penulis selama penulis menempuh studi.
2. Bapak Iwan Sugiarto, M.Si. dan Bapak Rizky Reza Fauzi, D.Phil.Math. selaku dosen pembimbing yang telah mencurahkan ilmu, waktu, dan tenaga sehingga penulis dapat merampungkan skripsi ini.
3. Ibu Farah Kristiani, Ph.D. dan Bapak Benny Yong, Ph.D. selaku dosen penguji yang telah memeriksa skripsi penulis sekaligus memberikan kritik dan saran terhadap skripsi yang penulis susun sehingga skripsi ini menjadi lebih baik.
4. Bapak Liem Chin, M.Si. selaku dosen wali penulis selama menempuh studi sekaligus koordinator skripsi yang telah memberikan informasi, mengarahkan, dan memberikan saran kepada penulis, baik untuk keperluan perkuliahan maupun untuk keperluan penyusunan skripsi.
5. Seluruh dosen Fakultas Teknologi Informasi dan Sains (FTIS) Universitas Katolik Parahyangan yang telah mengajar dan memberikan ilmu pengetahuan yang sangat berharga bagi penulis.
6. Seluruh staf Tata Usaha FTIS yang telah menolong penulis dalam mengurus berbagai keperluan administratif selama penulis menempuh studi.
7. Adriel Luis Figo yang telah memberikan dukungan moril kepada penulis selama penulis menempuh studi sekaligus menjadi teman bertukar pikiran mengenai sejarah, sepakbola (khususnya Manchester United), investasi saham, dan *cryptocurrency*.
8. Kimberley Blessinda yang telah menyemangati penulis dalam menyelesaikan skripsi sekaligus menolong penulis dalam melakukan *proofreading* berbagai dokumen dan menjadi teman bercerita penulis selama ini.
9. Adrian dan Andrew yang banyak bertukar pikiran dan informasi dengan penulis seputar penyusunan draf skripsi dan teknis pelaksanaan sidang skripsi.
10. Teman-teman LED: David, Gaby, Shelly, Zaza, Agit, dan Jane untuk memori, kebersamaan, dan dukungan moril yang diberikan kepada penulis selama penulis menempuh studi.
11. Teman-teman ulala: Agung, Aditya, Mikha yang telah menjadi teman jalan-jalan penulis sekaligus teman bercerita selama penulis menempuh studi.
12. Teman-teman HappyPadoru: Malvin, Topas, Rhandy, Jonathan Dinhara (Jojo), Jonathan Prasetyo Johan (JPJ), Adrian, Novaldi, Yohanes Reinhart, Jason untuk kebersamaan, dukungan, dan *sharing* ilmu pengetahuan selama penulis menempuh studi.
13. Teman-teman meatology: Malvin, Topas, Jojo, JPJ, Adrian untuk memori dan kebersamaan dalam berwisata kuliner selama kurang lebih 3,5 tahun ini.
14. Teman-teman pengurus Komisi Pemuda dan Komisi Multimedia GKI Taman Cibunut Bandung untuk pengalaman berorganisasi selama ini dan dukungan moril yang diberikan selama penulis

menempuh studi.

15. Teman-teman Matematika angkatan 2018 yang tidak dapat penulis sebutkan satu per satu untuk kebersamaan selama ini.
16. Pihak-pihak lainnya yang tak dapat penulis sebutkan satu per satu, yang turut berperan serta atas keberhasilan penulis dalam menyelesaikan skripsi.

Penulis menyadari bahwa skripsi ini belumlah sempurna. Oleh karena itu, penulis dengan segala kerendahan hati memohon maaf atas segala kekurangan yang terdapat dalam skripsi ini dan penulis dengan senang hati menerima kritik dan saran yang membangun dari pembaca sekalian. Akhir kata, mudah-mudahan skripsi yang penulis susun ini memberikan manfaat dan inspirasi kepada pihak-pihak yang membutuhkan. Terima kasih.

