

PERBANDINGAN ESTIMASI PREMI KLAIM ASURANSI KESEHATAN BERDASARKAN WILAYAH DAN STATUS PEROKOK MENGGUNAKAN MODEL BÜHLMANN–STRAUB

*Berlian Setiawaty, Annisa Aulia Putri, Bagus Handoko, Dendy Aditya, Monica Dewi Putri Kusmana, dan Ruhiyat

Program Studi Aktuaria, Sekolah Sains Data, Matematika, dan Informatika
Institut Pertanian Bogor, Jl. Meranti, Kampus IPB Dramaga Bogor.

berlianse@apps.ipb.ac.id *corresponding author, aauliaputri@apps.ipb.ac.id,
bagushandoko@apps.ipb.ac.id, dendy_aditya@apps.ipb.ac.id, monicaakusmana@apps.ipb.ac.id,
ruhiyat-mat@apps.ipb.ac.id

Abstrak

Peningkatan klaim asuransi kesehatan menuntut perusahaan asuransi menetapkan premi yang lebih sesuai dengan tingkat risiko. Penelitian ini bertujuan untuk mengestimasi dan membandingkan premi kredibilitas murni pada portofolio asuransi kesehatan menggunakan model semiparametrik Bühlmann–Straub, dengan fokus pada kombinasi faktor risiko status perokok dan wilayah tempat tinggal. Data yang digunakan adalah data besar klaim asuransi kesehatan dari Kaggle tahun 2023 yang dikelompokkan menjadi empat kelompok berdasarkan status perokok, bukan perokok dan *region* yaitu *Southeast* dan *Northeast*. Data besar klaim keempat kelompok tersebut dimodelkan menggunakan sebaran lognormal dan Weibull. Hasil penelitian menunjukkan bahwa model kredibilitas Bühlmann–Straub efektif menghasilkan premi asuransi kesehatan yang lebih akurat pada portofolio heterogen, dengan status perokok terbukti menjadi faktor risiko paling dominan dibandingkan wilayah. Faktor kredibilitas (Z) yang dihasilkan sangat tinggi, berkisar pada 0,99, menandakan bahwa pengalaman klaim tiap kelompok dapat diandalkan dalam penentuan premi. Hasil estimasi premi akhir tertinggi diperoleh kelompok perokok di *Southeast* (34841,838), sedangkan yang terendah adalah bukan perokok di *Southeast* (8296,787).

Kata kunci: Asuransi Kesehatan, Bühlmann–Straub, Estimasi Premi, Lognormal, Weibull

1 Pendahuluan

Kondisi perekonomian yang cenderung fluktuatif, ditandai dengan peningkatan biaya hidup dan inflasi memberikan tekanan berkelanjutan pada berbagai sektor industri, termasuk perusahaan asuransi. Dalam menghadapi situasi tersebut, asuransi berperan sebagai upaya untuk mengalokasikan atau mengalihkan risiko individu kepada perusahaan asuransi. Oleh karena itu, perusahaan asuransi dituntut untuk melakukan peninjauan harga atau *repricing* secara berkala demi menjaga premi tetap relevan dengan risiko yang ditanggung. Asuransi terdiri atas beberapa macam jenis, seperti asuransi kesehatan, asuransi jiwa, dan asuransi kendaraan. Di negara berkembang

2020 Mathematics Subject Classification: 91B05, 91G05.

Diajukan: 15/05/2026, diterima: 23/06/2026. DOI: <https://doi.org/10.29244/milang.22.1.15-26>

MILANG Journal of Mathematics and Its Applications, Vol.22, No.1, pp.15-26

ISSN: 2963-5233

seperti Indonesia, kesadaran masyarakat terhadap pentingnya asuransi semakin meningkat, khususnya dalam penggunaan asuransi kesehatan.

Peningkatan kesadaran ini diiringi dengan tantangan serius dari sisi manajemen risiko perusahaan. Di Indonesia, Asosiasi Asuransi Jiwa Indonesia (AAJI) melaporkan bahwa klaim asuransi kesehatan menembus Rp6,72 triliun hanya pada kuartal I-2026, melonjak 15,3% secara tahunan (*year-on-year*). Pada beberapa lini produk individu, rasio klaim (*loss ratio*) bahkan sempat melewati angka 100% hingga 200%, yang berarti nilai klaim jauh lebih besar daripada premi yang dikumpulkan [1]. Kondisi ini akan berdampak secara langsung terhadap tekanan finansial perusahaan asuransi. Salah satu faktor utamanya yang diidentifikasi sebagai pendorong kenaikan klaim ini adalah inflasi biaya medis, yang dipicu oleh kenaikan harga bahan baku farmasi serta kebutuhan perlengkapan medis [2].

