Model Cox *Proportional Hazard* Pada Kejadian Bersama *(Ties)* dengan Metode Breslow (Studi Kasus: Pasien Rawat Inap Demam Berdarah *Dengue* (DBD) di Rumah Sakit Dirgahayu Samarinda Periode Juli 2016 s.d Juni 2017)

Cox Proportional Hazard Model At The Tied Incident With Breslow Method (Case Study: Inpatient Dengue Hemorrhagic Fever (DHF) at Dirgahayu Hospital Samarinda July 2016 - June 2017)

Nazmi Soraya[1], Yuki Novia Nasution[2], dan Sri Wahyuningsih[3]
[1,3]Laboratorium Statistika Terapan Jurusan Matematika FMIPA Universitas Mulawarman
[3]Laboratorium Matematika Komputasi Jurusan Matematika FMIPA Universitas Mulawarman
[1]E-mail: nazmisoraya.ns@gmail.com

Abstract

Cox proportional hazard model at the tied incident is a modification from Cox model while there are two or more individual that experienced tied. Parameter estimation in procedure formation of cox generally uses maximum partial likelihood estimation (MPLE) that maximised the function of partial likelihood. In incident with ties case, the modification at partial likelihood is done with breslow approach. The data used in this research were Dengue Hemorrhagic Fever patients (DHF) who were hospitalized at Dirgahayu Hospital Samarinda from July 2016 until June 2017. There are 100 patients with 5 independent variables that are suspected to affect the healing of Dengue Hemorrhagic Fever (DHF) patients, namely sex, age, platelet count, hematocrit count, and duration of fever before hospitalized. From the calculation of data used R 2.11.1 software, Cox proportional hazard model using Breslow method, it is obtained that significant variable, are the count of platelets and hematocrit, and the duration of fever before hospitalized, patient who had normal platelet count had a chance of healing 2.359 times faster than patients had a thrombocytopenia (low platelet). For the amount of heamatocrit each patient who had a normal hematocrit had a chance of healing 2.364 times faster than patients who had under-normal hematocrit, and each one day addition to duration of patient's long history fever before hospitalized, decreased the chance of healing by 0.849 times.

Keywords: *Breslow, Cox Proportional Hazard, DHF, Ties.*

Pendahuluan

Waktu merupakan suatu kejadian atau keadaan yang sering dijumpai dalam kehidupan sehari-hari. Seperti lamanya sembuh dari suatu penyakit, lamanya masa studi seseorang, kegagalan mesin dan lain-lain. Salah satu ilmu statistik yang berkaitan dengan waktu kejadian-kejadian tersebut adalah metode *survival* yang menganalisis data yang berhubungan dengan waktu, mulai dari awal sampai terjadinya suatu peristiwa khusus.

Analisis *survival* merupakan istilah untuk mendeskripsikan analisis data dari *time origin* hingga suatu kejadian terjadi atau *end-point*. Pada penelitian kesehatan, *time origin* dapat merujuk pada pemilihan individu pada suatu penelitian sebagai contoh percobaan untuk membandingkan dua atau lebih perlakuan. Sedangkan, *end-point* adalah waktu pencatatan terakhir berguna untuk mengetahui status tersensor atau tidak tersensor pada suatu penelitian sebagai contoh kejadian meninggalnya pasien. Jika *end-point* berupa kejadian meninggal, sembuh, atau kambuhnya pasien maka data tersebut disebut waktu *survival*. Analisis *survival* digunakan pada data *survival* karena sering terjadi kejadian tersensor. Kejadian tersensor yaitu jika pengamatan waktu *survival* hanya sebagian atau sampai batas periode penelitian (Collett, 2003).

Untuk mengetahui hubungan antara waktu *survival* dengan variabel-variabel yang diduga mempengaruhi waktu *survival* dapat digunakan Model Cox *proportional hazard*. Waktu kejadian atau waktu *survival* dalam analisis *survival* terbagi menjadi dua macam yaitu waktu kejadian tanpa *ties* dan waktu kejadian dengan *ties. Ties* yaitu keadaan dimana terdapat dua individu atau lebih yang mengalami kejadian pada waktu bersamaan. Ada beberapa pendekatan yang dapat digunakan untuk mengatasi terjadinya *ties*, yaitu metode Efron, metode Breslow dan metode *Exact*. Metode Breslow merupakan metode yang lebih sederhana dan perhitungannya lebih cepat (Allison, 2010).

Beberapa penelitian yang mengkaji tentang analisis *survival* pada kejadian bersama telah ditulis oleh beberapa orang, antara lain Vitriana (2016) yang membahas prosedur pembentukan dan penerapan model Cox *extended* untuk mengatasi *nonproportional hazard* pada kejadian bersama dengan pendekatan metode Breslow. Susetyo (2014) yang membahas analisis *survival* data kejadian *ties* dengan membandingan antara metode Efron dan *Exact*.

Berdasarkan beberapa kajian di atas, penulis tertarik untuk mengambil studi kasus dalam bidang kesehatan. Salah satu bidang kesehatan yang akan digunakan adalah Demam Berdarah *Dengue* (DBD). Menurut Dinkes RI 2016, DBD menjadi masalah kesehatan masyarakat Indonesia selama 47 tahun terakhir. Sejak tahun 1968 terjadi peningkatan jumlah Provinsi dan Kabupaten/Kota dari 2 Provinsi dan Kota, menjadi 34 Provinsi dan 435 (85%) Kabupaten/Kota pada tahun 2015. Kalimantan Timur adalah salah satu Provinsi yang masuk ke dalam lima Provinsi dengan angka kesakitan DBD tertinggi tahun 2015 yaitu sebesar 186,12 per 100.000 penduduk.

Peningkatan dan penyebaran DBD tersebut dapat disebabkan oleh faktor epidemiologi. Penyakit ini juga dapat didiagnosis berdasarkan rekam medik yang ada di rumah sakit seperti umur, jenis kelamin, dan pemeriksaan darah seperti hematokrit dan trombosit. Untuk mengetahui faktor-faktor resiko penyakit DBD pada kejadian bersama, maka peneliti akan melakukan pemodelan Cox *proportional hazard* pada kejadian bersama dengan metode Breslow.

