

BAB I PENDAHULUAN

1.1 Latar Belakang

BPJS kesehatan adalah lembaga sosial yang dibentuk untuk menyelenggarakan program jaminan kesehatan di seluruh wilayah Indonesia. Dijelaskan dalam Undang-Undang Nomor 40 Tahun 2004 tentang Sistem Jaminan Sosial Nasional bahwa Jaminan kesehatan diselenggarakan secara nasional dengan tujuan menjamin agar peserta memperoleh manfaat pemeliharaan kesehatan dan perlindungan dalam memenuhi kebutuhan dasar kesehatan.

Kabupaten Pangkajene dan Kepulauan adalah kabupaten di Provinsi Sulawesi Selatan, Indonesia dengan luas wilayah 1.112,29 km² dengan jumlah penduduk pada tahun 2020 adalah 345.775 jiwa, yang merupakan salah satu kabupaten dari 334 kabupaten penerima penghargaan Universal Health Coverage (UHC) pada tahun 2023 dari 514 kabupaten se-Indonesia. UHC merupakan sistem penjaminan kesehatan yang memastikan setiap warga dalam populasi memiliki akses yang adil terhadap pelayanan kesehatan. Bupati Pangkajene dan Kepulauan, *Muhammad Yusran Lalogau* mengatakan saat ini 96,7 persen penduduk Pangkajene dan Kepulauan telah tercatat sebagai peserta Jaminan Kesehatan Nasional (JKN) yang diselenggarakan BPJS Kesehatan (Tribun Timur:2023).

Peraturan BPJS Kesehatan Nomor 4 Tahun 2014 mewajibkan seluruh Warga Negara Indonesia untuk mendaftarkan diri beserta seluruh anggota keluarga yang tercantum dalam satu Kartu Keluarga sebagai peserta BPJS Kesehatan. Oleh karena itu, setiap anggota keluarga wajib terdaftar dan membayar iuran BPJS Kesehatan setiap bulannya.

Berdasarkan Statistik Madya Badan Pusat Statistik (BPS) Provinsi Sulawesi Selatan, yang disampaikan oleh Suri Handayani, tingkat kemiskinan di Sulawesi Selatan pada Maret 2023 tercatat sebesar 8,70 persen. Angka ini menunjukkan adanya peningkatan sebesar 0,04 persen dibandingkan dengan kondisi pada September 2022, serta meningkat sebesar 0,07 persen dibandingkan dengan Maret 2022. Dari seluruh wilayah kabupaten/kota di Sulawesi Selatan, Kabupaten Pangkajene dan Kepulauan mencatatkan persentase penduduk miskin tertinggi, yaitu sebesar 13,92 persen. Artinya diperkirakan sekitar 48,146 individu hidup di bawah garis kemiskinan sehingga ini dapat meningkatkan risiko seseorang mengalami penunggakan iuran BPJS Kesehatan dikarenakan Kemiskinan yang sering kali terkait erat dengan keterbatasan finansial dan akses terhadap sumber daya yang diperlukan untuk memenuhi kebutuhan dasar, termasuk layanan kesehatan. Kabupaten Pangkajene dan Kepulauan, meskipun telah menerima penghargaan UHC pada tahun 2023, Tingkat kemiskinan yang tinggi tersebut dapat memengaruhi kemampuan masyarakat untuk an BPJS Kesehatan secara konsisten, sehingga meningkatkan risiko gagal nan merupakan masalah yang kompleks dan dipengaruhi oleh berbagai faktor rkaitan, seperti tingkat pendapatan, kesehatan, pendidikan, akses terhadap ia, lokasi geografis, jenis kelamin, dan kondisi lingkungan (KThomas, Ali, dkk.,



Pada bulan Juli 2024, data yang diperoleh dari Kantor BPJS Kesehatan Makassar menunjukkan bahwa jumlah masyarakat yang menunggak iuran BPJS Kesehatan di Kabupaten Pangkep mencapai 54.291 orang, dengan total tunggakan sebesar Rp90.798.437.059. Secara spesifik, di wilayah Kecamatan Minasa Tene, terdapat 6.991 masyarakat yang menunggak dengan total tunggakan sebesar Rp12.084.915.640. Sementara itu, di wilayah kecamatan Bonto Perak, tercatat 1.429 masyarakat yang menunggak dengan total tunggakan sebesar Rp2.654.190.493. Besarnya angka tunggakan ini menunjukkan adanya masalah dalam kepatuhan masyarakat terhadap kewajiban pembayaran iuran BPJS Kesehatan yang berpotensi mengganggu kelancaran pendanaan dan kualitas pelayanan kesehatan. Sehingga perlu mengetahui untuk mengidentifikasi faktor-faktor yang memengaruhi tingginya angka tunggakan BPJS Kesehatan di Kabupaten Pangkajene dan kepulauan.

Faktor yang memengaruhi kepatuhan dalam membayar iuran jaminan kesehatan yaitu pengetahuan, pendidikan, pekerjaan, ketersediaan tempat pelayanan, jarak ke tempat pelayanan, kemampuan dan kemauan membayar iuran, persepsi terhadap tempat pelayanan kesehatan dan motivasi. Variabel kemampuan dan kemauan dalam membayar iuran JKN tidak ditelusuri karena ketika telah menjadi peserta BPJS Kesehatan maka seseorang telah dianggap mampu dan mau untuk membayar iuran namun masih banyak peserta BPJS Kesehatan yang tidak patuh dalam membayar iuran.

Mengacu pada Peraturan Presiden Nomor 82 Tahun 2018 tentang Jaminan Kesehatan, yang menetapkan bahwa peserta BPJS Kesehatan dengan tunggakan iuran dapat melunasi kewajibannya dalam batas maksimal 24 bulan atau setara dengan dua tahun, penelitian ini menerapkan waktu observasi selama dua tahun untuk menganalisis risiko gagal bayar secara lebih menyeluruh.

Dengan demikian pada penelitian ini digunakan analisis survival. Analisis survival atau analisis ketahanan hidup adalah istilah yang digunakan untuk menggambarkan analisis data yang berhubungan dengan waktu, mulai dari waktu awal yang terdefinisi dengan baik hingga terjadinya suatu peristiwa atau titik akhir tertentu (Collett, 2015). Sehingga dalam penelitian ini waktu bertahannya peserta BPJS kesehatan di kabupaten pangkajene dan kepulauan untuk tidak mengalami penunggakan selama masa pengamatan, seperti tidak adanya keinginan untuk melanjutkan (*loss to follow up*), keluar (*drop out*), penghentian (*termination*) dan lain sebagainya yang dapat memberhentikan peserta BPJS kesehatan melakukan pembayaran iuran.

Sehingga dalam rangka meningkatkan pemahaman tentang faktor-faktor yang memengaruhi gagal bayar BPJS Kesehatan di Kabupaten Pangkajene dan Kepulauan, penelitian ini bertujuan untuk melakukan analisis menggunakan pendekatan model Cox Proportional Hazard. Pendekatan ini memungkinkan untuk mengidentifikasi dan memperkirakan dampak relatif dari berbagai variabel risiko terhadap waktu terjadinya gagal



1 memahami faktor-faktor yang berkontribusi terhadap gagal bayar BPJS arapkan dapat meningkatkan kepatuhan pembayaran iuran BPJS Kesehatan atkan akses masyarakat terhadap layanan kesehatan yang berkualitas di gkajene dan Kepulauan.

