ISSN: 2339-2541
JURNAL GAUSSIAN, Volume 14, Nomor 2, Tahun 2025, Halaman 378 - 389
Online di: https://ejournal3.undip.ac.id/index.php/gaussian/

PERBANDINGAN MODEL *HAZARD* MULTIPLIKATIF DAN ADITIF PADA LAJU PERBAIKAN KONDISI KLINIS PASIEN STROKE DI RS MH THAMRIN CILEUNGSI TAHUN 2021

Zulfa Luthfiyyah Ayunda[1], Triastuti Wuryandari[2], Suparti[3]
[1,2,3] Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro
***e-mail** : zulfayunda17@gmail.com

DOI: 10.14710/j.gauss.14.2.378-389

Article Info:
Received: 2024-07-26
Accepted: 2025-10-10
Available Online: 2025-10-14

Keywords:
Cox Proportional Hazard; Stroke; Survival; Adittive Hazard Regression; Lin-Ying

Abstract: Stroke is a condition that can result in permanent brain damage or even death. In Indonesia, the prevalence of stroke increased from 7% in 2013 to 10.9% in 2018. Numerous factors can affect a stroke patient's ability to recover. Survival analysis is a method that can be used to identify the variables that influence stroke patient's ability to recover. Cox proportional hazard and Lin-ying additive hazard approaches were utilized in this study to analyze stroke patient data from MH Thamrin Cileungsi Hospital. The most widely used regression model for survival data is the Cox proportional hazard which makes the assumption that the ratio between the hazard functions of various people is constant. In contrast, in additive hazard regression there is no assumption of proportionality. Age and cardiac history are the factors that have an impact on how well stroke patients recover, according to the findings. The Lin-Ying additive hazard approach yields the best results since its RMSE value is lower (0.3808777) than that of the cox proportional hazard model (0.9248512).

1. PENDAHULUAN

Stroke adalah jenis penyakit tidak menular yang dapat menyerang siapa pun dan harus segera diatasi karena berpotensi mengancam kesehatan penderita. Angka kejadian stroke di Indonesia telah mengalami peningkatan dari 7% menjadi 10,9% sejak Riskesdas 2013. Prevalensi stroke sebesar 11,4% di Provinsi Jawa Barat dan diperkirakan sekitar 131.846 orang, yang melebihi rata-rata prevalensi stroke di Indonesia sebesar 10,9%, atau diperkirakan sekitar 713.783 orang (Riskesdas, 2018).

Banyak faktor yang dapat mempengaruhi kesembuhan pasien stroke, termasuk usia, jenis kelamin, riwayat merokok, diabetes, hipertensi, dan masalah jantung. Faktor-faktor ini disebut sebagai variabel independen. Sedangkan status pasien dan lama rawat inap adalah variabel dependennya.

Analisis regresi *hazard* adalah jenis analisis *survival* yang sering digunakan dalam studi mengenai pengaruh variabel independen terhadap waktu kejadian. Terdapat dua jenis model *hazard* yang umum yaitu model *hazard* multiplikatif dan aditif (Klein dan Moeschberger, 2003). Model hazard multiplikatif atau yang sering dikenal dengan regresi Cox memiliki persyaratan proporsionalitas yang harus dipenuhi. Sebaliknya, model *hazard* aditif tidak menggunakan asumsi.

Terdapat dua jenis model *hazard* aditif, yaitu model Aalen dan model Lin-Ying. Dalam model Aalen, koefisien regresi memungkinkan menjadi fungsi yang bergantung pada waktu, sementara dalam model Lin-Ying, koefisien regresinya tetap konstan (Klein dan Moeschberger, 2003). Salah satu peneliti yang membahas regresi *hazard* aditif Lin-Ying dan *cox proportional hazard* yaitu Wuryandari, dkk (2020) mengenai analisis *survival* yang menggunakan model regresi *hazard* aditif Lin-Ying pada durasi proses kelahiran.

Tujuan dari penelitian ini adalah untuk mencari apa saja faktor yang mempengaruhi kesembuhan pasien stroke dan menemukan model yang terbaik untuk mengukur tingkat kesembuhan pasien stroke.

2. TINJAUAN PUSTAKA

Menurut *World Health Organization* (WHO), stroke adalah suatu kondisi di mana terdapat gejala klinis yang timbul secara cepat seperti gangguan neurologis fokal atau menyeluruh yang dapat menetap selama 24 jam atau lebih bahkan mengakibatkan kematian. Apabila pembuluh darah di otak tersumbat atau pecah maka dapat terjadi stroke karena hal itu menghilangkan darah dan oksigen di bagian otak yang mengakibatkan kematian sel atau jaringan (Kementerian Kesehatan RI, 2019).

Analisis *survival* merupakan metode yang dipergunakan untuk mengkaji dan menganalisis data yang memiliki hubungan dengan waktu yang dimulai dari *time origin* hingga terjadinya suatu peristiwa istimewa yang disebut dengan *failulre event* (Collett, 2003). Peristiwa khusus (*failure event*) seperti kesembuhan, kekambuhan, maupun peristiwa lainnya. Metode ini juga berguna untuk mengidentifikasi hubungan antara variabel-variabel independen dan dependen, dengan variabel dependennya berfokus pada waktu survival. (Klein dan Kleinbaum, 2012).