Tantangan besar yang dihadapi oleh industri asuransi, yaitu menilai risiko secara akurat, menetapkan premi dengan adil, serta mengelola klaim dan cadangan secara efisien. Kurang tepat dalam penilaian risiko dapat menyebabkan ketidakseimbangan antara premi dan klaim sehingga berpengaruh pada stabilitas keuangan perusahaan dan reputasi perusahaan karena pemegang polis merasa dirugikan. Masalah ini dapat diatasi dengan sebuah metode yang dapat mengidentifikasi perbedaan karakteristik antar-kelompok pemegang polis, seperti Model Bühlmann–Straub. Model Bühlmann–Straub merupakan salah satu metode dengan pendekatan kredibilitas yang dapat mengestimasi klaim dengan mempertimbangkan variasi risiko dan eksposur antarkelompok secara efektif [3]. Karakteristik risiko seperti wilayah tempat tinggal dan status perokok memiliki pengaruh signifikan terhadap besar klaim asuransi, khususnya pada asuransi kesehatan. Oleh karena itu, membandingkan estimasi klaim berdasarkan kedua karakteristik risiko tersebut menggunakan model Bühlmann–Straub diharapkan dapat memberikan hasil estimasi yang lebih adil, akurat, dan mendukung pengambilan keputusan dalam penetapan premi yang proporsional.

Dengan demikian masalah yang dihadapi perusahaan asuransi yaitu tentang bagaimana menetapkan premi yang adil dan akurat di tengah keberagaman portofolio ini. Setiap tertanggung tentunya memiliki profil risiko yang berbeda, sehingga memberlakukan premi secara rata akan menyebabkan subsidi silang yang tidak adil. Faktor risiko seperti gaya hidup dan faktor demografis sangat mempengaruhi besar klaim karena dapat meningkatkan risiko penyakit [4]. Salah satu gaya hidup yang berpengaruh kuat terhadap peningkatan risiko kesehatan adalah kebiasaan merokok. Hasil penelitian menunjukkan bahwa wilayah tempat tinggal berpengaruh terhadap perilaku merokok [5]. Oleh karena itu, diperlukan sebuah model estimasi yang mampu mengakomodasikan perbedaan karakteristik risiko ini.

Dalam melakukan peninjauan ulang tarif premi, perusahaan asuransi dapat menggunakan teori kredibilitas. Teori ini menawarkan kerangka kerja matematis untuk mengatasi masalah heterogenitas tersebut [6]. Prinsip teori kredibilitas adalah bekerja dengan menentukan sejauh mana pengalaman klaim masa lalu suatu kelompok dapat dipercaya (kredibel) untuk digunakan dalam menentukan premi di masa mendatang [7]. Tentunya model ini menyeimbangkan pengalaman klaim historis dari suatu kelompok data individu dengan pengalaman klaim kolektif dari seluruh portofolio. Teori kredibilitas memiliki beberapa metode, yaitu *limited fluctuation (classical credibility)*, *greatest accuracy* yang terdiri dari model Buhlmann, Bühlmann–Straub, dan *Empirical Bayes*, dan metode Bayesian [8]. Dalam penelitian ini, metode yang akan digunakan adalah metode *greatest accuracy*. Pemilihan metode ini karena mampu mengestimasi

premi secara optimal melalui minimisasi *mean squared error* tanpa bergantung pada asumsi subjektif seperti pada metode Bayesian maupun batas eksposur pada model *limited fluctuation* [9].

Model yang dipilih dari metode *greatest accuracy* adalah model Bühlmann–Straub. Model Bühlmann–Straub dipilih karena kemampuannya yang teruji dalam menangani volume eksposur data yang bervariasi antara kelompok risiko [10]. Penelitian ini menunjukkan bahwa model Bühlmann–Straub dapat menghasilkan faktor kredibilitas yang bervariasi dengan estimasi premi yang lebih proposional terhadap eksposur risiko, sehingga sangat direkomendasikan untuk portofolio yang bersifat heterogen [10]. Sementara itu, berbagai penelitian telah fokus pada analisis sebaran klaim atau faktor inflasi medis, penerapan model ini tentunya diaplikasikan untuk membandingkan dampak kombinasi faktor gaya hidup perokok dan demografis dalam konteks data asuransi di Indonesia masih perlu diekplorasi lebih dalam.

Oleh karena itu, penelitian ini bertujuan untuk menganalisis dan membandingkan estimasi premi kredibilitas murni pada portofolio asuransi kesehatan dengan menggunakan model Bühlmann–Straub. Fokus perbandingan diarahkan pada estimasi besarnya klaim untuk kelompok yang dibentuk berdasarkan kombinasi dua faktor risiko utama, yaitu status perokok dan wilayah.

2 Metode Penelitian

2.1 Data Penelitian

Penelitian ini menggunakan data *healthcare insurance* di negara berkembang di dunia tahun 2023 yang bersumber dari Kaggle. Data tersebut terdiri atas 7 variabel yang terdiri dari data *age*, *sex*, *bmi*, *children*, *smoker*, *region*, dan *charges* (besar klaim). Data yang akan digunakan dalam penelitian ini, yaitu data pemegang polis dengan kategori *smoker* (*smoker* dan *nonsmoker*) dan *region* (*Southeast* dan *Northeast*).

2.2 Sebaran Besar Klaim

Besar klaim merupakan data dengan sebaran kontinu. Sebaran kontinu yang cocok untuk digunakan antara lain sebaran lognormal, dan Weibull [11]. Penelitian ini, mengasumsikan bahwa data besar klaim menyebar lognormal dan Weibull.