1.2 Rumusan Masalah

Berdasarkan konteks masalah yang telah dijelaskan, maka rumusan permasalahan dalam penelitian ini adalah sebagai berikut:

1. Faktor demografi apa saja yang secara signifikan memengaruhi peserta mengalami gagal bayar BPJS Kesehatan di Kabupaten Pangkajene dan Kepulauan?
2. Bagaimana memodelkan laju gagal bayar iuran BPJS Kesehatan menggunakan model Cox dengan kovariat time-dependent untuk menganalisis faktor-faktor demografi yang memengaruhi risiko gagal bayar peserta di Kabupaten Pangkajene dan Kepulauan?

1.3 Tujuan dan Manfaat

1.3.1 Tujuan penelitian

Berdasarkan pada rumusan masalah yang telah dijelaskan sebelumnya, maka tujuan yang ingin dicapai adalah untuk mengetahui faktor yang signifikan mempengaruhi risiko gagal bayar iuran BPJS Kesehatan dan mengidentifikasi pola gagal bayar iuran berdasarkan karakteristik demografi masyarakat di kabupaten pangkajene dan kepulauan

1.3.2 Manfaat penelitian

Manfaat dari penelitian ini yaitu sebagai berikut:

1. Manfaat Teoretis
 - a. Menambah wawasan dan literatur ilmiah mengenai analisis survival dengan model Cox Proportional Hazard berbasis time-dependent dalam mempelajari risiko gagal bayar iuran BPJS Kesehatan.
 - b. Memberikan kontribusi akademik dalam pengembangan metode statistik untuk analisis risiko pada data yang melibatkan variabel waktu.
2. Manfaat Praktis
 - a. Memberikan informasi kepada pihak BPJS Kesehatan mengenai faktor-faktor utama yang memengaruhi risiko gagal bayar peserta, khususnya di Kabupaten Pangkajene dan Kepulauan, sehingga dapat digunakan sebagai dasar untuk ancing kebijakan yang lebih efektif.
 - b. Membantu pemerintah daerah dan instansi terkait dalam menyusun program intervensi untuk meningkatkan tingkat pembayaran iuran peserta BPJS Kesehatan kabupaten pangkajene dan kepulauan.



1.4 Teori

1.4.1 BPJS Kesehatan

BPJS Kesehatan merupakan lembaga sosial yang dibentuk dengan tujuan untuk menyelenggarakan program-program jaminan sosial di Indonesia. Berdasarkan UU No. 24 Tahun 2011, BPJS akan menggantikan sejumlah lembaga jaminan sosial yang ada, seperti lembaga asuransi kesehatan PT Askes Indonesia yang akan diubah menjadi BPJS Kesehatan. BPJS adalah badan hukum publik milik negara yang non-profit dan bertanggung jawab langsung kepada Presiden. Terdapat dua jenis BPJS, yaitu BPJS Kesehatan dan BPJS Ketenagakerjaan.

1. Jaminan Kesehatan Nasional (JKN)

Jaminan sosial adalah bentuk perlindungan sosial yang bertujuan memastikan seluruh masyarakat dapat memenuhi kebutuhan dasar hidup yang layak. Sistem Jaminan Sosial Nasional (SJSN) adalah mekanisme penyelenggaraan program jaminan sosial yang dikelola oleh Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan dan BPJS Ketenagakerjaan. SJSN menggunakan sistem asuransi kesehatan yang bersifat wajib, yang diatur dalam UU No. 40 Tahun 2004 tentang SJSN. Program ini bertujuan untuk memenuhi kebutuhan dasar kesehatan bagi seluruh masyarakat, dengan memberikan perlindungan kesehatan kepada setiap individu yang telah membayar iuran atau kepada mereka yang iurannya dibayar oleh pemerintah, seperti bagi penerima bantuan iuran (PBI).

2. Risiko Gagal Bayar.

Mengacu pada peraturan yang mengatur kewajiban pembayaran iuran BPJS Kesehatan, yang tercantum dalam Peraturan Presiden No. 82 Tahun 2018 dan peraturan-peraturan terkait lainnya. Jika peserta gagal membayar iuran tepat waktu, peserta dapat kehilangan akses terhadap layanan kesehatan yang dijamin oleh BPJS Kesehatan, serta status kepesertaan mereka bisa menjadi tidak aktif. Selain itu, peserta yang menunggak juga harus melunasi tunggakan dan dapat dikenakan denda sebelum dapat kembali mendapatkan layanan. Oleh karena itu, menjaga kelancaran pembayaran iuran sangat penting untuk memastikan perlindungan kesehatan tetap terjaga.

1.4.2 Analisis *survival*

Analisis *survival* adalah metode statistik untuk menilai durasi hingga kejadian terjadi. Dalam konteks ini, waktu biasanya diukur mulai dari titik awal tertentu hingga terjadinya peristiwa yang



penelitian seperti kegagalan, kematian, kekambuhan penyakit ataupun peristiwa. Tujuan utama analisis ini yaitu estimasi distribusi waktu hingga terjadi peristiwa atau perbandingan pola *survival* antar kelompok dan pemodelan pengaruh variabel terhadap risiko kejadian dengan menggunakan model regresi *Cox proportional hazards* yang memberikan kerangka yang kuat untuk memahami dinamika waktu kejadian berdasarkan inferensi dari data *Survival* (Kleinbaum & Klein, 2012).

Analisis *survival* menggunakan dua fungsi utama, yaitu *survival function*, yang mengukur probabilitas bertahan lebih lama dari waktu tertentu tanpa mengalami peristiwa, dan *hazard function*, yang menunjukkan laju atau risiko terjadinya peristiwa pada waktu tertentu (Hosmer et al., 2008). Untuk menentukan waktu *survival*, terdapat tiga aspek penting yang harus diperhatikan, yaitu:

1. Waktu awal pencatatan (*time origin atau start-point*)
Waktu awal dalam analisis *survival*, yaitu waktu awal terjadinya kejadian seperti waktu awal seseorang didiagnosa penyakit atau waktu awal melakukan pengobatan dari penyakitnya sehingga waktu awal yang digunakan untuk penelitian ini yaitu waktu awal peserta mengalami penunggakan iuran dari Januari 2022 hingga Desember 2023.
2. Waktu akhir pencatatan (*event time atau end-point*)
Waktu kegagalan dalam analisis *survival*, yaitu masa dimana peristiwa terjadi pada subjek selama penelitian berlangsung, seberapa lama terjadinya penunggakan sejak hingga peserta kembali melakukan pembayaran iuran BPJS Kesehatannya kembali.
3. Skala waktu diskrit, satuan pengukuran waktu dalam interval hari, bulan, atau tahun.

Variabel utama dalam analisis ini adalah waktu kejadian yang dihitung dari awal pengamatan hingga terjadi peristiwa tersebut. Salah satu karakteristik unik dari analisis *survival* adalah adanya data yang tersensor (*censored data*), yaitu saat di mana informasi lengkap mengenai waktu hingga peristiwa tidak tersedia, misalnya karena subjek tidak mengalami kejadian tersebut hingga akhir pengamatan atau keluar dari penelitian sebelumnya. Sehingga tidak setiap objek dapat dihitung data tahan hidupnya. Kondisi seperti ini membutuhkan tipe data tersensor, yaitu objek tidak dapat ditentukan data tahan hidupnya. Penyebab data tersensor antara lain:

1. *The study ends*, yaitu objek tidak mengalami kejadian selama penelitian.
2. *Loss to follow up*, yaitu objek menghilang selama masa penelitian.
3. *Withdraws*, yaitu objek mengundurkan diri selama masa penelitian.