Menurut Kleinbaum & Klein (2012) berikut adalah macam-macam penyensoran pada analisis *survival*, yaitu:

1. Penyensoran kanan
2. Penyensoran kiri
3. Penyensoran selang

Dalam analisis *survival* dikenal 3 fungsi yaitu fungsi densitas, *survival* dan *hazard*. Fungsi densitas merupakan peluang seseorang mengalami event dalam rentang waktu t sampai t + Δt dengan rumus sebagai berikut:

$$f(t) = \lim_{\Delta t \to 0} \left[\frac{P(t \leq T < (t + \Delta t)}{\Delta t} \right] \quad (1)$$

Fungsi *survival* merupakan fungsi yang menggambarkan probabilitas individu dapat bertahan hingga atau lebih dari waktu *t* (Cook & Lawless, 2007):

$$S(t) = P(T \geq t) = \int_t^\infty f(x)dx = 1 - P(T \leq t) = 1 - F(t) \quad (2)$$

Fungsi *hazard h(t)* menggambarkan tingkat kejadian suatu individu pada interval waktu dari t hingga t + Δt dengan catatan masih hidup hingga waktu t.

$$h(t) = \lim_{\Delta t \to 0} \left[\frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \right] = \frac{f(t)}{S(t)} \quad (3)$$

Model regresi cox merupakan model yang paling sering digunakan untuk menganalisis data *survival* yang memiliki asumsi proporsionalitas dan menyatakan bahwa fungsi *hazard* dari individu yang berbeda adalah proporsional atau memiliki rasio konstan. Dengan kata lain, laju kejadian risiko dari individu yang berbeda relatif konstan seiring berjalannya waktu (Lee & Wang, 2003). Persamaan regresi cox dapat dinyatakan sebagai berikut:

$$h_i(t|X) = h_0(t)\exp(\boldsymbol{\beta^t x_i}) \quad (4)$$

dengan:

$h_0(t)$: Fungsi dasar *hazard*

$h_i(t|X)$: Fungsi *hazard* individu *ke-i*

$\boldsymbol{\beta^t}$: $(\beta_1, \beta_2, \dots, \beta_p)$

$\boldsymbol{x_i}$: $\begin{pmatrix} x_{i1} \\ x_{i2} \\ \dots \\ x_{ip} \end{pmatrix}$

β_j : Koefisien regresi ke-j

x_{ji} : Nilai variabel independen ke-j dari individu ke-i, dengan *j=1,2,…,p* dan *i=1,2,…n*

Maximum Partial Likelihood Estimation (MPLE) dipakai untuk memperkirakan parameter regresi pada persamaan regresi Cox. Rumus fungsi partial likelihood adalah:

$$L(\beta) = \prod_{i=1}^{r} \frac{\exp(\boldsymbol{\beta^t x_i})}{\sum_{l \in R(t_{(i)})} \exp(\boldsymbol{\beta^t x_l})} \quad (5)$$

dengan:

$\boldsymbol{x_i}$: Vektor variabel yang mengalami kejadian di waktu ke t_i

$\boldsymbol{x_l}$: Vektor variabel yang tidak mengalami kejadian dan elemen dari $R_{(t_{(i)})}$

β : Koefisien regresi

$R_{(t_{(i)})}$: Himpunan yang beresiko mengalami kejadian di waktu ke t_i

Persamaan (5) dapat diubah menjadi bentuk ln, sehingga persamaannya menjadi:

$$\ln L(\beta) = \sum_{i=1}^{r} \left[(\boldsymbol{\beta^t x_i}) - ln\left(\sum_{l \in R(t_{(i)})} \exp(\boldsymbol{\beta^t x_l}) \right) \right] \quad (6)$$

Nilai taksiran β dapat diperoleh dengan mencari nilai maksimum dari fungsi *log partial likelihood* dengan melakukan turunan persamaaan (6) terhadap $\beta = 0$.

Model cox mengandung asumsi yang dikenal sebagai asumsi *proportional hazard*. Untuk menguji kecocokan model digunakan metode taksiran *Goodness of Fit* dengan statistik uji *schoenfeld residuals* (Kleinbaum dan Klein, 2012). Rumus mencari nilai *schoenfeld residuals* sebagai berikut:

$$R_{ji} = \delta_i\{x_{ji} - \hat{a}_{ji}\} \quad (7)$$

dengan,

$$\hat{a}_{ji} = \frac{\sum_{l \in Rt_i} x_{jl} \exp(\widehat{\boldsymbol{\beta^t}} \boldsymbol{x_l})}{\sum_{l \in Rt_i} \exp(\widehat{\boldsymbol{\beta^t}} \boldsymbol{x_l})} \quad (8)$$

Dalam model cox dan Lin-Ying, terdapat pengujian signifikansi parameter menggunakan uji simultan dan parsial untuk menentukan apakah variabel independen berpengaruh pada model.