Analisis dan Waktu *Survival*

Analisis *survival* merupakan suatu teknik statistik untuk menganalisis variabel acak bernilai positif. Analisis *survival* meliputi variasi metode yang luas untuk menganalisis waktu suatu peristiwa. Pada umumnya analisis ini digunakan pada bidang medis yaitu untuk menganalisis sebuah kematian. Tetapi analisis ini juga dapat digunakan dalam beberapa jenis peristiwa seperti kegagalan mesin, perceraian, dan lamanya masa studi (Miller, 1981).

Menurut Kleinbaum dan Klein (2005), waktu *Survival* adalah catatan waktu yang dicapai suatu objek sampai terjadinya peristiwa tertentu yang disebut sebagai *failure event*. Untuk menentukan waktu *survival* T, secara tepat terdapat tiga elemen yang harus diperhatikan yaitu:

1. Waktu awal (*time origin*)
2. Definisi *failure event* keseluruhan harus jelas.
3. Skala waktu dalam (*time scale*) sebagai satuan pengukuran harus jelas.

Fungsi Kepadatan Peluang

Fungsi kepadatan peluang adalah peluang suatu individu mati atau gagal dalam interval waktu t sampai $t+\Delta t$. Fungsi kepadatan peluang dinotasikan dengan $f(t)$ dan dirumuskan dengan.

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

Misalkan T adalah variabel *random* tidak negatif pada interval $\left[0,\infty\right)$ yang menunjukkan waktu hidup pada suatu populasi dan $f(t)$ merupakan fungsi kepadatan peluang t maka fungsi distribusi kumulatif F(t) adalah (Lawless, 2003):

$$F(t)=P(T\le t)=\int_0^t f(x)\,dx \qquad (2)$$

sehingga dari Persamaan (2) diperoleh

$$f(t)=\frac{d\left(F(t)\right)}{dt}=F'(t) \qquad (3)$$

Fungsi *Survival*

Jika T merupakan variabel *random* tidak negatif pada interval $\left[0,\infty\right)$ yang menunjukkan waktu individu sampai mengalami kejadian pada populasi, $f(t)$ merupakan fungsi kepadatan peluang dari t maka peluang suatu individu tidak mengalami kejadian sampai waktu t dinyatakan dengan fungsi *survival* S(t), (Lawless, 2003):

$$\begin{aligned} S(t)&=P(T\ge t)=1-P(T\le t)\\ &=1-F(t) \end{aligned} \qquad (4)$$

Fungsi *Hazard*

Jika T variabel acak positif pada interval $t=0$ menuju $t=\infty$ yang menunjukkan waktu hidup individu suatu populasi, maka risiko individu tersebut mengalami kegagalan pada interval t hingga $[t,t+\Delta t]$, dan jika diketahui bahwa individu tersebut masih bertahan hingga waktu ke-t adalah :

$$\begin{aligned} h(t)&=\lim_{\Delta t\to 0}\frac{P(t\le T<\left(t+\Delta t\right)|T>t)}{\Delta t}\\ &=\frac{f(t)}{S(t)} \end{aligned} \qquad (5)$$

Hazard Kumulatif

Menurut Lee dan Wang (2003), fungsi kumulatif *hazard* didefinisikan sebagai

$$H(t)=\int_0^t h(x)\,dx \qquad (6)$$

maka hubungan antara fungsi kumulatif *hazard* dengan fungsi *survival* yaitu

$$S(t)=\exp\left[-H(t)\right] \qquad (7)$$

Kurva Fungsi *Survival* dan Metode Kaplan-Meier

Kurva fungsi *survival* Kaplan-Meier membentuk kurva yang menggambarkan hubungan antara estimasi fungsi *survival* dengan waktu *survival* (Lee dan Wang, 2003).

Metode Kaplan-Meier mengelompokkan data ke dalam selang tertentu dan kemudian data disusun dalam suatu tabel. Secara umum susunan tabel Kaplan-Meier sebagai berikut:

1. Kolom waktu (t), merupakan selang waktu yang digunakan.

2. Kolom n_i, merupakan jumlah objek yang dapat bertahan sampai dengan waktu t.
3. Kolom d_i, merupakan jumlah objek yang mengalami kejadian pada dengan waktu t.
4. Kolom $\hat{S}(t)$, merupakan fungsi *survival* di

   mana $\hat{S}(t) = \prod_{t_i < t} \frac{n_i - d_i}{n_i}$, $t_i \le t \le t_{i+1}, i = 1, 2, \ldots \mathrm{m}$.

Kejadian Bersama

Dalam analisis *survival* terkadang ditemukan adanya kejadian bersama atau yang sering disebut *ties*. Kejadian bersama atau *Ties* adalah kejadian dimana terdapat dua individu atau lebih yang mengalami kejadian pada waktu yang bersamaan. Jika suatu data terdapat *ties*, maka akan menimbulkan permasalahan dalam membentuk *partial likelihood* yaitu saat menentukan anggota dari himpunan risikonya (Vitriana , 2016).

Menurut Allison (2010), terdapat tiga metode yang bisa digunakan untuk mengatasi kejadian bersama dalam analisis *survival* yaitu metode Breslow, metode Efron, dan metode Exact. Metode Breslow dan metode Efron keduanya memiliki perhitungan yang sederhana dan cepat, sedangkan metode *Exact* perhitungannya lebih rumit.

Asumsi *Proportional Hazard*

Ada pendekatan umum untuk menaksir asumsi *proportional hazard* yaitu grafik dan Uji *Goodness Of Fit* (GOF).

1. Pendekatan Grafik $\ln\left[-\ln(S(t))\right]$

Membuat *plot* $\ln\left[-\ln(S(t))\right]$ dari fungsi *survival*. Cara ini hanya dapat digunakan untuk variabel kategorik. Jika garis antar kategorik sejajar maka dapat dikatakan asumsi terpenuhi (Collet,2003).