Pada analisis *survival* akan selalu mengandung data tersensor (*censored data*) yaitu adanya informasi mengenai waktu ketahanan individu tetapi tidak diketahui secara pasti berapa lama waktu ketahanannya (Kleinbaum & Klein, 2005), sehingga penggunaan data tersensor dilakukan untuk pengamatan yang tidak sampai batas waktu kegagalan (*failure time*) kondisi dimana data yang diperoleh terhadap waktu terjadinya peristiwa tidak lengkap atau tidak munculnya kejadian yang diinginkan selama masa penelitian berakhir. (Maruddin, D., dkk. 2021) menyatakan bahwa Dalam analisis *survival* terdapat tiga tipe data sensing yaitu:

1. Data Lengkap

Setiap objek masuk pada waktu pengamatan yang sama, diamati sampai dengan objek tersebut mengalami kejadian (*event*) yang telah ditentukan. Jika dipunyai n objek:

, x_n , maka masing-masing objek diamati sampai mengalami kejadian (*event*).
Jika dipunyai n data tahan hidup, $t_1, t_2, \dots, t_n$.

Data Tipe 1

dipunyai n objek: $x_1, x_2, \dots, x_n$, dan waktu pengamatan selama T waktu. Maka masa pengamatan, akan diperoleh r data tahan hidup $t_1 \leq t_2 \leq \dots \leq t_r \leq T$.



Pengamatan berakhir pada waktu T dengan $n - r$ data yang tidak mengalami kejadian (*event*).

3. Sensor Data Tipe 2

Jika dipunyai n objek: $x_1, x_2, \dots, x_n$, dan waktu pengamatan berakhir jika telah diperoleh r kejadian (*event*). Maka selama masa pengamatan, akan diperoleh r data tahan hidup $t_1 \leq t_2 \leq \dots \leq t_r$. Pengamatan berakhir setelah ada r kejadian (*event*) dengan $n - r$ data yang tidak mengalami kejadian (*event*).

4. Sensor Data Tipe 3

Jika dipunyai n objek: $x_1, x_2, \dots, x_n$, dengan awal masuk ke dalam pengamatan yang berbeda-beda dan waktu pengamatan berakhir selama waktu T . Selama periode pengamatan tersebut, ada beberapa kemungkinan yang terjadi pada objek penelitian: mengalami kejadian (*event*), tidak mengalami kejadian (*event*), atau keluar/hilang dari pengamatan. Maka selama masa pengamatan, akan diperoleh r data tahan hidup $t_1 \leq t_2 \leq \dots \leq t_r \leq T$. Pengamatan berakhir setelah ada waktu T dengan $n - r$ data yang tidak mengalami kejadian (*event*).

1.4.3 Fungsi *survival* dan Fungsi *hazard*

Distribusi data tahan hidup terdiri dari tiga fungsi, yaitu: fungsi tahan hidup (*survival*), fungsi densitas probabilitas, dan fungsi kegagalan (*hazard*). Jika T adalah variabel random yang menyatakan data tahan hidup, maka distribusi dari T dapat dinyatakan dalam 3 (tiga) buah fungsi yang ekuivalen, sebagai berikut.

Fungsi densitas *probabilitas*. Fungsi densitas *probabilitas* dinotasikan dengan $f(t)$. Fungsi densitas probabilitas didefinisikan sebagai probabilitas suatu objek mengalami kejadian (*event*) dalam interval waktu yang sangat kecil, yaitu antara $(t, t + \Delta t)$ per satuan waktu. Dengan kata lain, Fungsi densitas *probabilitas* menunjukkan seberapa besar probabilitas kegagalan atau peristiwa yang diinginkan terjadi dalam interval waktu kecil tersebut. Fungsi densitas ini dapat dinyatakan sebagai:

$$f(t) = \lim_{\Delta t \rightarrow 0} \left[\frac{P(\text{individu mengalami event dalam interval waktu } (t, t + \Delta t))}{\Delta t} \right]$$

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t}$$

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t}$$

(1)



bertahan hidup seorang individu t , dapat dianggap sebagai nilai yang diukur dari t dengan interval $[0, \infty)$. Variabel acak ini memiliki distribusi probabilitas yang

menggambarkan kemungkinan waktu bertahan hidup. Fungsi distribusi kumulatif $F(t)$ menunjukkan probabilitas bahwa waktu bertahan hidup seseorang kurang dari sama dengan waktu t , dan sering disebut sebagai fungsi distribusi kumulatif, karena menggambarkan probabilitas terjadinya suatu kejadian (*event*) sebelum waktu t . Sehingga secara matematis di tuliskan sebagai berikut:

$$F(t) = P(T \leq t) = \int_0^t f(t) dt$$

Turunan dari fungsi distribusi:

$$F'(t) = \frac{d(\int_0^t f(x)dx)}{dt} = f(t) \quad (2)$$

Fungsi *survival* adalah fungsi yang menggambarkan probabilitas suatu individu atau objek bertahan hidup melewati waktu tertentu (t) tanpa mengalami peristiwa tertentu (seperti kegagalan, kematian, atau kejadian lainnya yang menjadi fokus penelitian). Fungsi ini biasanya dinotasikan sebagai $S(t)$ dan secara matematis didefinisikan sebagai berikut:

$$S(t) = P(T > t) = 1 - P(T \leq t) = 1 - F(t) \quad (3)$$

Sifat fungsi *survival* diatas merupakan Probabilitas bertahan hidup pada waktu(t) yaitu 0 adalah 1, yang berarti individu dalam populasi diasumsikan berada dalam kondisi bertahan pada waktu pengamatan. Sebaliknya seiring berjalannya waktu probabilitas bertahan hingga waktu tak terhingga (∞) adalah 0, peluang individu dapat bertahan semakin kecil hingga akhirnya mencapai nol.

Fungsi *hazard*. Fungsi *hazard* dinotasikan dengan $h(t)$, adalah ukuran risiko terjadinya suatu kejadian (*event*) pada waktu t , dengan syarat individu atau objek telah bertahan hingga waktu tersebut. Fungsi ini didefinisikan sebagai probabilitas suatu objek mengalami kejadian pada interval waktu $(t, t + \Delta t)$, dengan syarat objek tersebut belum mengalami kejadian hingga waktu. Fungsi hazard tidak menyatakan probabilitas langsung, tetapi menunjukkan laju kejadian (*event rate*) pada waktu tertentu per satuan waktu. Fungsi hazard digunakan untuk memahami pola risiko sepanjang waktu (Collett, 1993). Dengan fungsi hazard dapat dinyatakan sebagai berikut:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}$$



rsamaan diatas dapat diperoleh hubungan antara *fungsi survival* dan *fungsi* menggunakan teori probabilitas bersyarat $P(A|B) = \frac{P(A \cap B)}{P(B)}$, maka fungsi sebagai berikut:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t \cdot P(T \geq t)} = \frac{f(t)}{S(t)} \quad (4)$$

Hubungan antara fungsi *hazard* dan *survival*, jika:

$$f(t) = -\frac{dS(t)}{dt} \quad (5)$$

maka:

$$h(t) = \frac{-\frac{dS(t)}{dt}}{S(t)} = -\frac{d}{dt} \ln S(t) \quad (6)$$

Fungsi hazard kumulatif. Dengan mengintegrasikan kedua ruas fungsi persamaan (6) diatas, maka diperoleh persamaan berikut:

$$\int_0^t h(u) du = -\int_0^t \frac{1}{S(u)} \cdot \frac{dS(u)}{du} du = -\ln S(t) + \ln S(0)$$