Uji simultan bertujuan untuk mengetahui apakah parameter β_j secara bersamaan mempengaruhi model. Hipotesisnya adalah:

H_0 : $\beta_1 = \beta_2 = \ldots = \beta_p = 0$ (parameter tidak signifikan)

H_1 : minimal ada satu $\beta_j \neq 0$ (parameter signifikan)

Statistik uji:

$$G = -2\left[lnL_R - lnL_f\right] \tag{9}$$

H_0 ditolak jika $G \geq \chi^2_{(\alpha:db=p)}$ atau *p-value* $\leq \alpha$ pada taraf signifikasi $\alpha = 5\%$ yang berarti parameter signifikan.

Uji parsial bertujuan untuk mengidentifikasi variabel independen yang secara signifikan mempengaruhi variabel dependen. Hipotesisnya adalah sebagai berikut:

H_0: $\beta_j = 0$ (variabel X_j tidak signifikan)

H_1: $\beta_j \neq 0$ (variabel X_j signifikan)

Statistik uji:

$$Z = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \tag{10}$$

H_0 ditolak jika $|Z| \geq Z_{\alpha/2}$ atau *p-value* $\leq \alpha$ pada taraf signifikasi $\alpha = 5\%$ yang berarti variabel signifikan.

Dalam interpretasi model regresi Cox, dapat dihitung *Hazard Ratio*. yang mampu mengindikasi apakah ada peningkatan atau penurunan risiko pada individu yang mengalami kejadian khusus (Lee dan Wang, 2003). Rumus Hazard Ratio:

$$HR = \frac{h(t|x=1)}{h(t|x=0)} = \frac{h_0(t)\exp(1 \text{ x } \beta)}{h_0(t)\exp(0 \text{ x } \beta)} = \exp(\beta) \tag{11}$$

Dalam analisis *survival* sesekali ditemui kejadian sama, yang biasa disebut sebagai *ties*. Keberadaan *ties* dalam data dapat menyulitkan pembentukan *partial likelihood*, terutama ketika menentukan anggota dari himpunan risiko (Xin, 2011). Salah satu metode untuk mengatasi ties adalah *partial likelihood* dengan pendekatan *Efron* yang memiliki perhitungan yang cepat, akurat, dan sederhana meskipun mengandung banyak *ties* (Allison, 2010). Secara umum metode *partial likelihood Efron* memiliki persamaan sebagai berikut:

$$L(\boldsymbol{\beta}_{efron}) = \prod_{i=1}^{D} \frac{\exp(\boldsymbol{\beta}^t \boldsymbol{S}_i)}{\prod_{j=1}^{d_i}\left[\sum_{l \epsilon R_{t_i}} \exp(\boldsymbol{\beta}^t \boldsymbol{X}_l) - \frac{j-1}{d_i}\sum_{k \in D_i}\exp(\boldsymbol{\beta}^t \boldsymbol{X}_i)\right]} \tag{12}$$

dengan,

$L(\boldsymbol{\beta}_{Efron})$: *maximum likelihood estimation* dari parameter $\boldsymbol{\beta}$ menggunakan *Efron*

$\boldsymbol{S_i}$: total nilai variabel independen x pada kasus *ties*

$\boldsymbol{X_l}$: variabel independen yang masih bertahan dan merupakan elemen dari R_t

d_i : banyak kasus kejadian sama (*ties*) pada waktu t_i

D : himpunan indeks i dari semua t_i yang mengalami *event*

Apabila di komponen linear model cox terdapat p variabel $X_1, \ldots, X_p$ dan taksiran koefisien variabel $\beta_1, \ldots, \beta_p$ sehingga dugaan fungsi *hazard* individu ke-i adalah:

$$\hat{h}_i(t) = \exp\left(\hat{\boldsymbol{\beta}}^{\boldsymbol{t}} \boldsymbol{x_i}\right) \hat{h}_0(t) \tag{13}$$

$\hat{h}_0(t)$ adalah taksiran *baseline hazard function, sebagai berikut:*

$$\hat{h}_0(t) = 1 - \xi \tag{14}$$

dengan

$$\xi_j = \left(1 - \frac{\exp(\hat{\boldsymbol{\beta}}^{\boldsymbol{t}} \boldsymbol{x}_{(j)})}{\sum_{l \epsilon R(t_{(j)})} \exp(\hat{\boldsymbol{\beta}}^{\boldsymbol{t}} \boldsymbol{x}_l)}\right)^{\exp(-\hat{\boldsymbol{\beta}}^{\boldsymbol{t}} \boldsymbol{x}_{(j)})} \tag{15}$$

digunakan ketika hanya ada satu individu yang mengalami kejadian di waktu ke-$t_{(i)}$.