Peubah acak X menyebar lognormal dengan parameter $\mu \in \mathbb{R}$, $\sigma > 0$ jika mempunyai fungsi kepekatan peluang [12]

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{\left(\frac{\ln x - \mu}{\sigma}\right)^2}{2}\right), x > 0 \quad (1)$$

dengan nilai harapan X adalah

$$\mathbb{E}(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right). \quad (2)$$

Peubah acak X menyebar Weibull dengan parameter $\tau, \theta > 0$ jika mempunyai fungsi kepekatan peluang [12]:

$$f(x) = \frac{\tau \left(\frac{x}{\theta}\right)^{\tau} e^{-(x/\theta)^{\tau}}}{x}, x > 0 \quad (3)$$

dengan nilai harapannya adalah

$$E(X) = \theta \Gamma(1 + 1/\tau). \quad (4)$$

2.3 Pendugaan Parameter Sebaran Besar Klaim

Parameter sebaran besar klaim diduga menggunakan *Metode Maximum Likelihood*. Jika x_1, x_2, \dots, x_n merupakan contoh dari peubah acak yang saling bebas dan identik X_1, X_2, \dots, X_n serta menyebar lognormal dengan parameter μ dan σ , maka menurut [8] fungsi *loglikelihood*-nya adalah

$$l(\theta) = \ln \left(\prod_{j=1}^n \frac{1}{x_j \sigma \sqrt{2\pi}} \exp \left(-\frac{\left(\frac{\ln(x_j) - \mu}{\sigma} \right)^2}{2} \right) \right) \quad (5)$$

dan dugaan parameter adalah

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \ln x_i \quad (6)$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (\ln x_i - \hat{\mu}_i)^2. \quad (7)$$

Jika x_1, x_2, \dots, x_n merupakan contoh dari peubah acak yang saling bebas dan identik X_1, X_2, \dots, X_n serta menyebar Weibull dengan parameter τ dan θ , maka menurut [12] fungsi *loglikelihood*-nya adalah

$$l(\theta) = \ln \left(\prod_{j=1}^n \frac{\tau \left(\frac{x_j}{\theta} \right)^{\tau} e^{-\left(x_j/\theta \right)^{\tau}}}{x_j} \right) \quad (8)$$

dan dugaan parameter adalah

$$\hat{\theta} = \left(\frac{1}{n} \sum_{i=1}^n x_i^{\tau} \right)^{\frac{1}{\tau}} \quad (9)$$

$$\frac{1}{\tau} = \frac{\sum_{i=1}^n x_i^{\tau} \ln x_i}{\sum_{i=1}^n x_i^{\tau}} - \frac{1}{n} \sum_{i=1}^n \ln x_i. \quad (10)$$

2.4 Uji Kebaikan Suai Kolmogorov-Smirnov

Uji Kolmogorov-Smirnov (K-S) merupakan metode statistik uji nonparametrik yang digunakan untuk menilai kesesuaian sebaran data empiris terhadap sebaran teoritis tertentu [13]. Pada penelitian ini uji Kolmogorov Smirnov mempunyai hipotesis sebagai berikut:

- H_0 : data menyebar lognormal atau Weibull,
 H_1 : data tidak menyebar lognormal atau Weibull.

Uji sebaran lognormal dan Weibull K-S membandingkan fungsi sebaran kumulatif empiris $F_e(x)$ dengan fungsi sebaran kumulatif teoritis sebaran lognormal atau Weibull $F_X(x)$. Statistik uji yang digunakan dapat dinotasikan dengan persamaan berikut [12]:

$$D = \max_x |F_e(x) - F_X(x)|. \quad (11)$$

Fungsi sebaran kumulatif empiris $F_e(x)$ dapat dinotasikan dengan persamaan berikut:

$$F_e(x) = \frac{1}{n} \sum_{i=1}^n I(X_i \leq x) \quad (12)$$

dengan, $I(X_i \leq x)$ menyatakan fungsi indikator yang bernilai 1 jika $X_i \leq x$ dan 0 selainnya, serta n menyatakan banyaknya data.

Pengambilan keputusan dalam uji K-S umumnya menggunakan nilai p -value. Nilai p -value menunjukkan tingkat signifikansi α perbedaan antara fungsi sebaran kumulatif empiris $F_e(x)$ dan fungsi sebaran kumulatif untuk sebaran lognormal dan weibull $F_X(x)$.

Penelitian ini menggunakan tingkat signifikansi sebesar $\alpha = 0,05$ pada uji Kolmogorov–Smirnov untuk menentukan apakah data besar klaim menyebar lognormal atau Weibull. Keputusan uji didasarkan pada nilai p -value yang dihasilkan. Apabila p -value lebih besar dari 0,05, maka H_0 tidak ditolak, sehingga data dianggap menyebar lognormal atau Weibull. Sebaliknya, apabila p -value lebih kecil atau sama dengan 0,05, maka H_0 ditolak, yang berarti data tidak menyebar lognormal atau Weibull.