Karena $S(0) = 1$, maka $\ln S(0) = 0$, sehingga:

$$H(t) = \int_0^t h(u) du = -\ln S(t) \quad (7)$$

Sehingga hubungan antara fungsi kumulatif *hazard* dan fungsi *survival* dapat dinyatakan pada persamaan berikut:

$$S(t) = \exp(-H(t)), \text{ dan } H(t) = -\ln S(t) \quad (8)$$

1.4.4 Model Cox *Proportional Hazard*.

Model Cox Proportional Hazards bertujuan untuk memodelkan hubungan antara waktu survival variabel penjelas (kovariat) yang diduga memengaruhi waktu kejadian. Model ini menyatakan bahwa rasio hazard antara dua individu dengan nilai kovariat yang berbeda terhadap waktu, atau dikenal sebagai asumsi proporsionalitas hazard (Lee & Wang, 2003). Model Cox dapat dinyatakan dalam bentuk:



$$h(t, X) = h_0(t) \exp(\beta^T X)$$

$$h(t, X) = h_0(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)$$

(9)

dengan,

$h(t, X)$	: fungsi <i>hazard</i> pada waktu t yang bergantung pada kovariat X .
$h_0(t)$	: fungsi <i>hazard</i> dasar (<i>baseline hazard</i>).
β_i	: koefisien regresi dari kovariat X_i untuk $i = 1, 2, \dots, p$
X_i	: nilai kovariat ke- i dari individu yang diamati
$\beta^T X$	: kombinasi linear dari parameter regresi (β) dan kovariat (X)
$\exp(\beta)$	: menyatakan perubahan relatif risiko (<i>hazard ratio</i>) untuk setiap perubahan satu satuan pada kovariat X

Dimana $X = (X_1 + X_2 + \dots + X_p)$ adalah vector dari variabel prediktor, dan $h_0(t)$ adalah fungsi *hazard* dasar atau fungsi *baselin hazard*.

Fungsi hazard dasar $h_0(t)$ adalah fungsi *hazard* pada waktu t untuk objek dengan nilai semua variabel (kovariat) bernilai nol, yang dikenal sebagai *baseline hazard*. Model Cox menghitung rasio *hazard* berdasarkan koefisien regresi (β) yang dikalikan dengan nilai kovariat X , dan diterapkan dengan fungsi eksponensial untuk mengukur dampaknya terhadap risiko. Dengan demikian, risiko kegagalan untuk subjek dengan perlakuan *nilai kovariat tertentu* adalah $\exp(\beta)$ kali risiko kegagalan untuk subjek tanpa perlakuan *nilai kovariat nol* (Asuroh, 2005).

Pada persamaan regresi Cox PH, fungsi *hazard* dasarnya tidak diketahui bentuk fungsionalnya. Akan tetapi persamaan Cox PH ini tetap dapat memberikan informasi yang berguna berupa *Hazard Ratio* (HR) yang tidak bergantung dari nilai $h_0(t)$. Namun apabila nilai $h_0(t)$ diketahui, Maka perhitungan nilai *hazard* mengikuti bentuk distribusi data waktunya yang menjadikan persamaan *hazard* merupakan persamaan parametrik. Akan tetapi dalam regresi Cox, *Hazard Ratio* didefinisikan sebagai *Hazard Rate* dari satu individu dibagi dengan *Hazard Rate* dari individu lain. Jika $h_1(t)$ dan $h_0(t)$ adalah *fungsi hazard*, dengan $h_0(t)$ adalah *fungsi baseline hazard*, maka kedua *fungsi hazard* tersebut dikatakan proporsional jika:

$$h(t|X) = h_0(t) \cdot r(X, \beta), \quad \forall t \geq 0$$

(10)

Dalam model Cox, Rasio *hazard* tidak bergantung pada fungsi *hazard* dasar $h_0(t)$. Rasio *hazard* antara dua individu dengan kovariat X_1 dan X_0 adalah:

$$HR(t; X_1, X_0) = \frac{h(t|X_1)}{h(t|X_0)} = \frac{h_0(t) \cdot r(X_1, \beta)}{h_0(t) \cdot r(X_0, \beta)} = \frac{r(X_1, \beta)}{r(X_0, \beta)}$$

(11)



Cox (1972) mengusulkan penggunaan fungsi eksponensial untuk fungsi risiko relatif, sehingga:

$$r(X, \beta) = \exp(X^T \beta) \quad (12)$$

Penggunaan fungsi eksponensial ini untuk memastikan bahwa $h(t|X)$ akan selalu bernilai positif. Dengan demikian, rasio hazard dapat dinyatakan sebagai:

$$HR = \frac{\exp(X_1^T \beta)}{\exp(X_0^T \beta)} = \exp[\beta^T(X_1 - X_0)] \quad (13)$$

Jika X adalah vector kovariat multivariat, maka berlaku:

$$HR = \exp[\beta^T(X_1 - X_0)] \quad (14)$$

Untuk model Cox dengan banyak kovariat, masing-masing kovariat X_i memiliki kontribusi sebagai berikut:

$$HR_i = \exp(\beta_i), \text{ sehingga } \beta_i = \ln(HR_i) \quad (15)$$

Dari persamaan (10), model Cox dengan multivariat dapat ditulis sebagai:

$$h(t|X) = h_0(t) \cdot \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p) \quad (16)$$

Model Cox PH sederhana dapat dinyatakan sebagai

$$\ln(HR) = \beta_1 X$$

Sehingga:

$$\beta_i = \ln(HR) \quad (17)$$

Model Cox PH berganda dapat dinyatakan sebagai:

$$\ln(HR) = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad (18)$$

an, untuk setiap kovariat X_i , berlaku:

$$\beta_i = \ln(HR_i) \quad (19)$$



Yang dimaksud dengan rasio pada HR adalah rasio *hazard* untuk nilai kovariat X terhadap *hazard baseline*. Nilai rasio *hazard* berkisar dari 0 sampai tak hingga. Interpretasi rasio hazard adalah sebagai berikut:

- | | |
|----------|--|
| $HR < 1$ | mengindikasikan bahwa kovariat X merupakan faktor preventif terhadap terjadinya kegagalan. |
| $HR = 1$ | mengindikasikan bahwa tidak ada asosiasi antara kovariat X dengan kejadian kegagalan. |
| $HR > 1$ | mengindikasikan bahwa kovariat X merupakan faktor risiko terhadap terjadinya kegagalan. |

Model *Cox proportional hazard* pada persamaan (9) menyatakan bahwa *hazard* dari suatu individu atau objek pada waktu t dengan diketahui kumpulan variabel independent X yang akan dimodelkan. Model *Cox proportional hazard* merupakan model semiparametric karena tidak memerlukan informasi tentang distribusi yang mendasari waktu *survival* dan fungsi *baseline hazard* adalah fungsi yang tidak ditentukan.