$$\xi_j = \exp\left(\frac{-m_j}{\sum_{l \epsilon R(t_{(j)})} \exp(\hat{\boldsymbol{\beta}}^{\boldsymbol{t}} \boldsymbol{x}_l)}\right) \tag{16}$$

digunakan jika terdapat data *ties*.
$\boldsymbol{x}_{(j)}$ adalah vektor dari variabel penjelas untuk objek yang mengalami *event* saat waktu $x_{(j)}$. m_j adalah jumlah individu yang mengalami kejadian di waktu ke-$t_{(j)}$ sedangkan $R_{(t_{(j)})}$ merupakan kumpulan individu yang berisiko mengalami kejadian pada waktu $t_{(j)}$ yang telah diurutkan. Taksiran *baseline survivor function* dapat dihitung dengan

$$\hat{S}_0(t) = \prod_{j=1}^{k} \hat{\xi}_j \tag{17}$$

untuk $t_{(k)} \leq t \leq t_{(k+1)}$, k = 1, 2, 3, … , $r - 1$.
Taksiran nilai dari fungsi survivor dasar adalah nol untuk $t > t_{(r)}$. Taksiran nilai dari fungsi *hazard* dasar kumulatif adalah:

$$\hat{H}_0(t) = -\log \hat{S}_0(t) = -\sum_{j=1}^{k} \log \hat{\xi}_j \tag{18}$$

dengan k $= 1,2,3, \ldots, r$

Alternatif untuk regresi cox adalah regresi *hazard* aditif Lin-ying. Koefisien regresi konstan pada model Lin-Ying memungkinkan taksiran langsung dan interpretasi data yang mudah (Klein dan Moeschberger, 2003). Berikut adalah vektor dari variabel independen $\boldsymbol{X_i(t)}$ pada model Lin-Ying untuk individu ke-i :

$$h(t|\boldsymbol{X_i(t)}) = \alpha_0(t) + \boldsymbol{\alpha^t X_i}(t) \tag{19}$$

dengan:

$\alpha_0(t)$: Fungsi *hazard baseline*
$\boldsymbol{\alpha^t} = (\alpha_1, \ldots, \alpha_p)$: Vektor dari koefisien regresi
$\boldsymbol{X_i(t)} = \begin{pmatrix} X_{i1}(t) \\ \vdots \\ X_{ip}(t) \end{pmatrix}$: Vektor dari variabel independen

Taksiran fungsi hazard dasar melalui teori *counting process* $N_i(t)$ sebagai berikut: (Klein dan Moeschberger, 2003)

$$N_i(t) = M_i(t) + \int_0^t Y_i(t) \left[\alpha_0(t) + \boldsymbol{\alpha^t X_i}(\mathrm{t})\right] dt \tag{20}$$

Taksiran fungsi *hazard* dasar kumulatif diperoleh dari penurunan *counting process* $N_i(t)$ *dan dapat* dicari taksiran fungsi *baseline hazard* kumulatif sebagai berikut:

$$\hat{A}_0 = \int_0^t \frac{\sum_{i=1}^n dN_i(t) - Y_i(t)\boldsymbol{\alpha}^t \boldsymbol{X_i}(\mathrm{t})\, dt}{\sum_{i=1}^n Y_i(t)} \tag{21}$$

Model partial likelihood Cox adalah:

$$L(\boldsymbol{\beta}) = \prod_{i=1}^n \prod_{0 \le t \le \infty} \left(\frac{Y_i(t) exp\{\boldsymbol{\beta}^t \boldsymbol{X_i}(t)\}}{\sum_{i=1}^n Y_i(t) exp\{\boldsymbol{\beta}^t \boldsymbol{X_i}(t)\}} \right)^{dNi(t)} \tag{22}$$

Untuk mendapatkan persamaan skor pada model Cox dapat melakukan turunan logaritma dari $\left(log\ L(\boldsymbol{\beta})\right)$ terhadap $\boldsymbol{\beta}$, diperoleh:

$$U(\boldsymbol{\beta}) = \sum_{i=1}^n \int_0^\infty \boldsymbol{X_i(t)} \left(dN_i(t) - Y_i(t) \exp\{\boldsymbol{\beta}^t \boldsymbol{X_i}(t)\}\, d\hat{A}_0(t, \boldsymbol{\beta}) \right) \tag{23}$$

Score equation model Lin-Ying didapatkan dengan mengganti fungsi $\left(\exp\{\boldsymbol{\beta}^t \boldsymbol{X_i}(t)\}\, d\hat{A}_0(t, \boldsymbol{\beta})\right)$ menjadi $\left(d\hat{A}_0(t) + \boldsymbol{\alpha}^t \boldsymbol{X_i}(t) dt\right)$:

$$U(\boldsymbol{\alpha}) = \sum_{i=1}^n \int_0^\infty \boldsymbol{X_i(t)} \left[dN_i(t) - Y_i(t)\, d\hat{A}_0(t) - Y_i(t) \boldsymbol{\alpha}^t \boldsymbol{X_i}(t) dt \right] \tag{24}$$