Bandung, Januari 2022

Penulis

DAFTAR ISI

KATA PENGANTAR	xv
DAFTAR ISI	xvii
DAFTAR GAMBAR	xix
DAFTAR TABEL	xxi
1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	2
1.4 Batasan Masalah	2
1.5 Sistematika Pembahasan	2
2 LANDASAN TEORI	5
2.1 Distribusi Peluang	5
2.1.1 Peubah Acak Kontinu	5
2.1.2 Jenis-jenis Distribusi Peluang	5
2.2 Metode Penaksiran Parameter dengan <i>Maximum Likelihood</i>	8
2.3 Uji Kecocokan Distribusi	8
2.3.1 Uji Kolmogorov-Smirnov	9
2.3.2 Uji Anderson-Darling	9
2.4 Nilai Kriteria Informasi	10
3 KARAKTERISTIK DISTRIBUSI SKEW-NORMAL, ALPHA-SKEW-NORMAL, DAN ALPHA-SKEW-LAPLACE	11
3.1 Distribusi <i>Unimodal Skewed</i>	11
3.2 Distribusi <i>Bimodal Skewed</i>	17
3.2.1 Distribusi Alpha-skew-normal (ASN)	17
3.2.2 Distribusi Alpha-skew-Laplace (ASLP)	24
4 SIMULASI PEMODELAN DATA BESAR KLAIM ASURANSI	33
4.1 Penentuan Nilai Penaksir Parameter Distribusi	34
4.1.1 Nilai Penaksir Parameter Distribusi Skew-normal (SN)	34
4.1.2 Nilai Penaksir Parameter Distribusi Alpha-skew-normal (ASN)	35
4.1.3 Nilai Penaksir Parameter Distribusi Alpha-skew-Laplace (ASLP)	36
4.2 Penyeleksian Kandidat Distribusi dengan Uji Kolmogorov-Smirnov	37
4.2.1 Uji Kolmogorov-Smirnov untuk Distribusi Skew-normal (SN)	37
4.2.2 Uji Kolmogorov-Smirnov untuk Distribusi Alpha-skew-normal (ASN)	39
4.2.3 Uji Kolmogorov-Smirnov untuk Distribusi Alpha-skew-Laplace (ASLP)	39
4.3 Penyeleksian Kandidat Distribusi dengan Uji Anderson-Darling	40
4.3.1 Uji Anderson-Darling untuk Distribusi Skew-normal (SN)	41

4.3.2	Uji Anderson-Darling untuk Distribusi Alpha-skew-normal (ASN)	42
4.3.3	Uji Anderson-Darling untuk Distribusi Alpha-skew-Laplace (ASLP)	42
4.4	Nilai Kriteria Informasi untuk Distribusi yang Bersaing	43
4.4.1	Nilai <i>AIC</i> dan <i>BIC</i> Distribusi Skew-normal (SN)	43
4.4.2	Nilai <i>AIC</i> dan <i>BIC</i> Distribusi Alpha-skew-Laplace (ASLP)	44
5	KESIMPULAN DAN SARAN	47
5.1	Kesimpulan	47
5.2	Saran	48
	DAFTAR REFERENSI	49

DAFTAR GAMBAR

3.1	Visualisasi Distribusi SN Standar	14
3.2	Visualisasi Distribusi ASN Standar	21
3.3	Visualisasi Distribusi ASLP Standar	29
4.1	Histogram dari Data Besar Klaim Asuransi Kesehatan	33

DAFTAR TABEL

4.1	Ringkasan Data Besar Klaim Asuransi Kesehatan	34
4.2	Ilustrasi Perhitungan $F_n(x)$	37
4.3	Informasi untuk pelaksanaan uji Kolmogorov-Smirnov bagi distribusi SN	38
4.4	Informasi untuk pelaksanaan uji Kolmogorov-Smirnov bagi distribusi ASN	39
4.5	Informasi untuk pelaksanaan uji Kolmogorov-Smirnov bagi distribusi ASLP	40
4.6	Informasi untuk Pelaksanaan Uji Anderson-Darling bagi Distribusi SN	41
4.7	Informasi untuk Pelaksanaan Uji Anderson-Darling bagi Distribusi ASN	42
4.8	Informasi untuk Pelaksanaan Uji Anderson-Darling bagi Distribusi ASLP	43
4.9	Nilai AIC dan BIC dari Distribusi SN dan ASLP	45

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Secara singkat, asuransi adalah perjanjian di antara dua pihak, yaitu pihak tertanggung dan pihak penanggung. Pihak tertanggung berkewajiban memberikan imbalan atas jasa yang disediakan pihak penanggung. Di sisi lain, pihak penanggung berkewajiban memberikan kompensasi atas kerugian dari peristiwa tertentu yang dialami oleh pihak tertanggung [1].