1.4.5 Estimasi parameter pada Model *Cox proportional hazard*

Estimasi parameter pada Model *Cox proportional hazard* menggunakan Maksimum Partial Likelihood Estimation (MPLE), yang dilakukan dengan memaksimalkan fungsi partial likelihood. Apabila terdapat n sampel dan D_i yang merupakan himpunan waktu yang terdiri dari semua waktu survivalnya, sehingga fungsi partial likelihood dari model Cox proportional hazard adalah sebagai persamaan berikut:

$$L(\beta) = \prod_{i=1}^n \frac{\exp(\beta^T x_{(i)})}{\sum_{k \in D_i} \exp(\beta^T x_{(k)})} \quad (20)$$

Dengan memaksimalkan *Likelihood* dalam bentuk fungsi *Log-Likelihood*, Logaritma dari $L(\beta)$ adalah:

$$\ell(\beta) = \sum_{i=1}^n \left(\beta^T x_{(i)} - \log \sum_{k \in D_i} \exp(\beta^T x_{(k)}) \right) \quad (21)$$



na *Gradient* $g(\beta)$, untuk mencari estimasi β dengan menggunakan turunan *g-likelihood* terhadap β :

$$g(\beta) = \frac{\partial \ell(\beta)}{\partial \beta} \quad (22)$$

Suku kedua pada persamaan (24) merupakan ekspektasi kovariat tertimbang dalam kelompok risiko. Dengan aturan turunan, menurunkan setiap suku:

$$\frac{\partial \ell(\beta)}{\partial \beta} = \sum_{i=1}^n \left(x_{(i)} - \frac{\sum_{k \in D_i} x_{(k)} \exp(\beta^T x_{(k)})}{\sum_{k \in D_i} \exp(\beta^T x_{(k)})} \right) \quad (23)$$

Selanjutnya untuk mendapatkan turunan kedua (*Hessian*), *Newton-Raphson* membutuhkan matriks *Hessian* $H(\beta)$ yang merupakan turunan kedua dari log-likelihood:

$$H(\beta) = \frac{\partial^2 \ell(\beta)}{\partial \beta \partial \beta^T} \quad (24)$$

Matriks *Hessian* pada persamaan (26) adalah matriks informasi Fisher negatif, yang digunakan dalam metode iteratif. Dengan turunan kedua diperoleh sebagai berikut:

$$H(\beta) = - \sum_{i=1}^n \left(\frac{\sum_{k \in D_i} x_{(k)} x_{(k)}^T \exp(\beta^T x_{(k)})}{\sum_{k \in D_i} \exp(\beta^T x_{(k)})} - \left(\frac{\sum_{k \in D_i} x_{(k)} \exp(\beta^T x_{(k)})}{\sum_{k \in D_i} \exp(\beta^T x_{(k)})} \right) \left(\frac{\sum_{k \in D_i} x_{(k)} \exp(\beta^T x_{(k)})}{\sum_{k \in D_i} \exp(\beta^T x_{(k)})} \right)^T \right) \quad (25)$$

Pada persamaan (21) hingga (26) diatas merupakan persamaan untuk memaksimalkan turunan $\ln L(\beta)$ karena estimasi parameter yang diperoleh implisit, maka menggunakan metode iterasi numerik yaitu metode *Newton-Raphson* (Collet, 1994). *Newton-Raphson* adalah metode iteratif untuk mencari akar dari turunan pertama (*gradient*). Berikut persamaannya:

$$\hat{\beta}^{(m+1)} = \hat{\beta}^{(m)} - H^{-1}(\hat{\beta}^{(m)}) \cdot g(\hat{\beta}^{(m)}) \quad (26)$$

n $\hat{\beta}^{(0)} = 0,0$ iterasiakan berhenti jika $\|\hat{\beta}^{m+1} - \hat{\beta}^m\| \leq \varepsilon$, dimana ε merupakan yang sangat kecil.



1.4.6 Model regresi Cox *time-dependent covariate*

Model regresi *Cox Proportional Hazard* dapat dimodifikasi untuk melibatkan variabel prediktor bebas terhadap waktu maupun yang bergantung pada waktu (Kleinbaum & Klein, 2005). Sama seperti model *Cox Proportional Hazard* modifikasi ini memuat fungsi *Baseline Hazard* $h_0(t)$ yang dikalikan dengan fungsi eksponensial. *Time-dependent* dilambangkan sebagai $x(t)$, dengan fungsi *hazard* yang bergantung pada waktu $ke - t$, sehingga hubungannya dengan $x(t)$ dapat dinyatakan sebagai berikut:

$$h(t|x(t)) = h_0(t) \cdot \exp\left(\sum_{j=1}^P \beta_j x_j(t)\right) \quad (27)$$

Dengan probabilitas survival dinyatakan dengan:

$$S(t|x(t)) = \exp[-H(t|x(t))]$$

Dimana $H(t|x(t))$ adalah fungsi kumulatif *hazard* yang dijelaskan di bawah. Fungsi kumulatif *hazard* dapat dihitung sebagai integral dari *hazard function*:

$$H(t|x(t)) = \int_0^t h_0(u) \exp(\beta^\top x(u)) du$$

Sehingga dapat ditulis sebagai untuk fungsi hazard:

$$H(t|x(t)) = H_0(t) \cdot \exp(\beta^\top x(t))$$

Nilai hazard kumulatif $x(t)$, yaitu sebagai berikut:

$$\hat{H}(t|x(t)) = \sum_{i=1}^n \int_0^t \frac{\exp(\beta^\top x(u)) d\delta_i(t)}{\sum_{k=1}^n Y_k(u) \exp(\beta^\top x_k(u))} \quad (28)$$

Pada model ini $\delta_i(t)$ merepresentasikan proses *Poisson*, sehingga $\delta_i(t) = I(T_i \geq t)$ dan I sebagai fungsi indikator yang bernilai 1 jika status kejadian (*event*) adalah kegagalan, dan bernilai 0 jika statusnya adalah bertahan hidup (*survive*). Selain itu, $Y_k(t)$ menggambarkan jumlah individu yang masih bertahan (*survive*) pada waktu ke- t (Thomas & Reyes, 2014).
ian, fungsi survival untuk variabel yang bergantung pada waktu t dapat
bagai berikut:

$$\hat{S}(t|x(t)) = \exp[-\hat{H}(t|x(t))] \quad (29)$$



1.4.7 Estimasi parameter pada model regresi Cox *time-dependent covariate*

Dalam model regresi Cox, data yang digunakan mencakup waktu kejadian atau pengamatan terakhir, status kejadian 1 jika peristiwa terjadi, 0 jika tidak, dan kovariat yang mempengaruhi risiko kejadian. Tujuan analisis adalah menentukan nilai $\hat{\beta}$, yang menunjukkan pengaruh kovariat terhadap risiko kejadian. Estimasi ini dilakukan dengan *partial likelihood*, yang memfokuskan pada *rasio hazard* tanpa memerlukan distribusi waktu kejadian secara penuh. *Rasio hazard* ini menggambarkan perbandingan risiko kejadian antara individu dengan kovariat berbeda. Hasil nilai $\hat{\beta}$, diinterpretasikan sebagai *log hazard ratio*, menunjukkan bagaimana perubahan kovariat mempengaruhi risiko kejadian, seperti misalnya peningkatan usia yang dapat meningkatkan risiko kematian. Proses ini dilakukan dengan memaksimalkan *partial likelihood* untuk mendapatkan nilai $\hat{\beta}$, yang memberikan informasi penting tentang faktor-faktor yang mempengaruhi risiko kejadian.