2. Uji *Goodness of Fit*

Pengujian asumsi *proportional hazard* dengan metode *goodness of fit* menggunakan residual Schoenfeld. Residual Schoenfeld terdefinisi pada setiap individu yang mengalami *event* untuk setiap variabel prediktor dalam model (Kleinbaum & Klein, 2005). Adapun langkah-langkah pengujian asumsi *proportional hazard* menggunakan residual Schoenfeld adalah sebagai berikut:

1. Membangun model Cox *proportional hazard* dan mencari taksiran residual Schoenfeld untuk setiap variabel bebas.
2. Membuat variabel *ranking survival time* di mana waktu *survival* diurutkan mulai dari individu yang mengalami *event* pertama kali.
3. Menguji korelasi antara variabel yang dihasilkan pada langkah pertama yaitu residual Schoenfeld dengan variabel yang dihasilkan pada langkah kedua yaitu *ranking survival time*.

Rumus umum residual Schoenfeld adalah sebagai berikut:

$$R_{ij} = \delta_i \left(x_{ij} - \frac{\sum_{l \in R(t_i)} x_{lj} \exp\left(\hat{\boldsymbol{\beta}}'\mathbf{x}_l\right)}{\sum_{l \in R(t_i)} \exp\left(\hat{\boldsymbol{\beta}}'\mathbf{x}_l\right)} \right), j = 1, 2, \ldots, p; i = 1, 2, \ldots, n \tag{8}$$

di mana,

R_{ij} : Residual Schoenfeld variabel bebas ke-*p* dari individu yang mengalami *event* pada waktu $t_{(i)}$

δ_i : Indikator penyensoran

$\hat{\boldsymbol{\beta}}$: Estimator *partial likelihood* maksimum dari β

x_{ij} : Nilai dari variabel bebas ke-*p* dari individu yang mengalami *event* pada waktu $t_{(i)}$

Asumsi *proportional hazard* terpenuhi jika tidak terdapat korelasi antara residual Schoenfeld dengan *ranking survival time.* Berikut adalah rumus korelasi yang digunakan:

$$r = \frac{\sum_{j=1}^{m}\left(R_{ij} - \bar{R}_{ij}\right)\left(RT_i - \overline{RT}_i\right)}{\sqrt{\sum_{j=1}^{m}\left(R_{ij} - \bar{R}_{ij}\right)^2}\sqrt{\sum_{j=1}^{m}\left(RT_i - \overline{RT}_i\right)^2}}$$

di mana,

m : Jumlah Individu ke-*i*

RT_i : *Rangking* waktu *survival* untuk individu ke-*i*

$\overline{RT_i}$:Rata-rata *rangking* waktu *survival* untuk individu ke-*i*

Hipotesis

H_0 : (Asumsi *proportional hazard* terpenuhi)

H_1 : (Asumsi *proportional hazard* tidak terpenuhi)

Statistik Uji

$$z = \frac{r}{\frac{1}{\sqrt{n-1}}} \tag{9}$$

Tolak H_0 jika nilai $|z| > z_{\alpha/2}$ atau *p-value* ≤ α, artinya asumsi *propotional hazard* tidak terpenuhi yang berarti terdapat korelasi antara residual Schoenfeld dengan *ranking survival time.*

Model Umum *Proportional Hazard*

Model Cox *proportional hazard* merupakan model berdistribusi semiparametrik karena model Cox tidak memerlukan informasi tentang distribusi yang mendasari waktu *survival* dan parameter regresi dapat diestimasi dari model Cox tanpa harus menentukan fungsi *baseline hazard* (Lee dan Wang, 2003). Model semiparametrik lebih sering digunakan karena walaupun bentuk fungsional *baseline hazard* tidak diketahui, tetapi model Cox *proportional hazard* dapat memberikan informasi berupa *hazard ratio* yang tidak bergantung pada *baseline hazard*. Melalui model Cox *proportional hazard* dapat dilihat hubungan antara variabel bebas terhadap variabel terikat yaitu waktu *survival* fungsi *hazard*nya. Persamaan model Cox *proportional hazard* dapat ditulis sebagai berikut:

$$h(t,\mathbf{x}) = h_0(t)\exp\left(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p\right) \quad (10)$$

dengan:

$h_0(t)$: *baseline hazard*

$(x_1, x_2, ..., x_p)$: vektor yang berisi p variabel bebas

$(\beta_1, \beta_2, ..., \beta_p)$: vektor pada parameter regresi

Estimasi Parameter Model Regresi Cox *Proportional Hazard*

Estimasi parameter β_j dengan $j = 1,2,...,p$ pada model regresi Cox *proportional hazard* dapat dilakukan salah satunya menggunakan metode *Maximum Partial likelihood Estimation (MPLE.)* Fungsi *likelihood* adalah fungsi dari parameter-parameter β yang tidak diketahui nilainya yang menggambarkan peluang bersama dari observasi data (Kleinbaum and Klein, 2005).

Menurut Collet (2003), misal data untuk m individu yang terdiri dari k waktu kejadian yang tidak tersensor dan $m-k$ individu tersensor kanan, diurutkan menjadi $t_1 < t_2 < \ldots < t_i < \ldots < t_k$ dengan t_i merupakan urutan waktu kejadian ke-i. Diasumsikan hanya terdapat satu individu yang mengalami kematian pada tiap waktu kegagalan, jadi tidak terjadi *ties* pada data. Hal lain, yang perlu dipertimbangkan adalah peluang kematian suatu individu yang mati pada waktu kegagalan t_i, dengan syarat t_i menjadi salah satu yang diamati dari k waktu kegagalan $t_1 < t_2 < \ldots < t_k$.