2.5 Model Bühlman Straub

Diketahui suatu portfolio asuransi dengan pengalaman n tahun. Untuk tahun ke- j terdapat m_j pemegang polis, $j = 1, \dots, n$. Besar klaim untuk peserta ke- k pada tahun ke j dilambangkan dengan $Y_{jk}, k = 1, \dots, m_j$. Misalkan Θ adalah peubah acak yang menyatakan Risiko. Diasumsikan $Y_{jk} | \Theta = \theta$ saling bebas dan identik untuk setiap j, k dengan

$$\mu(\theta) = \mathbb{E}[Y_{jk} | \Theta = \theta] \text{ dan } v(\theta) = \text{Var}(Y_{jk} | \Theta = \theta).$$

Parameter struktural didefinisikan sebagai

$$\mu = \mathbb{E}[\mu(\Theta)], \quad v = \mathbb{E}[v(\Theta)], \quad a = \text{Var}(\mu(\Theta)).$$

Misalkan X_j adalah rata-rata klaim pemegang polis pada tahun ke- j , yaitu

$$X_j = \frac{1}{m_j} \sum_{k=1}^{m_j} Y_{jk}.$$

Berdasarkan pengalaman masa lalu X_1, \dots, X_n , akan diestimasi X_{n+1} sebagai kombinasi linear dari rata-rata klaim di masa lalu, yaitu

$$X_{n+1} = \alpha_0 + \sum_{i=1}^n \alpha_i X_i.$$

Premi kredibilitas didefinisikan sebagai

$$P = \hat{\alpha}_0 + \sum_{i=1}^n \hat{\alpha}_i X_i$$

dengan koefisien $(\hat{\alpha}_0, \dots, \hat{\alpha}_n)$ dipilih sehingga meminimumkan rata-rata kuadrat galat

$$Q(\alpha_0, \dots, \alpha_n) = \mathbb{E} \left[\left(X_{n+1} - \alpha_0 - \sum_{i=1}^n \alpha_i X_i \right)^2 \right].$$

Menurut [12], koefisien $(\hat{\alpha}_0, \dots, \hat{\alpha}_n)$ memenuhi persamaan normal berikut:

$$\mathbb{E}[X_{n+1}] = \hat{\alpha}_0 + \sum_{i=1}^n \hat{\alpha}_i \mathbb{E}[X_i],$$

$$\text{Cov}(X_j, X_{n+1}) = \sum_{i=1}^n \hat{\alpha}_i \text{Cov}(X_i, X_j), j = 1, \dots, n.$$

Teorema 1. (Klugman *et al.* [12]) Untuk model Bühlman Straub, diperoleh:

$$\mathbb{E}[X_j|\Theta] = \mu(\Theta), \quad \text{Var}(X_j|\Theta) = \frac{v(\Theta)}{m_j}.$$

$$\mathbb{E}[X_j] = \mu, \quad \text{Var}(X_j) = \frac{v}{m_j} + a, \quad \text{Cov}(X_i, X_j) = a, \text{ untuk } i \neq j,$$

dan premi kredibilitas Bühlman Straub adalah

$$P = Z \bar{X} + (1 - Z)\mu,$$

dengan

$$Z = \frac{m}{\frac{v}{a} + m}, \quad \bar{X} = \sum_{i=1}^n \frac{m_i}{m} X_i \text{ dan } m = \sum_{i=1}^n m_i.$$

Masalah berikutnya adalah menentukan penduga parameter struktural μ, v dan a pada premi kredibilitas Bühlman Straub. Akan digunakan data masa lalu untuk menduga parameter tersebut. Metode yang digunakan adalah *Empirical Bayes Estimation*.

Misalkan diberikan notasi-notasi berikut:

- Terdapat r kelompok dan n_i tahun pengamatan untuk kelompok ke- i , $i = 1, \dots, r$.
- m_{ij} menyatakan banyaknya anggota kelompok ke i pada tahun ke- j , $\underline{m}_i = (m_{i1}, \dots, m_{in_i})$, $m_i = \sum_{j=1}^{n_i} m_{ij}$ dan $m = \sum_{i=1}^r m_i$.
- X_{ij} menyatakan rata-rata klaim untuk kelompok ke- i pada tahun ke- j , $\underline{X}_i = (X_{i1}, \dots, X_{in_i})$, $\bar{X}_i = \frac{1}{m_i} \sum_{j=1}^{n_i} m_{ij} X_{ij}$, dan $\bar{X} = \frac{1}{m} \sum_{i=1}^r \bar{X}_i$.

Diasumsikan bahwa $\{\bar{X}_i\}$ saling bebas, $\{\Theta_i\}_{i=1}^r$ saling bebas dan identik, serta $\{X_{ij}|\Theta_i\}_{i=1}^r$ saling bebas dengan *mean* $\mathbb{E}[X_{ij}|\Theta_i] = \mu(\Theta_i)$ dan *varians* $V(X_{ij}|\Theta_i) = \frac{v(\Theta_i)}{m_{ij}}$. Parameter strukturalnya adalah

$$\mu = \mathbb{E}[\mu(\Theta_i)], \quad v = \mathbb{E}[v(\Theta_i)], \quad a = \text{Var}(\mu(\Theta_i)).$$

Teorema 2. (Klugman *et al.* [12]) Non-parametric estimation untuk model Bühlman Straub adalah:

- Penduga tak bias untuk μ adalah $\hat{\mu} = \bar{X}$.
- Penduga tak bias untuk v adalah $\hat{v} = \frac{1}{\sum_{i=1}^r (n_i - 1)} \sum_{i=1}^r (n_i - 1) \hat{v}_i$, dengan $\hat{v}_i = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} m_{ij} (X_{ij} - \bar{X}_i)^2$ juga merupakan penduga tak bias untuk v .
- Penduga tak bias untuk a adalah $\hat{a} = \frac{m}{m^2 - \sum_{i=1}^r m_i^2} \left(\sum_{i=1}^r m_i (\bar{X}_i - \bar{X})^2 - (r - 1) \hat{v} \right)$.