Dalam model Cox *proportional time-dependent covariate*, fungsi *likelihood parsial* mengalami modifikasi karena kovariat $X(t)$ berubah seiring waktu. Model Cox dengan *time-dependent covariates*, model dasarnya tetap sama yaitu:

$$h(t|X(t)) = h_0(t) \exp(\beta^T X(t)) \quad (30)$$

Likelihood Parsial Tetap Sama, Tetapi dengan Kovariat yang Berubah, Struktur *likelihood* tetap mengikuti model Cox yaitu:

$$L(\beta) = \prod_{i=1}^n \left(\frac{\exp(\beta^T x_i(t_i))}{\sum_{j \in R(t_i)} \exp(\beta^T x_j(t_i))} \right)^{\delta_i} \quad (31)$$

Untuk kemudian, digunakan *log-likelihood parsial*, sebagai berikut:

$$\ell(\beta) = \sum_{i=1}^n \delta_i \left(\beta^T x_{(i)}(t_i) - \log \sum_{j \in R(t_i)} \exp(\beta^T x_{(j)}(t_i)) \right) \quad (32)$$



terdapat n peserta dalam kurun waktu penelitian, sehingga data terdiri dari dimana $i = 1, \dots, n$, dimana t_i adalah waktu pengamatan (data pencatatan peserta, i dan δ_i menunjukkan status peserta yang bernilai 1 untuk status *event* atau survive untuk peserta $ke - i$ pada waktu $ke - t$, serta x_i adalah kovariat

dari peserta $ke - i$ pada saat pengamatan $ke - t$. Estimasi β berdasarkan nilai fungsi *partial likelihood* adalah sebagai berikut:

$$L(\beta) = \sum_{i=1}^n \delta_i \left(X_i(t_i) - \frac{\sum_{j \in R(t_i)} X_j(t_i) \exp(\beta^T X_j(t_i))}{\sum_{j \in R(t_i)} \exp(\beta^T X_j(t_i))} \right) \quad (33)$$

Di mana D_i adalah himpunan individu yang memiliki waktu survival lebih panjang dibandingkan dengan waktu pengamatan pada titik $ke - i$. Untuk menghitung nilai $\hat{\beta}$, digunakan *Estimasi maximum partial likelihood*. Diketahui bahwa $\hat{\beta}$ konsisten dan mengikuti distribusi normal *asimtotik* yaitu bersifat kontinu, dengan matriks kovarians $I^{-1}(\hat{\beta})$ (Faisher & Lin, 1999). Persamaan dibawah ini merupakan turunan kedua dari *log-likelihood parsial model Cox Proportional Hazards* dengan *Time-Dependent Covariates*. Ini disebut sebagai matriks informasi Fisher (*Hessian Matrix*) atau *second derivative of log-likelihood*, Maka sebagai berikut:

$$I(\beta) = \sum_{i=1}^n \delta_i \left(\frac{\sum_{j \in R(t_i)} X_j(t_i) X_j(t_i)^T \exp(\beta^T X_j(t_i))}{\sum_{j \in R(t_i)} \exp(\beta^T X_j(t_i))} - \frac{\sum_{j \in R(t_i)} X_j(t_i) \exp(\beta^T X_j(t_i))}{\sum_{j \in R(t_i)} \exp(\beta^T X_j(t_i))} (.)^T \right) \quad (34)$$

1.4.8 AIC (*Akaike Information Criterion*)

Akaike Information Criterion (AIC) adalah suatu kriteria yang dikembangkan oleh Hirotugu Akaike pada tahun 1974 dengan tujuan untuk pemilihan model statistic terbaik dari semua model yang tersedia. Sehingga AIC dapat mengevaluasi model yang meminimalkan kehilangan informasi dengan memberikan prediksi yang akurat. AIC secara umum menggabungkan dua aspek dalam pemilihan model statistic yaitu kecocokan model dengan data dan kompleksitas model dengan kriteria memberikan penalti terhadap model yang menggunakan terlalu banyak parameter agar tidak terjebak dalam overfitting, yaitu saat model terlalu kompleks sehingga dapat mengakomodasi noise pada data dan mengurangi kemampuan prediksi pada data. Secara matematis, AIC dihitung menggunakan rumus:

$$AIC = -2 \log(L(\hat{\theta})) + 2k$$

$$AIC = -2 \sum_{i=1}^n \log(\hat{h}(t_i; \hat{\beta}))$$

(35)



1.4.9 Hazard Ratio

Hazard ratio adalah ukuran yang digunakan untuk melakukan analisis *survival* yang membandingkan risiko terjadinya suatu peristiwa antara dua kelompok pada suatu waktu tertentu. Menurut National cancer institute, *hazard ratio* mengukur seberapa sering suatu peristiwa terjadi dalam satu kelompok dibandingkan dengan kelompok lain selama periode waktu tertentu. *Hazard ratio* sebagai peluang seseorang individu misalkan sembuh lebih cepat dalam pengobatan tertentu, namun tidak memberikan informasi mengenai seberapa cepat peristiwa tersebut terjadi (Spurance et al.2004). Persamaan *Hazard ratio* sebagai berikut:

$$\frac{h(t, X_{j+1})}{h(t, X_j)} = \frac{h_0(t) \exp(\beta X_{j+1})}{h_0(t) \exp(\beta X_j)} = \exp((X_{j+1} - X_j)\beta) \quad (36)$$

Dengan:

X_j = objek dalam variabel independen pada kategori $j + 1$

X_{j+1} = objek dalam variabel independen pada kategori j

Dengan Hubungan antara *hazard ratio* dengan β untuk menginterpretasi laju terjadinya event, sebagai berikut:

1. Jika $\beta_i > 0$, peningkatan nilai x akan meningkatkan nilai *hazard*, sehingga risiko suatu objek mengalami kejadian (*event*) menjadi lebih tinggi.
2. Jika $\beta_i < 0$, peningkatan nilai x akan menurunkan nilai *hazard*, yang berarti risiko suatu objek mengalami kejadian (*event*) akan semakin kecil.
3. Jika $\beta_i = 0$, perubahan nilai x tidak mempengaruhi *hazard*, sehingga risiko suatu objek mengalami kejadian sama dengan risiko objek yang tidak mengalaminya.

1.4.10 Uji Asumsi

Uji asumsi merupakan tahapan penting dalam analisis statistik yang bertujuan untuk memastikan bahwa data atau model yang digunakan memenuhi syarat-syarat tertentu agar hasil analisis valid dan dapat diinterpretasikan secara tepat. Langkah ini dilakukan untuk mengevaluasi apakah data sesuai dengan karakteristik yang diperlukan oleh metode analisis yang dipilih. Beberapa asumsi yang sering diuji mencakup normalitas apakah data mengikuti distribusi normal, homoskedastisitas (*varians residual tetap konstan*), independensi (*data tidak saling berkaitan*), linearitas (*hubungan antara variabel bersifat linear*), serta proporsional (*analisis survival, menguji apakah rasio hazard tetap konstan seiring waktu*). Untuk menguji asumsi kegagalan proporsional, antara lain melalui uji korelasi antara variabel dependen dengan nilai *Schoenfeld Residual* dari masing-masing responden dengan Langkah-langkah sebagai berikut:



Hipotesis

$H_0: \rho = 0$ Kedua variabel tak memiliki korelasi (asumsi *proporsional hazard* terpenuhi)

$H_1: \rho \neq 0$ Kedua variabel memiliki korelasi (asumsi *proporsional hazard* tidak terpenuhi)

Statistik uji

$$Z_j = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (37)$$

dimana,

$\hat{\beta}_j$ adalah koefisien estimasi untuk kovariat Z_j ,

$SE(\hat{\beta}_j)$ adalah standar error dari koefisien estimasi.

Penolakan H_0 terjadi jika $p\text{-value} < \alpha$

Tidak ada penolakan terhadap H_0 jika $p\text{-value} \geq \alpha$

$$p\text{-value} = 2 \cdot (1 - \Phi(|Z_j|)) \quad (38)$$

Jika nilai *residual Schoenfeld* untuk setiap kovariat tidak berkorelasi dengan peringkat waktu *survival* maka asumsi *Proportional Hazard* terpenuhi. Hipotesis nol (H_0) untuk uji asumsi *Proportional Hazard* dengan nilai residual Schoenfeld, yaitu terdapat korelasi antara residual Schoenfeld dengan rank waktu *survival* sama dengan nol (0). Penolakan hipotesis nol (H_0) berarti asumsi *Proportional Hazard* tidak terpenuhi.