Misalkan kejadian A adalah individu dengan variabel x_i meninggal pada saat t_i dan kejadian B adalah satu kematian pada saat t_i, maka

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} \quad (11)$$

Pembilang pada Persamaan (12) adalah bentuk sederhana dari risiko kematian pada waktu t_i untuk individu dengan variabel x_i. Jika pembilang tersebut adalah individu ke-i yang meninggal pada saat t_i, fungsi *baseline hazard* ini dapat ditulis menjadi $h_i(t_i)$. Penyebutnya adalah penjumlah dari peluang kematian pada waktu t_i, dari semua individu yang mempunyai risiko kematian pada waktu t_i. Apabila individu-individu tersebut dinotasikan dengan l dan R_{ti} adalah himpunan dari individu-individu yang berisiko pada waktu t_i. Maka peluang dalam Persamaan (11) menjadi $\frac{h_i(t_i)}{\sum_{l\in R(t_i)} h_i(t_i)}$, Dengan mengambil hasil peluang bersyarat di atas, memberikan fungsi *partial likelihood* sebagai berikut:

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{r} \frac{\exp(\sum_{j=1}^{p} \beta_j x_{ij})}{\sum_{l\in R(t_i)} \exp(\sum_{j=1}^{p} \beta_j x_{lj})} \quad (12)$$

Dari Persamaan (12) diperoleh fungsi *ln partial likelihood* sebagai berikut:

$$\ln L(\boldsymbol{\beta}) = \ln\left[\prod_{i=1}^{r} \frac{\exp(\sum_{j=1}^{p} \beta_j x_{ij})}{\sum_{l\in R(t_i)} \exp(\sum_{j=1}^{p} \beta_j x_{lj})}\right]$$

$$= \sum_{i=1}^{r}\left[\left(\sum_{j=1}^{p}\beta_j x_{ij}\right) - \left(\ln \sum_{l\in R(t_i)} \exp\left(\sum_{j=1}^{p}\beta_j x_{lj}\right)\right)\right] \quad (13)$$

Estimasi Parameter Metode Breslow

Menurut Klein dan Moeschberger (2003), fungsi *partial likelihood* dengan metode Breslow sebagai berikut:

$$L(\boldsymbol{\beta})_{Breslow} = \prod_{i=1}^{k} \frac{exp\left(\sum_{l\grave{o}D(t_i)}\sum_{j=1}^{p} x_{lj}\beta_j\right)}{\left(\sum_{l\in R(t_i)} \exp\left(\sum_{j=1}^{p}\beta_j x_{lj}\right)\right)^{d_i}} \quad (14)$$

di mana:

$D(t_i)$: himpunan individu yang mengalami *event* pada waktu t_i

R_{ti} : himpunan individu yang berisiko mengalami *event* pada saat t_i

k : jumlah individu yang tidak tersensor

d_i : banyaknya *ties* pada saat t_i

Dari Persamaan (14) diperoleh fungsi *log partial likelihood* sebagai berikut:

$$\ln L(\boldsymbol{\beta}) = ln\prod_{i=1}^{k} \frac{exp\left(\sum_{l\grave{o}D(t_i)}\sum_{j=1}^{p} x_{lj}\beta_j\right)}{\left(\sum_{l\in R(t_i)} \exp\left(\sum_{j=1}^{p}\beta_j x_{lj}\right)\right)^{d_i}}$$

$$= \sum_{i=1}^{k}\left[\left(\sum_{l\in D(t_i)}\sum_{j=1}^{p}\beta_j x_{lj}\right) - d_i ln\left(\sum_{l\in R(t_i)} exp\left(\sum_{j=1}^{p}\beta_j x_{lj}\right)\right)\right] \quad (15)$$

Pengujian Parameter

Pengujian parameter digunakan untuk menentukan apakah variabel terikat dalam model signifikan terhadap variabel bebas. Pengujian dilakukan secara serentak maupun tiap variabel bebas, uji yang digunakan adalah :

1. Uji *Likelihood Ratio*

Uji *likelihood ratio* digunakan untuk mengetahui pengaruh seluruh kovariat dalam model secara bersama-sama maka diperlukan suatu uji simultan.

Hipotesis dalam pengujian ini adalah:

$H_0 : \beta_1 = \beta_2 = \ldots = \beta_p = 0$

H_1 : minimal ada saat, $\beta_j \neq 0$ dengan $j = 1, 2, \ldots, p$

Statistik uji G adalah sebagai berikut:

$$G = -2\ln\left[\frac{L_0}{L_p}\right] \qquad (16)$$

di mana:

L_0 : Fungsi kemungkinan tanpa peubah penjelas

L_p : Fungsi kemungkinan dengan p peubah penjelas

Uji ini mengikuti sebaran *Chi-Square*, dimana H_0 ditolak jika nilai $G \geq X^2_{(\alpha;p)}$ atau *p-value* $\leq \alpha$ sehingga dapat disimpulkan secara simultan semua variabel bebas berpengaruh terhadap variabel terikat.

2. Uji Wald

Uji Wald digunakan untuk menguji pengaruh estimasi parameter masing-masing variabel bebas maka diperlukan uji parsial.

Hipotesis dalam pengujian ini adalah:

$H_0 : \beta_j = 0$

$H_1 : \beta_j \neq 0,$ dengan $j = 1, 2, \ldots, p$

Statistik uji Wald adalah sebagai berikut:

$$W^2 = \left(\frac{\hat{\beta}_j}{SE\left(\hat{\beta}_j\right)}\right)^2 \qquad (17)$$

di mana:

W^2 : statistik uji Wald

$\hat{\beta}_j$: koefisien variabel bebas ke-j

$SE\left(\hat{\beta}_j\right)$: *standar error* koefisien variabel bebas ke-j

Uji ini mengikuti sebaran *Chi-square* dimana H_0 ditolak jika nilai $W^2 \geq X^2_{(\alpha;p)}$ atau *p-value* $\leq \alpha$ sehingga dapat disimpulkan secara parsial variabel bebas berpengaruh terhadap variabel terikat (Collet, 2003).