Premi kredibilitas Bühlman Straub untuk kelompok ke- i adalah

$$P_i = Z_i \bar{X}_i + (1 - Z_i)\mu, \text{ dengan } Z_i = \frac{m_i}{\frac{v}{a} + m_i}, i = 1, \dots, r.$$

Penduga premi kredibilitas Bühlman Straub untuk kelompok ke- i adalah

$$\hat{P}_i = \hat{Z}_i \bar{X}_i + (1 - \hat{Z}_i)\hat{\mu}, \text{ dengan } \hat{Z}_i = \frac{m_i}{\frac{\hat{v}}{a} + m_i} i = 1, \dots, r.$$

Untuk pendekatan semiparametrik, \bar{X}_i dihipotesis oleh penduga parameter \hat{X}_i yang diperoleh dari *mean* sebaran yang sesuai untuk data kelompok ke- i .

3 Hasil dan Pembahasan

3.1 Statistik Deskriptif Data Klaim

Data penelitian ini mencakup 688 observasi pemegang polis dari wilayah *Northeast* dan *Southeast*. Analisis statistik deskriptif dilakukan untuk memahami karakteristik risiko klaim berdasarkan status perokok (*smoker*) dan wilayah. Ringkasan statistik disajikan pada Tabel 1.

Tabel 1. Statistik deskriptif besar klaim berdasarkan kelompok

Kelompok	Karakteristik	Mean (\bar{X})	Median	Std. Dev (s)	Min	Max	Skewness
1	<i>Smoker - Northeast</i>	29673,5	28101,3	11589,56	12829,4	58571,0	0,26
2	<i>Nonsmoker - Northeast</i>	9165,53	8342,91	6157,20	1694,80	32108,6	1,32
3	<i>Smoker - Southeast</i>	34845,0	37484,4	11324,77	16577,7	63770,4	-0,09
4	<i>Nonsmoker - Southeast</i>	8032,22	6652,53	6137,33	1121,87	36580,2	1,52

Berdasarkan Tabel 1, kelompok perokok (1,3) memiliki rata-rata klaim pada kisaran 29000-34000 yang jauh lebih tinggi dibandingkan kelompok bukan perokok (2,4) yaitu kisaran 8000-9000. Selain itu, nilai standar deviasi (s) pada kelompok perokok juga hampir dua kali lipat lebih besar. Hal ini mengindikasikan bahwa kelompok perokok tidak hanya memiliki ekspektasi klaim yang tinggi, tetapi memiliki risiko kesehatan yang lebih besar.

3.2 Sebaran Kelompok

Dalam penelitian ini, estimasi premi kredibilitas Bühlmann–Straub didasarkan pada asumsi bahwa besar klaim mengikuti sebaran yang sesuai dengan karakteristik data, sehingga perlu dilakukan identifikasi sebaran yang tepat sebelum perhitungan premi.

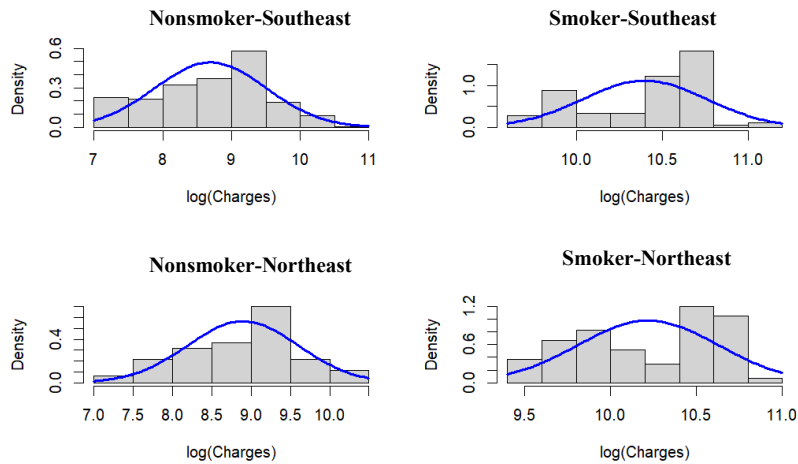
Tabel 2 menunjukkan sebaran yang sesuai untuk setiap kelompok pemegang polis. Dugaan parameter untuk setiap sebaran ditentukan menggunakan Metode *Maximum Likelihood*. Untuk menguji kesesuaian sebaran teoritis dengan data setiap kelompok digunakan Uji Kolmogorov-Smirnov.