1.4.11 Uji Signifikansi

Pengujian signifikansi dilakukan secara serentak dan secara parsial dengan tujuan untuk memastikan model memiliki kemampuan secara keseluruhan dan mengetahui variabel mana yang memberikan pengaruh signifikan, dengan membantu menyederhanakan atau menginterpretasikan model dengan baik. Berikut ini adalah pengujian signifikansi secara serentak dan secara parsial.

1. Uji Serentak (*Uji rasio likelihood*)

Secara serentak hipotesis yang digunakan pada uji signifikansi parameter, yaitu:

H_0 Semua koefisien regresi ($\beta = \beta_1, \beta_2, \dots, \beta_p$) = 0

Minimal ada satu $\beta_j \neq 0$, dengan $j = 1, 2, \dots, p$



H_0 Ini berarti variabel-variabel independen secara bersama-sama tidak memiliki pengaruh yang signifikan terhadap variabel dependen dan H_1 Ini berarti minimal ada 1 variabel prediktor yang mempengaruhi terhadap variabel respon. Berikut

statistik *Uji rasio likelihoodnya*:

$$G^2 = -2 \ln \frac{L(\hat{\omega})}{L(\hat{\Omega})} \quad (39)$$

Kriteria uji Tolak H_0 jika $G_{hit}^2 > X_{(a,p)}^2$ atau $p\text{-value} < \alpha$

$L(\hat{\omega})$: Likelihood dari model kosong (tanpa variabel independen).

$L(\hat{\Omega})$: Likelihood dari model penuh (dengan semua variabel independen).

2. Uji parsial (*Uji wald*)

Secara parsial hipotesis yang digunakan pada uji signifikansi parameter, yaitu:

$H_0 : \beta_j = 0$ (Variabel prediktor tidak berpengaruh terhadap variabel respon)

$H_1 : \beta_j \neq 0$ (Variabel prediktor berpengaruh terhadap variabel respon)

$$W^2 = \frac{(\hat{\beta}_j)^2}{(SE(\hat{\beta}_j))^2} \quad (40)$$

Kriteria keputusan uji Tolak H_0 jika $W_{hit}^2 > X_{(a,1)}^2$ atau $p\text{-value} < \alpha$

Jika variabel prediktor tidak berpengaruh terhadap variabel respon, maka perubahan pada nilai variabel prediktor tidak menyebabkan perubahan yang signifikan pada variabel respon. Sebaliknya, jika variabel prediktor berpengaruh terhadap variabel respon, maka perubahan pada nilai variabel prediktor akan memengaruhi variabel respon secara signifikan (Lee & Wang, 2003).

3. Uji *residuals martingale* dan Uji *residuals deviance*

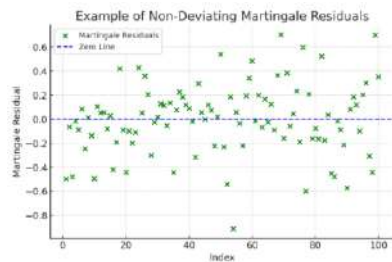
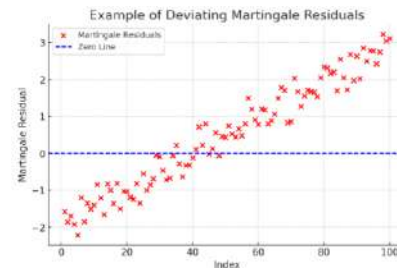
Dalam analisis *survival* menggunakan model *Cox Proportional Hazards*, *residuals* digunakan untuk memeriksa kualitas model dan mengevaluasi kesesuaian data dengan model yang dihasilkan. Dua jenis residual yang umum digunakan adalah residual *Martingale* dan residual *Deviance*, yang keduanya dapat dianalisis melalui plot untuk melihat apakah model sudah fit dengan data.

a. *Residual Martingale*

Residual Martingale adalah jenis *residual* yang digunakan untuk menilai sejauh mana kecocokan model Cox. Residual ini mengukur perbedaan antara nilai yang diharapkan dan nilai yang teramati dari kejadian (misalnya, waktu kegagalan) pada setiap individu. *Residual Martingale* menggambarkan seberapa besar kesalahan model dalam memprediksi waktu kejadian.

Plot Residual Martingale digunakan untuk menggambarkan hubungan antara dual *Martingale* dan kovariat dalam model. Jika model sesuai, plot ini arusnya tidak menunjukkan pola yang jelas. Adanya pola dalam plot iunjukkan kemungkinan masalah dalam model, seperti kovariat yang belum asukkan atau distribusi yang tidak tepat. Berikut ini merupakan contoh plotnya:



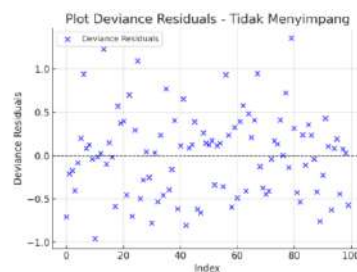
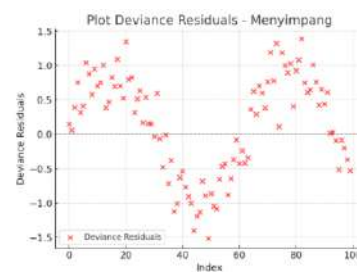
Gambar 1. Plot tidak menyimpang**Gambar 2.** Plot menyimpang

Sumber: OriginLab

b. *Residual Deviance*

Residual Deviance adalah jenis residual yang dihitung berdasarkan *deviance*, yaitu selisih antara *log-likelihood* model dengan *log-likelihood* model yang lebih baik. *Residual* ini digunakan untuk menilai sejauh mana model dapat menjelaskan data yang diamati.

Plot Residual Deviance digunakan untuk memeriksa apakah model telah mencakup semua informasi penting dalam data. Dalam plot ini, nilai *residual Deviance* untuk setiap individu digambarkan terhadap nilai yang diprediksi atau kovariat dalam model. Jika model sesuai, *plot residual deviance* seharusnya tidak menunjukkan pola yang jelas, dan nilai residual akan tersebar secara acak. Namun, jika pola tertentu terlihat, ini mengindikasikan bahwa model belum sepenuhnya sesuai dengan data, dan perbaikan model mungkin diperlukan. Berikut ini merupakan contoh plotnya

Gambar 1. Plot tidak menyimpang**Gambar 2.** Plot menyimpang

Sumber: OriginLab



Pada kedua jenis visualisasi, jika pola yang terlihat acak dan tidak terstruktur, menandakan bahwa model telah diterapkan dengan baik. Sebaliknya, jika terdapat pola yang teratur, outlier ekstrem, atau penyimpangan dari garis horizontal, menunjukkan adanya masalah dalam model yang perlu ditangani. Analisis residual juga dapat membantu dalam mengidentifikasi pengaruh observasi tertentu

yang dapat memberikan dampak signifikan terhadap estimasi parameter model (Therneau & Grambsch, 2000).



BAB II

METODE PENELITIAN

2.1 Pendekatan dan Jenis Penelitian

Pendekatan dan Jenis Penelitian yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif, yang bertujuan untuk mengukur dan menganalisis faktor-faktor risiko gagal bayar iuran pada peserta BPJS Kesehatan di Kabupaten Pangkajene dan Kepulauan. Pendekatan ini dipilih karena fokus penelitian adalah pada pengolahan dan analisis data numerik untuk mengidentifikasi pola, hubungan, dan pengaruh variabel independen terhadap waktu gagal bayar menggunakan model *Cox Proportional Hazard* berbasis *time-dependent*.