Hasil dan Pembahasan

Data berupa data sekunder yang diambil dari hasil rekam medis di Rumah Sakit Dirgahayu Samarinda dengan jumlah penderita DBD sebanyak 100 pasien selama periode Juli 2016 sampai Juni 2017. Variabel terikat yang digunakan pada penelitian ini yaitu variabel waktu (t) menunjukkan lama rawat inap pasien (dalam satuan hari) dan berada dalam batas periode penelitian dengan ketentuan:

$$\text{Status} = \begin{cases} 1, \text{Jika lama rawat inap pasien DBD kurang} \\ \quad \text{dari atau sama dengan 7 hari} \\ 0, \text{Jika lama rawat inap pasien DBD lebih dari} \\ \quad \text{7 hari atau pasien DBD dinyatakan meninggal} \end{cases}$$

Adapun variabel bebas yang digunakan adalah sebagai berikut:

a. Jenis kelamin (x_1) dengan kategori sebagai berikut:

$$x_1 = \begin{cases} 1, \text{ untuk jenis kelamin laki-laki} \\ 0, \text{ untuk jenis kelamin perempuan} \end{cases}$$

b. Umur (x_2)

c. Jumlah Trombosit (x_3), dengan kategori sebagai berikut:

$$x_3 = \begin{cases} 1, \text{ normal (150.00-400.000/mm}^3\text{)} \\ 0, \text{ di bawah normal (<150.000/mm}^3\text{)} \end{cases}$$

d. Jumlah Hematokrit (x_4)

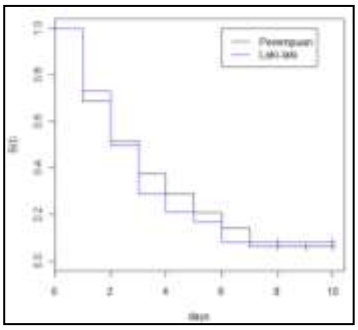
$$x_4 = \begin{cases} 1, \text{ normal} \begin{pmatrix} \text{38\%-42\% untuk perempuan} \\ \text{40-47\% untuk laki-laki} \end{pmatrix} \\ 0, \text{ di bawah normal} \begin{pmatrix} \text{<38\% untuk perempuan} \\ \text{<40\% untuk laki-laki} \end{pmatrix} \end{cases}$$

e. Lama Demam Sebelum Rawat Inap (x_5)

Kurva Fungsi *Survival* Kaplan-Meier dan kurva Fungsi Kumulatif *Hazard*

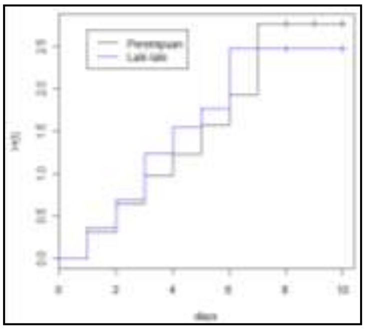
Karakteristik fungsi *survival* dan fungsi kumulatif *hazard* lama rawat inap pasien DBD di Rumah Sakit Dirgahayu Samarinda digambarkan dengan kurva berdasarkan variabel-variabel yang diduga mempengaruhi lama rawat inap pasien DBD. Estimasi fungsi *survival* dan fungsi kumulatif *hazard* dengan metode Kaplan-Meier dilakukan dengan bantuan *software* R.

1. Karakteristik berdasarkan variabel jenis kelamin.



Gambar 1. Kurva fungsi *survival* berdasarkan jenis kelamin

Berdasarkan Gambar 1 kurva fungsi *survival* dimulai dari atas ke bawah hal ini menunjukkan bahwa peluang pasien untuk tetap dirawat di rumah sakit semakin rendah. Dapat dilihat hari ke-1 dan hari ke-7 hingga hari ke-10 peluang pasien DBD yang berjenis kelamin perempuan untuk tetap dirawat di rumah sakit lebih rendah dibandingkan pasien DBD yang berjenis kelamin laki-laki. Sedangkan hari ke-2 hingga hari ke-6 peluang pasien DBD yang berjenis kelamin laki-laki untuk tetap dirawat di rumah sakit lebih rendah dibandingkan pasien DBD yang berjenis kelamin perempuan.



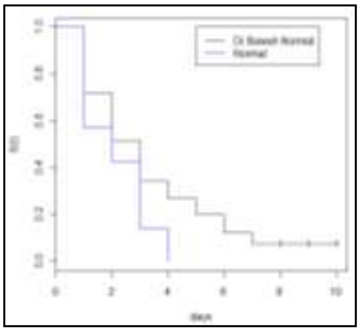
Gambar 2. Kurva fungsi kumulatif *hazard* berdasarkan jenis kelamin

Berdasarkan Gambar 2 dapat dilihat bahwa kurva dimulai dari bawah ke atas hal ini menunjukkan bahwa risiko pasien DBD untuk sembuh semakin tinggi. Dapat dilihat pada hari ke-1 dan hari ke-7 hingga ke-10 risiko pasien DBD yang berjenis kelamin perempuan untuk sembuh lebih tinggi dibandingkan pasien DBD yang berjenis kelamin laki-laki. Sedangkan hari ke-2 hingga hari ke-6 risiko pasien DBD yang berjenis kelamin laki-laki untuk sembuh lebih tinggi dibandingkan pasien DBD yang berjenis kelamin perempuan.

2. Karakteristik berdasarkan variabel jumlah trombosit.

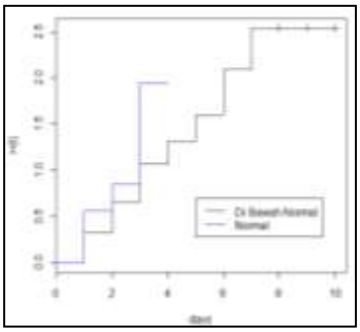
Berdasarkan Gambar 3 dapat dilihat bahwa kurva fungsi *survival* dimulai dari atas ke bawah hal ini menunjukkan bahwa peluang pasien untuk tetap dirawat di rumah sakit semakin rendah. Dapat dilihat pada hari ke-1 hingga hari ke-4 peluang pasien DBD yang memiliki jumlah trombosit normal untuk tetap dirawat di rumah sakit lebih rendah dibandingkan pasien DBD yang memiliki jumlah trombosit di bawah normal.