Dengan substitusi persamaan (20) ke dalam persamaaan (23) maka diperoleh:

$$U(\boldsymbol{\alpha}) = \sum_{i=1}^n \int_0^\infty [\boldsymbol{X_i(t)} - \overline{\boldsymbol{X}}\boldsymbol{(t)}][dN_i(t) - Y_i(t) \boldsymbol{\alpha}^t \boldsymbol{X_i}(t) dt] \tag{25}$$

Taksiran dari koefisien regresi $\boldsymbol{\alpha}^t$, didapatkan dengan menyelesaikan persamaan $U(\boldsymbol{\alpha}) = 0$ dan didapatkan taksiran koefisien model Lin-ying sebagai berikut:

$$\hat{\alpha} = \left(\sum_{i=1}^n \int_0^\infty Y_i(t) [\boldsymbol{X_i(t)} - \overline{\boldsymbol{X}}\boldsymbol{(t)}]\, [\boldsymbol{X_i(t)} - \overline{\boldsymbol{X}}\boldsymbol{(t)}]\, dt \right)^{-1} \left(\sum_{i=1}^n \int_0^\infty [\boldsymbol{X_i(t)} - \overline{\boldsymbol{X}}\boldsymbol{(t)}] dN_i(t) \right) \tag{26}$$

taksiran dari varian $\hat{\alpha}$ adalah:

$$\widehat{\boldsymbol{V}} = \widehat{V}\,(\hat{\alpha}) = \boldsymbol{A^{-1} C\, A^{-1}} \tag{27}$$

dengan

$$\boldsymbol{A} = \sum_{i=1}^n \sum_{j=1}^i (T_j - T_{j-1}) \left[\boldsymbol{X_i} - \overline{\boldsymbol{X}}(T_j)\right]^{\boldsymbol{t}} \left[\boldsymbol{X_i} - \overline{\boldsymbol{X}}(T_j)\right] \tag{28}$$

$$\boldsymbol{C} = \sum_{i=1}^n \delta_i \left[\boldsymbol{X_i} - \overline{\boldsymbol{X}}(T_j)\right]^{\boldsymbol{T}} \left[\boldsymbol{X_i} - \overline{\boldsymbol{X}}(T_j)\right] \tag{29}$$

Uji simultan bertujuan untuk mengetahui parameter pada setiap variabel indepnden berpengaruh signifikan terhadap variabel dependen. Hipotesisnya adalah sebagai berikut:

H_0 : $\alpha_1 = \alpha_2 = \ldots = \alpha_p = 0$ (parameter tidak signifikan)

H_1 : minimal ada satu $\alpha_j \neq 0$ (parameter signifikan)

Statistik uji :

$$\chi^2 = [\widehat{\boldsymbol{\alpha}}_{\boldsymbol{j}} - 0] \widehat{\boldsymbol{V}}_{\boldsymbol{j}}^{\boldsymbol{-1}} [\widehat{\boldsymbol{\alpha}}_{\boldsymbol{j}} - 0] \tag{30}$$

H_0 ditolak jika $\chi^2 > \chi^2{}_{\alpha;db=p}$ atau *p-value* < α pada taraf signifikasi α = 5% yang berarti minimal satu variabel signifikan.

Uji parsial bertujuan untuk mengetahui variabel independen apa yang berpengaruh secara signifikan. Hipotesisnya adalah sebagai berikut:
H_0: $\alpha_j = 0$ untuk suatu j (variabel X_j tidak signifikan)
H_1: $\alpha_j \neq 0$, untuk suatu j (variabel X_j signifikan)
dengan j=1,2,…,p
Statistik uji :

$$Z = \frac{\hat{\alpha}_j}{\sqrt{\hat{V}_{jj}}} \tag{31}$$

H_0 ditolak jika $|Z| \geq Z_{\alpha/2}$ atau *p-value* $\leq \alpha$ pada taraf signifikasi $\alpha = 5\%$ yang berarti variabel independen berpengaruh signifikan.

Pemilihan model terbaik pada penelitian ini menggunakan RMSE. Model dengan nilai RMSE paling kecil merupakan model terbaik.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \; ; i = 1,2,\ldots,n \tag{32}$$

3. METODE PENELITIAN

Data pada penelitian ini diperoleh dari bagian rekam medis RS MH Thamrin Cileungsi dengan kasus pasien stroke rawat inap pada tahun 2021. Data yang diperoleh berjumlah 93 pasien dengan rincian 90 pasien dinyatakan membaik atau tidak tersensor dan 3 pasien dinyatakan tersensor. Variabel yang dipakai adalah lama pasien rawat inap, status, usia (X_1), jenis kelamin (X_2), jenis stroke (X_3), riwayat jantung (X_4), riwayat diabetes melitus (X_5), riwayat hipertensi (X_6), dan riwayat merokok (X_7).