Tabel 2. Sebaran besar klaim untuk setiap kelompok pemegang polis

Kelompok	Sebaran	AIC	Parameter	p -value uji K-S
1	Lognormal	1443	$\hat{\mu} = 10,21777, \hat{\sigma} = 0,48331$	0,0839
2	Weibull	5131	$\hat{\tau} = 1,58744, \hat{\theta} = 10282,1$	0,2013
3	Weibull	1957	$\hat{\tau} = 3,51139, \hat{\theta} = 38814,49$	0,1395
4	Lognormal	5410	$\hat{\mu} = 8,69693, \hat{\sigma} = 0,807059$	0,0548

Dari Tabel 2, terlihat bahwa data kelompok 1 dan 4 menyebar lognormal, serta kelompok 2 dan 3 menyebar Weibull. Hasil uji K-S menunjukkan bahwa nilai p -value semua kelompok di atas tingkat kepercayaan $\alpha = 0,05$. Ini berarti bahwa tidak terdapat

bukti yang cukup kuat untuk menolak sebaran tersebut. Plot fungsi kepekatan peluang vs histogram peluang disajikan pada Gambar 1.



Gambar 1. Plot fungsi kepekatan peluang vs histogram frekuensi setiap kelompok

Dari Gambar 1 terlihat bahwa grafik fungsi kepekatan peluang setiap kelompok cukup menghampiri histogram frekuensinya.

3.3 Pendugaan Semiparametrik Model Bühlman Straub

Pendugaan semiparametrik kredibilitas Bühlmann-Straub pada penelitian ini diawali dengan menghitung parameter-parameter statistik dari keempat kelompok pemegang polis berdasarkan dua karakteristik yang menjelaskan tingkat risiko untuk pemegang polis, yaitu status perokok (*smoker* atau *non-smoker*) dan wilayah tempat tinggal (*northeast* atau *southeast*). Parameter-parameter statistik dari keempat kelompok pemegang polis pada penelitian diuraikan sebagai berikut.

3.3.1 Hypothetical mean kelompok \bar{X}_i

Pada penelitian ini jumlah kelompok (r) adalah 4. Nilai *hypothetical mean* \bar{X}_i untuk $i = 1, 2, 3, 4$ dihitung menggunakan Persamaan (2) dan (4) dan dugaan parameter sebaran yang terdapat pada Tabel 2. Total pemegang polis pada kelompok ke- i yaitu m_i dan nilai penduga nilai *hypothetical mean* \bar{X}_i yaitu \hat{X}_i ditunjukkan pada Tabel 3 berikut.

Kelompok	Karakteristik	Total Pemegang Polis (m_i)	\hat{X}_i
1	<i>Smoker-Northeast</i>	67	29766,537
2	<i>Nonsmoker-Northeast</i>	257	9225,498
3	<i>Smoker-Southeast</i>	91	34929,256
4	<i>Nonsmoker-Southeast</i>	273	8288,356

Berdasarkan data pada Tabel 3 diperoleh nilai m sebagai berikut:

$$m = \sum_{i=1}^4 m_i = 67 + 257 + 91 + 273 = 688.$$

Menurut Tabel 3, kelompok *smoker-southeast* memiliki nilai *hypotetical mean* tertinggi sebesar 34929,256 kemudian kelompok *smoker-northeast* sebesar 29766.537 sehingga kedua kelompok tersebut tergolong kelompok risiko tinggi dengan ekspektasi klaim terbesar. Sebaliknya, kelompok *non-smoker-northeast* memiliki nilai *hypotetical mean* sebesar 9225,498 dan kelompok *non-smoker-southeast* sebesar 8288.356 memiliki nilai *hypotehtical mean* lebih rendah sehingga kedua kelompok tersebut memiliki risiko relatif kecil dan ekspektasi klaim yang sedikit. Hal ini mengindikasikan bahwa karakteristik perokok meningkatkan ekspektasi klaim sedangkan wilayah tempat tinggal hanya memperkuat variasi.

3.3.2 Hypothetical mean $\hat{\mu}$

Rataan klaim untuk *hypothetical mean* kelompok pemegang polis didapatkan dari persamaan berikut:

$$\begin{aligned}\hat{\mu} &= \bar{X} = \frac{\sum_{i=1}^4 m_i \bar{X}_i}{\sum_{i=1}^4 m_i} \\ &= \frac{(67 \times 29766,537) + (257 \times 9225,498) + (91 \times 34929,256) + (273 \times 8288.356)}{67 + 257 + 91 + 273} \\ &= 14253,77.\end{aligned}$$

Estimasi rata-rata klaim untuk *hypothetical mean* kelompok pemegang polis pada penelitian ini sebesar 14253,77. Nilai ini mengindikasikan besar klaim rata-rata yang diajukan oleh setiap pemegang polis selama periode pengamatan.

3.3.3 Variance process \hat{v}

Variance process didapatkan dari persamaan berikut:

$$\hat{v} = \frac{\sum_{i=1}^4 \sum_{j=1}^{n_i} m_{ij} (X_{ij} - \bar{X}_i)^2}{\sum_{i=1}^4 (n_i - 1)} = 59032426.$$

Estimasi *variance process* untuk semua kelompok pemegang polis pada penelitian ini sebesar 59032426. Nilai ini mengindikasikan bahwa variasi klaim di dalam masing-masing kelompok pemegang polis tinggi sehingga besar klaim antar-pemegang polis dalam kelompok pemegang polis yang sama berbeda secara signifikan.