.



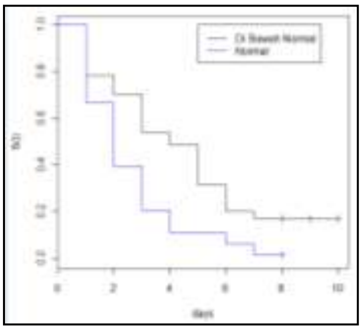
Gambar 3. Kurva fungsi *survival* berdasarkan jumlah trombosit

Berdasarkan Gambar 4 dapat dilihat bahwa kurva fungsi kumulatif *hazard* dimulai dari bawah ke atas hal ini menunjukkan bahwa risiko pasien DBD untuk sembuh semakin tinggi. Dapat dilihat pada hari ke-1 hingga hari ke-4 risiko pasien DBD yang memiliki jumlah trombosit normal untuk sembuh lebih tinggi dibandingkan pasien DBD yang memiliki jumlah trombosit di bawah normal.



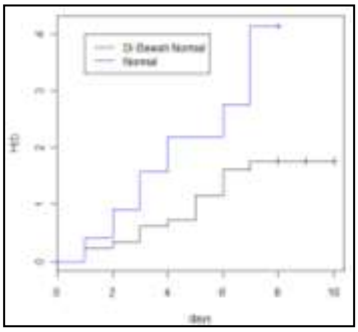
Gambar 4. Kurva fungsi kumulatif *hazard* berdasarkan jumlah trombosit

3. Karakteristik berdasarkan variabel jumlah hematokrit



Gambar 5. Kurva fungsi *survival* berdasarkan jumlah hematokrit

Berdasarkan Gambar 5 dapat dilihat bahwa kurva fungsi *survival* dimulai dari atas ke bawah hal ini menunjukkan bahwa peluang pasien untuk tetap dirawat di rumah sakit semakin rendah. Dapat dilihat pada hari ke-1 hingga hari ke-8 peluang pasien DBD yang memiliki jumlah hematokrit normal untuk tetap dirawat di rumah sakit lebih rendah dibandingkan pasien DBD yang memiliki jumlah hematokrit di bawah normal.



Gambar 6. Kurva fungsi kumulatif *hazard* berdasarkan jumlah hematokrit

Berdasarkan Gambar 6 dapat dilihat bahwa kurva fungsi kumulatif *hazard* dimulai dari bawah ke atas hal ini menunjukkan bahwa risiko pasien

DBD untuk sembuh semakin tinggi. Dapat dilihat pada hari ke-1 hingga hari ke-8 risiko pasien DBD yang memiliki jumlah hematokrit normal untuk sembuh lebih tinggi dibandingkan pasien DBD yang memiliki jumlah hematokrit di bawah normal.

Pengujian Asumsi *Proportional Hazard*

Pengujian asumsi *proportional hazard* dilakukan dengan uji *goodness of fit* menggunakan residual Schoenfeld pada persamaan (8). Adapun hasil uji *goodness of fit* dengan bantuan *software* R sebagai berikut:

Tabel 1. Hasil Pengujian Asumsi *Proportional Hazard* dengan *Goodness of Fit*

Variabel	*p-value*	Keputusan
Jenis Kelamin	0,786	H_0 gagal ditolak
Jumlah Trombosit	0,642	H_0 gagal ditolak
Jumlah Hematokrit	0,729	H_0 gagal ditolak

Hipotesis yang digunakan untuk pengujian asumsi *proportional hazard* dengan *goodness of fit* adalah sebagai berikut:

H_0 : Asumsi *proportional hazard* terpenuhi

H_1 : Asumsi *proportional hazard* tidak terpenuhi

Berdasarkan Tabel 1, diperoleh *p-value* lebih besar dari $a = 0,05$ untuk semua variabel sehingga dari hasil pengujian H_0 gagal ditolak. Dengan tingkat kepercayaan 90% dapat disimpulkan bahwa asumsi *proportional hazard* terpenuhi.

Estimasi Parameter Model Cox dengan Pendekatan Breslow *Partial Likelihood*

Membentuk fungsi *partial likelihood* dengan pendekatan Breslow yang didefinisikan sebagai berikut:

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{k} \frac{exp\left(\sum_{l\grave{o}D(t_i)}\sum_{j=1}^{p} x_{lj}\beta_j\right)}{\left(\sum_{l\in R(t_i)}\exp\left(\sum_{j=1}^{p}\beta_j x_{lj}\right)\right)^{d_i}}$$

Selanjutnya membentuk fungsi *log partial likelihood* dari *L*(**β**) sehingga diperoleh sebagai berikut:

$$\ln L(\boldsymbol{\beta}) = \ln\prod_{i=1}^{92} \frac{exp\left(\sum_{l\grave{o}D(t_i)}\sum_{j=1}^{5} x_{lj}\beta_j\right)}{\left(\sum_{l\in R(t_i)}\exp\left(\sum_{j=1}^{5}\beta_j x_{lj}\right)\right)^{d_i}}$$

$$= \sum_{i=1}^{92}\left[\left(\sum_{l\in D(t_i)}\sum_{j=1}^{5}\beta_j x_{lj}\right) - d_i \ln\left(\sum_{l\in R(t_i)} exp\left(\sum_{j=1}^{5}\beta_j x_{lj}\right)\right)\right]$$