Berikut adalah tahapan analisis data:

1. Melakukan analisis deskriptif
2. Melakukan Uji Asumsi Proporsional *Hazard*
3. Membentuk model awal cox dan Lin-Ying
4. Melakukan uji parameter secara simultan dan parsial pada model cox dan Lin-Ying
5. Menentukan model akhir cox dan Lin-Ying
6. Menginterpretasikan model cox dan Lin-Ying
7. Menentukan model terbaik menggunakan nilai RMSE yang merujuk pada nilai rata-rata kuadrat kesalahan dalam model

4. HASIL DAN PEMBAHASAN

Pengolahan data menggunakan *Software Rstudio* dan Microsoft Excel. Berikut hasil analisis deskriptif data kontinu dan kategorik disajikan pada Tabel 1 dan Tabel 2.

Tabel 1. Analisis Deskriptif Data Kontinu

Variabel	Mean	Min	Max
Waktu *Survival* (hari)	4,849	1	13

Berdasarkan Tabel 1 didapatkan informasi bahwa rata-rata pasien menjalani rawat inap adalah 4,849 ≈ 5 hari.

Tabel 2. Analisis Deskriptif Data Kategorik

Variabel	Kategori	Jumlah
Usia	0 = < 60 tahun	65
(X_1)	1 = ≥ 60 tahun	28
Jenis Kelamin	0 = Perempuan	44
(X_2)	1 = Laki-laki	49
Jenis Stroke	0 = Hemoragik	15
(X_3)	1 = Iskemik	78
Riwayat Jantung	0 = Tidak Ada	81
(X_4)	1 = Ada	12
Riwayat Diabetes Melitus	0 = Tidak Ada	78
(X_5)	1 = Ada	15
Riwayat Hipertensi	0 = Tidak Ada	50
(X_6)	1 = Ada	43
Riwayat Merokok	0 = Tidak Ada	74
(X_7)	1 = Ada	19

Setelah dilakukan analisis deskriptif, dilanjutkan dengan uji asumsi proportional hazard dengan *goodness of fit* di yang disajikan pada Tabel 3.

Tabel 3. Uji Asumsi Proportional Hazard

Variabel	r_{hitung}	*P-value*	Keputusan
X_1	0,06772839	0,3652414	Gagal tolak H_0
X_2	0,01753667	0,8434465	Gagal tolak H_0
X_3	-0,02865912	0,8281940	Gagal tolak H_0
X_4	-0,09846517	0,2574144	Gagal tolak H_0
X_5	0,09108644	0,3262884	Gagal tolak H_0
X_6	0,11306631	0,1203903	Gagal tolak H_0
X_7	0,05784204	0,7491457	Gagal tolak H_0

Berdasarkan Tabel 3 dapat diketahui H_0 gagal ditolak untuk semua variabel karena nilai $|r_{hitung}| < r_{(91;0,025)} = 0{,}2039$ dan *p-value* > α = 0,05 sehingga kesimpulannya semua variabel memenuhi asumsi *hazard* proporsional. Model awal cox adalah:

$$h(t|X) = h_0(t)\exp(-0{,}9779\,X_1 - 0{,}2701\,X_2 + 0{,}2976\,X_3 - 0{,}9524\,X_4 \\ -0{,}5718\,X_5 - 0{,}3207\,X_6 + 0{,}4732\,X_7)$$

Tabel 4. Uji Cox Proportional Hazard Model Awal

Variabel	β	Exp(β)	Se(β)	Z	P-value	Keputusan
X_1	-0,9779	0,3761	0,2773	-3,526	0,000421	Tolak H_0
X_2	-0,2701	0,7633	0,2238	-1,207	0,227386	Gagal Tolak H_0
X_3	0,2976	1,3466	0,3018	0,986	0,324053	Gagal Tolak H_0
X_4	-0,9524	0,3858	0,3664	-2,600	0,009333	Tolak H_0
X_5	-0,5718	0,5645	0,3092	-1,849	0,064417	Gagal Tolak H_0
X_6	-0,3207	0,7256	0,2270	-1,413	0,157724	Gagal Tolak H_0
X_7	0,4732	1,6052	0,2885	1,641	0,100888	Gagal Tolak H_0

Likelihood ratio test=29,17 on 7 df, p=1e-04

Berdasarkan Tabel 4 disimpulkan minimal ada satu variabel independen yang berpengaruh secara signifikan karena $G\ (29{,}17) \geq \chi^2_{0.05;7}\ (14{,}067)$ atau *p-value* 1e-04 $\leq (\alpha = 0{,}05)$ serta X_2, X_3, X_5, X_6, X_7 gagal tolak H_0 karena $|Z| < Z_{\alpha/2} = 1{,}96$ atau *p-value* > $(\alpha = 0{,}05)$. Sehingga variabel tersebut dieliminasi *backward*.