3.3.4 Varians *hypothetical mean* \hat{a}

Varians *hypothetical mean* pada penelitian ini didapatkan dari persamaan berikut:

$$\begin{aligned}\hat{a} &= \frac{\sum_{i=1}^4 m_i (\bar{X}_i - \bar{X})^2 - (r - 1) \hat{v}}{m - \frac{1}{m} \sum_{i=1}^4 m_i^2} \\ &= \frac{\sum_{i=1}^4 m_i (\bar{X}_i - \bar{X})^2 - 3 \hat{v}}{m - \frac{1}{m} \sum_{i=1}^4 m_i^2} \\ &= 152779527.\end{aligned}$$

Estimasi varians *hypothetical mean* pada penelitian sebesar $\hat{a} = 152779527$. Nilai ini mengindikasikan bahwa terdapat perbedaan yang sangat besar pada rata-rata klaim antar-kelompok pemegang polis.

3.3.5 Koefisien kredibilitas Bühlmann-Straub

Koefisien kredibilitas Bühlmann-Straub didapatkan dari persamaan berikut:

$$\hat{k} = \frac{\hat{v}}{\hat{a}} = \frac{59032426}{152779527} = 0,3863896.$$

Nilai ini mengindikasikan bahwa variasi rata-rata besar klaim antar-pemegang polis dalam kelompok pemegang polis yang sama lebih besar 0.39 kali lebih besar dibandingkan dengan variasi rata-rata besar klaim antar-kelompok pemegang polis. Hal ini menunjukkan bahwa fluktuasi klaim terjadi karena risiko individu pemegang polis dibandingkan pada kelompok pemegang polis.

3.3.6 Faktor kredibilitas Bühlmann-Straub \hat{Z}_i

Faktor kredibilitas Bühlmann-Straub untuk setiap kelompok pemegang polis ke- i didapatkan menggunakan persamaan berikut:

$$\hat{Z}_i = \frac{m_i}{m_i + \hat{k}}, \quad i = 1, 2, 3, 4.$$

Faktor kredibilitas Bühlmann-Straub \hat{Z}_i pada penelitian ini ditunjukkan pada Tabel 4 berikut.

Tabel 4. Faktor kredibilitas Bühlmann-Straub (BS)

Kelompok	Karakteristik	Faktor Kredibilitas BS
1	<i>Smoker-Northeast</i>	0,9942661
2	<i>Nonsmoker-Northeast</i>	0,9984988
3	<i>Smoker-Southeast</i>	0,9957719
4	<i>Nonsmoker-Southeast</i>	0,9985867

Menurut Tabel 4, faktor kredibilitas Bühlmann-Straub \hat{Z}_i untuk keempat kelompok pemegang polis berada pada kisaran 0,99. Nilai ini mengindikasikan bahwa besar klaim masing-masing kelompok pemegang polis memiliki tingkat kepercayaan yang sangat tinggi dalam penentuan premi untuk masing-masing kelompok pemegang polis. Dengan demikian, premi untuk masing-masing kelompok pemegang polis akan lebih dipengaruhi oleh rata-rata klaim kelompok pemegang polis dibandingkan rata-rata klaim seluruh pemegang polis sehingga menunjukkan karakteristik risiko setiap kelompok secara lebih akurat.

3.3.7 Premi kredibilitas kelompok pemegang polis \hat{P}_i

Premi kredibilitas kelompok pemegang polis untuk setiap kelompok pemegang polis ke- i didapatkan menggunakan persamaan berikut:

$$\hat{P}_i = \hat{Z}_i \bar{X}_i + (1 - \hat{Z}_i) \bar{X}, \quad i = 1, 2, 3, 4.$$

Premi kelompok pemegang polis untuk setiap kelompok pemegang polis pada penelitian ini ditunjukkan pada Tabel 5 berikut.

Tabel 5. Premi kelompok pemegang polis

Kelompok	Karakteristik	Premi
1	<i>Smoker-Northeast</i>	29677,587
2	<i>Non-smoker-Northeast</i>	9233,046
3	<i>Smoker-Southeast</i>	34841,838
4	<i>Non-smoker-Southeast</i>	8296,787

Menurut Tabel 5, kelompok *smoker-southeast* memiliki premi terbesar yaitu sebesar 34841,838 kemudian kelompok *smoker-northeast* dengan premi sebesar 29677,587 yang mencerminkan kelompok dengan karakteristik penjas risiko *smoker* memiliki rata-rata klaim yang tinggi hanya dibedakan dengan wilayah tempat tinggal. Sebaliknya, kelompok *non-smoker-northeast* memiliki premi lebih rendah, yaitu sebesar 9233,046 dan *non-smoker-southeast* dengan premi terendah sebesar 8296,787 yang mencerminkan kelompok dengan karakteristik penjas risiko *non-smoker* memiliki rata-rata klaim yang rendah hanya dibedakan dengan wilayah tempat tinggal.

Sebagai perbandingan, rata-rata besar klaim seluruh pemegang polis adalah 14253,77. Nilai ini merupakan estimasi untuk premi tunggal bersih bagi seluruh pemegang polis. Pemegang polis yang tidak merokok mendapatkan premi yang lebih besar dibandingkan dengan seharusnya. Sebaliknya, pemegang polis perokok justru diuntungkan, sebab beban premi yang mereka bayar menjadi lebih rendah dibandingkan risiko klaim aktual. Kondisi ini mengurangi keadilan aktuaria karena premi tidak lagi mencerminkan *fair pricing* sesuai ekspektasi kerugian tiap individu.