Setelah membentuk fungsi *log partial likelihood*, selanjutnya menentukan turunan pertama dan kedua dari ln *L*(**β**) terhadap β_j yaitu sebagai berikut:

$$\frac{\partial \ln L(\boldsymbol{\beta})}{\partial\beta_j} = \frac{\partial\sum_{i=1}^{92}\left[\left(\sum_{l\in D(t_i)}\sum_{j=1}^{5}\beta_j x_{lj}\right) - d_i ln\left(\sum_{l\in R(t_i)} exp\left(\sum_{j=1}^{5}\beta_j x_{lj}\right)\right)\right]}{\partial\beta_j}$$

$$= \sum_{i=1}^{92}\left[\sum_{l\in D(t_i)}\sum_{j=1}^{5} x_{lj} - d_i\frac{\left(\sum_{l\in R(t_i)}\sum_{j=1}^{5} x_{lj}\exp\left(\sum_{j=1}^{5}\beta_j x_{lj}\right)\right)}{\left(\sum_{l\in R(t_i)}\exp\left(\sum_{j=1}^{5}\beta_j x_{lj}\right)\right)}\right]$$

$$\frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial^2\beta_j} = \frac{\partial^2\sum_{i=1}^{92}\left[\left(\sum_{l\in D(t_i)}\sum_{j=1}^{5}\beta_j x_{lj}\right) - d_i ln\left(\sum_{l\in R(t_i)} exp\left(\sum_{j=1}^{5}\beta_j x_{lj}\right)\right)\right]}{\partial^2\beta_j}$$

$$= \sum_{i=1}^{92} d_i\left(\begin{array}{l}\dfrac{\left(\sum_{l\in R(t_i)} x_{lj}\exp\left(\sum_{j=1}^{5}\beta_j x_{lj}\right)\right)\left(\sum_{l\in R(t_i)}\sum_{j=1}^{5} x_{lj}\exp\left(\sum_{j=1}^{5}\beta_j x_{lj}\right)\right)}{\left(\sum_{l\in R(t_i)}\exp\left(\sum_{j=1}^{5}\beta_j x_{lj}\right)\right)^2} - \\ \dfrac{\left(\sum_{l\in R(t_i)}\sum_{j=1}^{5} x_{lj}^2\exp\left(\sum_{j=1}^{5}\beta_j x_{lj}\right)\right)\left(\sum_{l\in R(t_i)}\exp\left(\sum_{j=1}^{5}\beta_j x_{lj}\right)\right)}{\left(\sum_{l\in R(t_i)}\exp\left(\sum_{j=1}^{5}\beta_j x_{lj}\right)\right)^2}\end{array}\right)$$

Berdasarkan hasil turunan *partial* kedua yang diperoleh dengan pendekatan Breslow *partial likelihood*, dapat dilihat bahwa fungsi tidak dapat diselesaikan secara analitik untuk mendapatkan hasil estimasi parameter. Salah satu metode yang sering digunakan untuk menyelesaikan sistem *non liniear* tersebut adalah iterasi Newton-Raphson, untuk mengestimasi parameter model Cox dilakukan dengan bantuan *software* R.

Analisis Regresi Cox *Proportional Hazard*

Setelah melakukan pengujian asumsi *proportional hazard*, selanjutnya adalah melakukan uji *likelihood ratio* (Secara Simultan) dan uji Wald (Secara Parsial) untuk mengetahui variabel apa saja yang masuk ke dalam model. Hasil estimasi parameter model regresi Cox *proportional hazard* dengan pendekatan Breslow *partial likelihood* dengan bantuan *software* R sebagai berikut:

Parameter yang diperoleh akan diuji secara simultan untuk menguji parameter hasil estimasi antara kelima variabel terhadap waktu *survival*. Adapun hipotesis untuk pengujian secara simultan adalah sebagai berikut:

$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$

H_1 : minimal ada satu $\beta_j \neq 0$, dengan $j = 1,2,3,4,5$

Berdasarkan Tabel 2, diperoleh *p-value* sebesar 0,001 ≤ $\alpha = 0,05$ sehingga dapat disimpulkan bahwa secara simultan minimal ada satu variabel bebas yang berpengaruh terhadap lama rawat inap pasien DBD dirumah Sakit Dirgahayu Samarinda.

Setelah dilakukan pengujian parameter secara simultan, selanjutnya dilakukan pengujian signifikansi parameter secara partial untuk melihat hasil estimasi masing-masing variabel dengan hipotesis sebagai berikut:

$H_0 :\ \beta_j = 0,$ dengan $j = 1,2,3,4,5$

$H_1 :\ \beta_j \neq 0,$ dengan $j = 1,2,3,4,5$

Berdasarkan Tabel 4, diperoleh *p-value* untuk variabel jumlah trombosit, jumlah hematokrit dan lama demam sebelum rawat inap nilainya lebih kecil dibandingkan nilai $\alpha = 0,05$ sehingga dapat disimpulkan bahwa jumlah trombosit, jumlah hematokrit dan lama demam sebelum rawat inap signifikan berpengaruh terhadap lama rawat inap pasien DBD di Rumah Sakit Dirgahayu Samarinda sehingga variabel jumlah trombosit, jumlah hematokrit dan lama demam sebelum rawat ianp dimasukkan ke dalam model.

Tabel 2. Hasil Estimasi Parameter Model Cox Proportional Hazard

Variabel	β	*p-value*	Keputusan
Jenis Kelamin	0,106	0,647	H_0 gagal ditolak
Umur	0,008	0,284	H_0 gagal ditolak
Jumlah Trombosit	0,930	0,037	H_0 ditolak
Jumlah Hematokrit	0,813	0,001	H_0 ditolak
Lama Demam Sebelum Rawat Inap	-0,184	0,008	H_0 ditolak
Likelihood Ratio Test		0,001	H_0 ditolak

Kemudian dilakukan analisis kembali dengan mengeluarkan variabel yang tidak signifikan yaitu variabel jenis kelamin dan umur sebagai berikut:

Tabel 3. Hasil Estimasi Parameter Model Cox Proportional Hazard ke 2

Variabel	β	*p-value*	Keputusan
Jumalah Trombosit	0,858	0,048	H_0 ditolak
Jumlah Hematokrit	0,861	0,000	H_0 ditolak
Lama Demam Sebelum Rawat Inap	-0,164	0,849	H_0 ditolak
Likelihood Ratio Test		0,000	H_0 ditolak