Tabel 5. Uji Cox Proportional Hazard Model Akhir

Variabel	β	Exp(β)	Se(β)	Z	*P-value*	Keputusan
X_1	-0,9545	0,3850	0,2491	-3,831	0,000127	Tolak H_0
X_4	-0,7274	0,4832	0,3497	-2,080	0,037494	Tolak H_0

Likelihood ratio test=18,92 on 2 df, p=8e-05

Berdasarkan Tabel 5 didapatkan model akhir dan disimpulkan minimal ada satu variabel independen yang berpengaruh secara signifikan karena $G\ (18{,}92) \geq \chi^2_{0.05;2}\ (5{,}991)$ atau *p-value* $8e-05 \leq (\alpha = 0{,}05)$ serta X_1, X_4 menolak H_0 karena $|Z| > Z_{\alpha/2} = 1{,}96$ atau *p-value* $< (\alpha = 0{,}05)$. variabel usia (X_1) dan riwayat jantung (X_4) secara parsial berpengaruh secara signifikan. Didapatkan model akhir sebagai berikut:

$$h(t|X) = h_0(t)\exp\left(-0{,}9545\, X_1 - 0{,}7274\, X_4\right)$$

Tabel 6. Nilai $\hat{\hat{\xi}}_j, \hat{h}_0(t), \hat{S}_0(t), \hat{H}_0(t)$

Time	$\hat{\hat{\xi}}_j$	$\hat{h}_0$	$\hat{S}_0(t)$	$\hat{H}_0(t)$
1	0,9756	0,0243	0,9756	0,0246
2	0,9255	0,0744	0,9030	0,1019
3	0,9625	0,0374	0,8692	0,1401
4	0,8586	0,1413	0,7464	0,2924
5	0,7836	0,2163	0,5849	0,5362
6	0,7693	0,2306	0,4500	0,7985
7	0,6492	0,3507	0,2921	1,2304
8	0,8781	0,1218	0,8781	0,1299
9	0,8534	0,1465	0,7495	0,2883
10	0,6767	0,3232	0,1977	1,6209
11	0,6146	0,3853	0,4606	0,7750
13	0,2735	0,7264	0,126	2,0713

Tabel 6 berisi inferensi statistik meliputi taksiran nilai $\hat{\xi}_j, \hat{h}_0(t), \hat{S}_0(t)$, dan $\widehat{H}_0(t)$ dimana nilai $\hat{h}_0(t)$ yang didapatkan merupakan nilai *baseline hazard function* di persamaan Cox *proportional hazard*.

Interpretasi model cox diperolah hasil bahwa pada variabel usia nilai HR = 0,3850 artinya pasien stroke yang berusia ≥ 60 tahun mempunyai risiko sembuh lebih kecil dibandingkan pasien stroke berusia < 60 tahun sedangkan pada variabel riwayat jantung nilai HR = 0,4832 artinya pasien stroke yang memiliki riwayat jantung mempunyai risiko sembuh lebih kecil dibandingkan pasien yang tidak memiliki riwayat jantung.

Model selanjutnya adalah model Lin-Ying. Model awalnya adalah:

$$\begin{aligned} h(t|X_i(t)) = \alpha_0(t) &- 0{,}17231X_1(t) - 0{,}02994X_2(t) + 0{,}06735X_3(t) \\ &- 0{,}14631X_4(t) - 0{,}09151X_5(t) - 0{,}06315X_6(t) \\ &+ 0{,}06999X_7(t) \end{aligned}$$

Tabel 7. Uji Hazard Aditif Lin-Ying Model Awal

Variabel	*Estimate*	*Std. Error*	Z	*p-value*	Keputusan
X_1	-0,17231	0,04625	-3,725	0,000195	Tolak H_0
X_2	-0,02994	0,04192	-0,714	0,475057	Gagal Tolak H_0
X_3	0,06735	0,05920	1,138	0,255216	Gagal Tolak H_0
X_4	-0,14631	0,05658	-2,586	0,009705	Tolak H_0
X_5	-0,09151	0,05333	-1,716	0,086174	Gagal Tolak H_0
X_6	-0,06315	0,04420	-1,429	0,153062	Gagal Tolak H_0
X_7	0,06999	0,06259	1,118	0,263490	Gagal Tolak H_0

Wald test=23,6476 on 7 df, p=0,001314

Berdasarkan Tabel 7 disimpulkan minimal ada satu variabel independen yang berpengaruh secara signifikan karena χ^2 (23,6474) ≥ ${\chi^2}_{0.05;7}$ (14,067) atau *p-value* sebesar 0,001314 ≤ $(\alpha = 0{,}05)$ serta X_2, X_3, X_5, X_6, X_7 gagal tolak H_0 karena $|Z| < Z_{\alpha/2} = 1{,}96$ atau *p-value* > $(\alpha = 0{,}05)$. Sehingga variabel tersebut dieliminasi *backward*.