Pada skripsi ini, jenis asuransi yang menjadi topik utama adalah asuransi kesehatan. Asuransi kesehatan dipilih atas pertimbangan populasi manusia yang meningkat dari waktu ke waktu. Ditambah lagi dengan hal yang merugikan kondisi kesehatan dari populasi manusia yang bertambah secara terus-menerus seperti penyakit, wabah, atau pandemi sehingga kesadaran masyarakat akan pentingnya asuransi kesehatan pun meningkat.

Dalam praktek berasuransi, tentu saja rangkaian pengajuan klaim dari pihak tertanggung kepada perusahaan asuransi untuk mengganti kerugian yang dialami tertanggung membentuk suatu data besar klaim asuransi yang perlu dimodelkan dengan suatu distribusi tertentu. Pemodelan data besar klaim asuransi ini penting untuk dilakukan dikarenakan perusahaan asuransi harus menaksir besar total klaim dan besar cadangan asuransi demi kelancaran kegiatan perasuransian di masa yang akan datang.

Secara garis besar, dalam dunia perasuransian banyak dibahas topik yang berkaitan dengan statistika dan peluang, sehingga tentu sudah tidak asing bagi pelaku usaha perasuransian dengan distribusi-distribusi yang sudah umum, seperti misalnya distribusi normal, Poisson, beta, gamma, binomial, dan lain sebagainya [2]. Namun, dalam dunia perasuransian, terkadang terdapat faktor-faktor atau peristiwa luar biasa yang menyebabkan data besar klaim asuransi mengalami kemencengan/*skewness* dan memiliki ekor/*tail* yang tebal.

Faktor atau peristiwa luar biasa yang dimaksud pun juga bisa saja menyebabkan data besar klaim asuransi yang sebelumnya unimodal (terdapat satu modus dalam distribusi) menjadi bimodal (terdapat dua modus pada distribusi). Dalam hal data besar klaim asuransi mengalami kemencengan, perubahan tebal ekor, dan perubahan jumlah modus, distribusi yang sudah umum digunakan tidak lagi mampu untuk memodelkan data besar klaim asuransi tersebut dengan akurat.

Oleh karena itulah, diperlukan suatu alternatif dari distribusi yang telah umum digunakan untuk memodelkan data besar klaim asuransi yang terimbas faktor atau peristiwa yang dimaksud seakurat mungkin. Dalam hal ini, distribusi Skew-normal, Alpha-skew-normal, dan Alpha-skew-Laplace yang termasuk dalam distribusi *flexible skewed* dapat menjadi alternatif bagi distribusi yang telah umum digunakan dikarenakan distribusi *flexible skewed* mampu memodelkan data besar klaim asuransi unimodal dan bimodal [3].

Distribusi *Skew-normal*, *Alpha-skew-normal*, dan *Alpha-skew-Laplace* yang menjadi kandidat distribusi untuk memodelkan data besar klaim asuransi pun nantinya akan dicari nilai penaksir parameternya berdasarkan data yang ada. Hal tersebut diperlukan untuk melakukan uji kecocokan agar diketahui mana saja distribusi yang cocok dengan data besar klaim asuransi. Nilai kriteria informasi Akaike (AIC) dan Bayesian (BIC) juga dicari dari distribusi-distribusi yang dinyatakan cocok untuk mengetahui distribusi yang terbaik dalam memodelkan suatu data besar klaim asuransi.

1.2 Rumusan Masalah

Skripsi ini akan berfokus pada seputar distribusi *Skew-normal*, Alpha-skew-normal, dan Alpha-skew-Laplace, yang mencakup penyelidikan atas properti dari distribusi yang bersangkutan. Skripsi ini juga berfokus pada pemodelan data besar klaim asuransi menggunakan distribusi Skew-normal, Alpha-skew-normal, dan Alpha-skew-Laplace. Adapun hal-hal yang akan digali dalam pengerjaan skripsi ini dapat disusun dalam suatu rumusan masalah sebagai berikut.