Berdasarkan Tabel 3, diperoleh *p-value* sebesar 0,000 ≤ $\alpha = 0,05$ sehingga dapat disimpulkan bahwa secara simultan minimal ada satu variabel bebas berpengaruh terhadap lama rawat inap pasien DBD di Rumah Sakit Dirgahayu Samarinda. Untuk pengujian secara parsial, berdasarkan Tabel 5, diperoleh *p-value* untuk variabel jumlah trombosit, jumlah hematokrit dan lama demam sebelum rawat inap nilainya lebih kecil dibandingkan nilai $\alpha = 0,05$ sehingga dapat disimpukan bahwa jumlah trombosit, jumlah hematokrit dan lama demam sebelum rawat inap signifikan berpengaruh terhadap lama rawat inap pasien DBD di Rumah Sakit Dirgahayu Samarinda. Sehingga diperoleh model regresi Cox *proportional hazard* sebagai berikut:

$$\hat{h}(t,\boldsymbol{x}) = \hat{h}_0(t)\exp(0,858x_3 + 0,861x_4 - 0,164x_5)$$

Berdasarkan model yang diperoleh diketahui bahwa nilai *hazard ratio* untuk variabel jumlah trombosit adalah $\exp(0,858)$ atau 2,359. Artinya pasien DBD yang jumlah trombositnya normal memiliki kesempatan 2,359 kali lebih cepat untuk mengalami *event* atau sembuh dibandingkan pasien DBD dengan jumlah trombosit di bawah normal. Nilai *hazard ratio* untuk variabel jumlah hematokrit adalah $\exp(0,861)$ atau 2,364. Artinya pasien DBD yang memiliki jumlah hematokrit normal memiliki kesempatan 2,364 kali lebih cepat untuk mengalami *event* atau sembuh dibandingkan pasien DBD yang memiliki jumlah hematokrit di bawah normal. Dan nilai *hazard ratio* untuk variabel lama sakit sebesar $\exp(1(-0,164))$ atau 0,849. Artinya setiap penambahan satu hari riwayat lama demam pasien DBD sebelum rawat inap, kesempatan untuk mengalami *event* atau sembuh adalah 0,849 kali. Dengan kata lain semakin lama riwayat pasien demam sebelum dirawat inap di rumah sakit maka kesempatan untuk sembuh semakin kecil.

Kesimpulan

Berdasarkan hasil analisis dan pembahasan maka diperoleh kesimpulan bahwa, setelah dilakukan estimasi parameter menggunakan metode Breslow diperoleh bahwa variabel jumlah trombosit, jumlah hematokrit, dan lama sakit secara parsial dan simultan berpengaruh. Sehingga diperoleh model Cox dengan variabel yang berpengaruh signifikan sebagai berikut:

$$\hat{h}(t,\boldsymbol{x}) = \hat{h}_0(t)\exp(0,858x_3 + 0,861x_4 - 0,164x_5)$$

Berdasarkan model yang telah didapatkan, dapat diperoleh bahwa nilai *hazard ratio* untuk variabel jumlah trombosit adalah $\exp(0,858)$ atau 2,359. Artinya pasien DBD yang jumlah trombositnya normal memiliki kesempatan 2,359 kali lebih cepat untuk mengalami *event* atau sembuh dibandingkan pasien DBD dengan jumlah trombosit di bawah normal. Nilai *hazard ratio* untuk variabel jumlah hematokrit adalah $\exp(0,861)$ atau 2,364. Artinya pasien DBD yang memiliki jumlah hematokrit normal memiliki kesempatan 2,364 kali lebih cepat untuk mengalami *event* atau sembuh dibandingkan pasien DBD yang memiliki jumlah hematokrit di bawah normal. Dan nilai *hazard ratio* untuk variabel lama sakit sebesar $\exp(1(-0,164))$ atau 0,849. Artinya setiap penambahan satu hari riwayat lama sakit pasien DBD sebelum di rawat inap, risiko untuk mengalami *event* atau sembuh adalah 0,849 kali. Dengan kata lain semakin lama riwayat pasien sakit sebelum dirawat inap di rumah sakit maka kesempatan untuk sembuh semakin kecil.

Daftar Pustaka

Allison, P. (2010). *Survival Analysis Using SAS: A Practical Guide*. Cary: SAS Institute Inc.

Collett, D. (2003). *Modelling Survival Data in Medical Research*. London: Chapman & Hall/CRC.

Depkes RI. (2016). *Info DATIN Pusat Data dan Informasi Kementrian Kesehatan RI Situasi DBD*. Jakarta: Depkes RI. Dari www.dipkes.go.id diaskes pada 27 Febuari 2017 pukul 12.36 WITA.

Kleinbaum, D. G., dan Klein, M. (2005). *Survival Analysis: A Self-Learning*. New York: Springer.

Klein, J. P., dan Moeschberger, M. L. (2003). *Survival Analysis: Techiques for Censored and Truncated Data* (Second Edition). New York: Springer-Verlag.

Lee, E. T., dan Wang, J. W. (2003). *Statistical Methods for Survival Data Analysis* (Third Edition). Hoboken: John Wiley & Sons.

Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data* (Second Edition). Hoboken: John Wiley & Sons, Inc.

Miller, Jr.,RG. (1981). *Survival Analysis*. New York: John Weley & Sons.

Susetyo, A. B. (2014). Analisis *Survival* Data Kejadian *Ties* dengan *Exact Partial Likelihood* Cox *Regression* (Studi Kasus: Data Siswa Putus Sekolah Tingkat Menengah Pertama). *Seminar Nasional Matematika dan Pendidikan Matematika Universitas Jember, Jember.*

Vitriana, A. N., Kusumawati, R. (2016). Model Cox *Extended* dengan g(t)=1 untuk Mengatasi *Nonproportional Hazard* pada Kejadian Bersama. *Prosiding Seminar Nasional Matematika dan Pendidikan Matematika Universitas Negeri Yogyakarta*, Yogyakarta.