Tabel 8. Uji Hazard Aditif Lin-Ying Model Akhir

Variabel	*Estimate*	*Std. Error*	Z	*p-value*	Keputusan
X_1	-0,17311	0,04288	-4,037	5,4e-05	Tolak H_0
X_4	-0,11950	0,05503	-2,172	0,0299	Tolak H_0

Wald test=18,3378 on 2 df, p=0,0001042

Berdasarkan Tabel 8 didapatkan model akhir dan disimpulkan minimal ada satu variabel independen yang berpengaruh signifikan karena χ^2 (18,3378) ≥ ${\chi^2}_{0.05;2}$ (5,991) atau *p-value* 0,0001042 ≤ $(\alpha = 0{,}05)$ serta X_1, X_4 menolak H_0 karena $|Z| > Z_{\alpha/2} = 1{,}96$ atau *p-value* < $(\alpha = 0{,}05)$. variabel usia (X_1) dan riwayat jantung (X_4) secara parsial berpengaruh secara signifikan. Didapatkan model akhir sebagai berikut:

$$h(t|X_i(t)) = \alpha_0(t) - 0{,}17311X_1(t) - 0{,}11950X_4(t)$$

Interpretasi model regresi hazard aditif lin-ying diperolah hasil taksiran koefisien regresi variabel usia sebesar -0,17311. Nilai 0,17311 menunjukkan perbedaan risiko antara pasien berusia ≥ 60 tahun dan berusia < 60 tahun sedangkan hasil taksiran koefisien regresi variabel riwayat jantung sebesar -0,11950. Nilai 0,11950 menunjukkan perbedaan risiko antara pasien yang memiliki riwayat jantung dan tidak memiliki riwayat jantung.

Pemilihan model terbaik dilakukan menggunakan nilai RMSE yang disajikan pada Tabel 9.

Tabel 9. Pemilihan Model Terbaik

Model	Variabel yang Signifikan	RMSE
Cox Proportional Hazard	X_1, X_4	0,9248512
Hazard Aditif Lin-Ying	X_1, X_4	0,3808777

Berdasarkan Tabel 9 dapat diketahui model *Hazard* Aditif Lin-Ying memiliki nilai RMSE sebesar 0,3808777 yang lebih kecil dibanding model regresi *cox proportional hazard* yaitu 0,9248512 sehingga disimpulkan berdasarkan nilai RMSE terkecil maka model terbaik dan yang lebih sesuai yaitu model *Hazard* Aditif Lin-Ying yang terdiri dari dua variabel yaitu usia (X_1) dan riwayat penyakit jantung (X_4).

5. KESIMPULAN

Dari perbandingan kedua model didapatkan model akhir. Model akhir Regresi *Hazard* Aditif Lin-Ying:

$$h(t|X_i(t)) = \alpha_0(t) - 0{,}17311X_1(t) - 0{,}11950X_4(t)$$

Model akhir Regresi Cox Propotional *Hazard*:

$$h(t|X) = h_0(t)\exp\left(-0{,}9545\, X_1 - 0{,}7274\, X_4\right)$$

Variabel-variabel yang mempengaruhi kesembuhan pasien stroke adalah usia (X_1) dan riwayat jantung (X_4). Interpretasi menggunakan Lin-Ying diperoleh nilai $0{,}17311$ menunjukkan perbedaan risiko antara pasien stroke yang berusia ≥ 60 tahun dan < 60 tahun. Sedangkat nilai $0{,}1195$ menunjukkan perbedaan risiko antara pasien stroke yang memiliki riwayat penyakit jantung dan tidak. Sedangkan interpretasi menggunakan cox diperoleh nilai HR = 0,3850 artinya pasien stroke yang berusia ≥ 60 tahun mempunyai risiko sembuh lebih kecil dibandingkan pasien stroke berusia < 60 tahun sedangkan pada variabel riwayat jantung nilai HR = 0,4832 artinya pasien stroke yang memiliki riwayat jantung mempunyai risiko sembuh lebih kecil dibandingkan pasien yang tidak.

Untuk menentukan model terbaik didapatkan model regresi *hazard* aditif Lin-Ying memiliki nilai RMSE sebesar 0,3808777 yang lebih kecil dari regresi cox proportional *hazard* sebesar 0,9248512 dan merupakan model terbaik untuk mengidentifikasi apa saja variabel yang mempengaruhi kesembuhan pasien stroke.

DAFTAR PUSTAKA

Allison, P. D. 2010. *Survival Analysis Using SAS: A Practical Guide, Second Edition*. Cary, NC: SAS Institute Inc.

Collett, D. 2003. *Modelling Survival Data in Medical Research Second Edition*. US: Chapman & Hall.

Cook, R. J., & Lawless, J. F. 2007. *The statistical analysis of recurrent events*. Springer.

Danardono. 2012. *Analisis Data Survival*. Diktat Kuliah. Universitas Gadjah Mada. Yogyakarta.

Kemenkes RI. 2019. *Laporan Nasional RISKESDAS 2018*. Jakarta: Kementrian Kesehatan Republik Indonesia.

Kleinbaum, D. G., & Klein, M. 2012. *Survival Analysis: A Self-Learning Text*. Spinger.

Klein, J.P. and Moeschberger, M.L., 2003. *Analysis Techniques for Censored and Truncated Data Second Edition*, New York.

Lee E.T., dan Wang, J.W. 2003. *Statistical Methods For Survival Data Analysis*. *Third Edition*. New Jersey : John Wiley And Sons.

Xin, Xin. 2011. *A Study Of Ties And Time-Varying covariates In Cox Proportional Hazards Model*. Tesis. Canada: University of Guelph.