4 Simpulan

Hasil penelitian menunjukkan bahwa penerapan model kredibilitas Bühlmann–Straub efektif dalam menghasilkan perhitungan premi asuransi kesehatan yang lebih adil dan akurat pada portofolio yang heterogen. Berdasarkan analisis data, status perokok teridentifikasi sebagai faktor risiko utama yang memicu lonjakan besar klaim yang jauh lebih signifikan dibandingkan faktor wilayah tempat tinggal.

Data klaim kelompok risiko paling cocok dimodelkan menggunakan sebaran lognormal dan Weibull. Hasil perhitungan menunjukkan faktor kredibilitas (Z) yang sangat tinggi, yakni berkisar pada 0,99. Angka ini mengindikasikan bahwa data historis dari masing-masing kelompok risiko memiliki tingkat kepercayaan yang sangat tinggi. Hal tersebut menggambarkan bahwa dalam menetapkan premi, perusahaan asuransi sebaiknya lebih memprioritaskan riwayat klaim spesifik dari kelompok tersebut (kredibilitas individu) dibandingkan sekadar menggunakan rata-rata klaim seluruh pemegang polis.

Hasil estimasi premi menempatkan kelompok perokok di wilayah *Southeast* sebagai pembayar premi tertinggi (34841,838) dan juga sebagai kelompok kategori bukan perokok dengan tarif terendah (8296,787). Penelitian ini menunjukkan bahwa penggunaan model Bühlmann–Straub dapat membantu perusahaan menghindari subsidi silang yang tidak adil. Nasabah risiko rendah tidak perlu menanggung beban biaya dari nasabah risiko tinggi sehingga stabilitas finansial perusahaan tetap terjaga.

Daftar Pustaka

- [1] Asosiasi Asuransi Jiwa Indonesia. 2026. *Laporan Kinerja Industri Asuransi Jiwa Indonesia Kuartal I-2026*. Jakarta: AAJI.
- [2] Bani P. 2025. Asimetri informasi dan moral hazard: tinjauan literatur tentang dampaknya terhadap klaim asuransi kesehatan. *Journal of Economics and Business UBS*. 14: 593–611.
- [3] Susanti D, Sukono S. 2023. calculating premium credibility using the Bühlmann–Straub model with nonparametric assessment. *International Journal of Global Optimization and Its Application*. 1(1): 1-7.

- [4] [Generali Indonesia]. 2022. 7 Faktor yang mempengaruhi harga tarif premi seseorang." [Daring]. <https://www.generali.co.id/id/healthyliving/2/7-faktor-yang-mempengaruhi-harga-tarif-premi-seseorang>.
- [5] Zeng Q, Zhang C, Su F, Wan Y, Tu W, Hu H. 2024. Prevalence, cessation, and geographical variation of smoking among middle-aged and elderly adults in China: a population-based study. *Tobacco Induced Diseases*. 22, art. no. 42. doi: [10.18332/tid/190247](https://doi.org/10.18332/tid/190247).
- [6] Pramono, S. (2016). Penerapan ilmu matematika dalam perlindungan kehidupan terhadap risiko. Di dalam: *Peran Matematika, Sains, dan Teknologi dalam Mendukung Gaya Hidup Perkotaan (Urban Lifestyle) yang Berkualitas*. Universitas Terbuka, Tangerang Selatan, pp. 243-274. ISBN 978-602-392-160-7 (e)
- [7] Sari DA. 2008. Kredibilitas pengalaman klaim asuransi kesehatan kelompok : studi kasus pada PT.(Persero) Asuransi Kesehatan Indonesia [Tesis]. Universitas Indonesia.
- [8] Atkinson DB. 2019. Overview of credibility, in: *Credibility Applications for Life and Health Insurers and Pension Plans*, US: Society of Actuaries, Chapter 1.
- [9] Atkinson DB. 2019. Greatest accuracy method, in: *Credibility Applications for Life and Health Insurers and Pension Plans*, US: Society of Actuaries, Chapter 3.
- [10] Abdul ARS, Putra RK, Budiantara IN. 2022. Estimasi premi kredibilitas asuransi umum menggunakan model Bühlmann dan Bühlmann-Straub dengan pendekatan empirical Bayes. *Jurnal Sains dan Seni ITS*. 11(1): 14-20.
- [11] Ahmad Z, Mahmoudi E, Hamedani G. 2022. A class of claim distributions: properties, characterizations and applications to insurance claim data. *Commun Stat Theory Methods*. 51(7): 2183–2208. doi: [10.1080/03610926.2020.1772306](https://doi.org/10.1080/03610926.2020.1772306).
- [12] Klugman SA, Panjer HH, Willmot GE. 2019. *Loss Models from Data to Decisions*, 5 ed. Hoboken, NJ : John Wiley and Sons, Inc.
- [13] Cardoso DO, Galeno TD. 2022. Online evaluation of the Kolmogorov–Smirnov test on arbitrarily large samples. *Journal of Computational Science*. 67: 101959. doi: [10.1016/j.jocs.2023.101959](https://doi.org/10.1016/j.jocs.2023.101959).