2. Jenis Penelitian

Penelitian ini termasuk dalam kategori penelitian analitik observasional, yang bertujuan untuk mengidentifikasi hubungan antara faktor-faktor tertentu (seperti jenis kelamin, usia, pekerjaan, pendidikan, kelas BPJS, segmen dan kepesertaan) dengan risiko gagal bayar iuran BPJS Kesehatan. Data yang digunakan bersifat retrospektif, yaitu data historis peserta yang dikumpulkan dari fasilitas kesehatan di Kabupaten Pangkajene dan Kepulauan. Penelitian ini menggunakan analisis survival untuk mengevaluasi risiko gagal bayar berdasarkan waktu dan variabel yang berubah seiring waktu (*time-dependent covariates*).

2.2 Waktu dan Tempat Penelitian

Penelitian ini akan dilakukan pada bulan Mei tahun 2024 dengan periode data mulai dari bulan Januari 2022 hingga bulan Desember 2023. Tempat penelitian berlokasi di RSUD Batara siang dan fasilitas kesehatan masyarakat (PKM) puskesmas minasatene dan puskesmas bonto perak

2.3 Objek Penelitian

Objek pada penelitian ini adalah data rekam medis pasien rawat jalan dari bulan Januari hingga Desember 2023 di RSUD Batara siang dan fasilitas Kesehatan minasa tene dan bonto perak. Sampel dalam penelitian ini adalah Segmen Pekerja Bukan Penerima Upah atau pekerja mandiri seperti pemilik usaha, petani, sopir angkot, mitra ojol, pedagang, nelayan. dan Segmen Pekerja Penerima Upah seperti Pegawai pemerintah non pegawai negeri sipil (PPNPN) termasuk pegawai ai honorer, staf khusus, dan pegawai lain yang dibayarkan oleh APBN atau yang bekerja di lembaga atau instansi swasta



umber Data

Data penelitian ini merupakan data kuantitatif, yang berupa informasi yang telah ada sebelumnya dan digunakan sebagai pelengkap kebutuhan data penelitian. Data kuantitatif yang digunakan adalah data riwayat pembayaran peserta beserta data demografi peserta BPJS. Data demografi peserta BPJS Kesehatan mencakup informasi seperti usia, jenis kelamin, pendidikan, pekerjaan, kelas BPJS dan segmen untuk menggambarkan profil peserta BPJS.

2.5 Metode Pengumpulan Data

Penelitian ini menggunakan pengumpulan data yang berupa studi literatur. Metode ini melibatkan pengumpulan informasi dari berbagai sumber literatur yang dapat diakses secara online, seperti artikel, jurnal-jurnal dan situs web. Berikut variabel yang digunakan pada penelitian ini:

1. Variabel respon atau dependen : Y
 - a. Time: Mengukur seberapa lama waktu yang dibutuhkan oleh peserta BPJS Kesehatan untuk melakukan pembayaran iurannya. Waktu dihitung sejak awal periode pengamatan hingga akhir periode pengamatan (Januari 2022 hingga Desember 2023).
 - b. Status: Menunjukkan apakah peserta mengalami gagal bayar iuran selama periode pengamatan. Nilainya berupa biner, di mana:
 - 1: Peserta pernah mengalami tunggakan.
 - 0: Peserta tidak mengalami tunggakan selama waktu pengamatan.
2. Variabel prediktor atau independen : X
 - a. Variabel Jenis kelamin X_1
 - 1 = Kategori laki-laki
 - 2 = Kategori perempuan
 - b. Variabel Usia X_2
 - 1 = Kategori Usia 0 -15 tahun
 - 2 = Kategori Usia 16-25 tahun
 - 3 = Kategori Usia 26-35 tahun
 - 4 = Kategori Usia 36-45 tahun
 - 5 = Kategori Usia 46-55 tahun
 - 6 = Kategori Usia 56-65 tahun
 - c. Variabel Pekerjaan X_3
 1. = Kategori peserta Belum/tidak bekerja
 2. = Kategori Bekerja
 - d. Variabel Pendidikan X_4
 1. = Kategori tingkat Pendidikan Tidak/Belum sekolah
 2. = Kategori tingkat Pendidikan Belum Tamat Sekolah dasar/Sederajat
 3. = Kategori tingkat Pendidikan Tamat Sekolah dasar/Sederajat
 4. = Kategori tingkat Pendidikan Sekolah lanjut tingkat pertama/Sederajat
 5. = Kategori tingkat Pendidikan Sekolah lanjut tingkat atas/Sederajat
 6. = Kategori tingkat Pendidikan Diploma III/Diploma IV/Sarjana Muda/Strata-II
 - e. Kepesertaan X_5



Variabel kepesertaan bersifat dinamis, dimana status peserta dapat berubah dari waktu ke waktu, tergantung pada pembayaran iuran setiap bulannya. Jika pada suatu bulan peserta melakukan pembayaran iuran, maka statusnya akan dianggap aktif untuk bulan tersebut. Sebaliknya, jika peserta tidak membayar iuran, maka statusnya menjadi non aktif. Sebagai contoh Jika seorang peserta tidak membayar iuran hingga Juli 2024, maka statusnya pada bulan tersebut adalah non aktif. Namun, jika peserta kemudian melakukan pembayaran pada bulan Juli 2024, maka statusnya berubah menjadi aktif untuk bulan itu.

1. = Aktif (Peserta membayar iuran pada bulan juli 2024)
 2. = Non Aktif (Peserta berhenti membayar iuran pada bulan juli 2024)
- f. Variabel Kelas X_6
1. = Kelas 1
 2. = Kelas 2
 3. = Kelas 3
- g. Variabel Segmen X_7
1. = Pekerja Mandiri
 2. = Pekerja Swasta

2.6 Metode Analisis Data

Penelitian ini dilakukan melalui beberapa tahap analisis untuk mengevaluasi faktor-faktor yang memengaruhi waktu kegagalan pembayaran iuran BPJS Kesehatan. Berikut adalah tahapan metode analisis data:

1. Pengumpulan data peserta rawat jalan BPJS Kesehatan dikumpulkan dari tiga fasilitas kesehatan, yaitu RSUD Batara Siang, Puskesmas Minasa Tene, dan Puskesmas Bonto Perak. Data meliputi informasi demografi, status pembayaran, dan variabel lainnya yang relevan.
2. Analisis deskriptif data demografi peserta BPJS Kesehatan.
3. Pembentukan Model Awal *Cox Proportional Hazard* untuk mengevaluasi pengaruh variabel independen terhadap waktu kegagalan.
4. Uji Asumsi *Proportional Hazards* untuk model awal menggunakan *goodnees of fit*.
5. Mengidentifikasi variabel independen yang tidak memenuhi asumsi *proportional hazards*.
6. Pembentukan Model *Cox time-dependent*, Model *Cox time-dependent* dilakukan melalui tahapan berikut:
 - a. Membangun model *time-dependent* pada variabel kepesertaan.
 - b. Mengestimasi parameter pada model dan menghitung *hazard ratio* untuk mengukur pengaruh variabel terhadap risiko gagal bayar dan uji asumsi model *time-dependent*.
7. Pemilihan model terbaik berdasarkan kriteria AIC terkecil untuk menyederhanakan model



variabel yang paling signifikan mempengaruhi risiko gagal bayar peserta BPJS

rsi Model untuk memastikan validitas hasil dan kekuatan model dalam pengaruh variabel terhadap waktu kegagalan.

2.7 Alur Kerja