1. Bagaimana properti yang ada pada distribusi Skew-normal, Alpha-skew-normal, dan Alpha-skew-Laplace?
2. Bagaimana sifat yang terdapat pada distribusi Skew-normal, Alpha-skew-normal, dan Alpha-skew-Laplace?
3. Bagaimana keabsahan dari distribusi Skew-normal, Alpha-skew-normal, dan Alpha-skew-Laplace?
4. Bagaimana memperoleh nilai penaksir parameter dari setiap kandidat distribusi?
5. Bagaimana menentukan distribusi yang terbaik untuk memodelkan data besar klaim asuransi?

1.3 Tujuan

Tujuan yang akan dicapai oleh skripsi ini adalah sebagai berikut.

1. Menyelidiki dan menjelaskan properti yang ada pada distribusi Skew-normal, Alpha-skew-normal, dan Alpha-skew-Laplace.
2. Menyelidiki dan menjelaskan sifat yang terdapat pada distribusi Skew-normal, Alpha-skew-normal, dan Alpha-skew-Laplace.
3. Menyelidiki dan menjelaskan keabsahan dari distribusi Skew-normal, Alpha-skew-normal, dan Alpha-skew-Laplace.
4. Memperoleh nilai penaksir parameter dari setiap kandidat distribusi.
5. Menentukan distribusi yang terbaik untuk memodelkan data besar klaim asuransi.

1.4 Batasan Masalah

Pada skripsi ini terdapat beberapa batasan masalah yang perlu untuk diperhatikan, yaitu distribusi Skew-normal, Alpha-skew-normal, dan Alpha-skew-Laplace yang akan dijelaskan merupakan bagian dari distribusi *flexible skewed* yang hanya mendukung data besar klaim asuransi unimodal dan bimodal. Selain itu, pemodelan data besar klaim asuransi hanya menggunakan sumber data dari [4] saja.

Tidak hanya itu, penentuan dan pengujian nilai penaksir parameter dari kandidat distribusi dikerjakan secara numerik dengan menggunakan metode Newton-Raphson pada peranti lunak R dengan tingkat signifikansi yang digunakan pada skripsi ini adalah sebesar 5%.

1.5 Sistematika Pembahasan

Sistematika pembahasan pada skripsi ini terbagi atas lima bab sebagai berikut.

Bab 1: Pendahuluan

Bab ini berisi latar belakang, rumusan masalah, tujuan, batasan masalah, dan sistematika pembahasan.

Bab 2: Landasan Teori

Bab ini berisi dasar teori yang mendukung topik dalam skripsi ini. Dasar teori ini nantinya akan dimanfaatkan dalam pemodelan data besar klaim asuransi.

Bab 3: Karakteristik dan Penyelidikan Distribusi Skew-normal, Alpha-skew-normal, dan Alpha-skew-Laplace

Bab ini berisi pembahasan akan properti dari distribusi Skew-normal, Alpha-skew-normal, dan Alpha-skew-Laplace yang nantinya akan digunakan untuk memodelkan data besar klaim asuransi. Bab ini juga memuat penyelidikan akan karakteristik dan properti dari distribusi Skew-normal, Alpha-skew-normal, dan Alpha-skew-Laplace.

Bab 4: Simulasi Pemodelan Data Besar Klaim Asuransi

Bab ini berisi mekanisme pencarian nilai penaksir parameter dari distribusi Skew-normal, Alpha-skew-normal, dan Alpha-skew-Laplace serta penyeleksian distribusi yang paling baik dalam memodelkan data besar klaim asuransi dari kandidat distribusi yang ada.

Bab 5: Kesimpulan dan Saran

Bab ini berisi kesimpulan yang dapat ditarik berdasarkan hasil simulasi pemodelan data yang telah dilakukan pada bab sebelumnya dan saran untuk pengembangan lebih lanjut dari topik skripsi ini.