# Model Regresi Cox Weibull Dengan Metode Penaksiran Parameter Efron *Partial Likelihood* (Studi Kasus : Lama Perawatan Pasien Penderita Tuberkulosis Di Puskesmas Loa Ipuh Tenggarong Tahun 2017)

## *Cox PH Weibull Regression Model with Efron Partial Likelihood Estimation Method (Case Study: Treatment Duration of Tuberculosis Patients in Health Services of Loa Ipuh Tenggarong 2017)*

**Ihsan Akbar, Suyitno, dan Sri Wahyuningsih**
Laboratorium Statistika Terapan Universitas Mulawarman
E-mail:ihsanakbar99@gmail.com

***Abstract***

*Survival analysis is an analysis that involves statistical tests to analyze data on the time or length of time until the occurrence of a particular event. One regression model for time duration data modeling is the Cox PH regression model. Cox PH regression applied to time duration data with Weibull distribution is called Cox PH Weibull regression. The purpose of this study is to obtain a Cox PH Weibull regression model and determine the factors that influence the length of treatment for tuberculosis patients. The parameter estimation in this study is Efron partial likelihood method. The Efron partial likelihood method is suitable for estimating Cox PH regression parameters to data containing ties. Based on the results of parameter estimation the best model is obtained by using AIC crtiteria. Based on the partial test, age is factor that influence to the length of treatment. The results of the study show that every increasing age patient one year, then length of treatment until the patient recovered, will increas 1,024 times.*
*Keywords: Cox PH Weibull, Efron Partial Likelihood, ties*

## Pendahuluan

Analisis *survival* atau yang biasa dikenal sebagai analisis ketahanan hidup merupakan prosedur statistika yang digunakan untuk menganalisis data waktu antar *event*. Analisis ini digunakan pada data durasi waktu hingga terjadi peristiwa tertentu (Kleinbaum dan Klein, 2005). Variabel yang menyatakan jangka waktu dari awal pengamatan dimulai hingga *event* yang diharapkan terjadi, disebut dengan waktu *survival* atau *failure time*. Disiplin ilmu *survival* ini diterapkan tidak hanya terbatas pada *event* yang dialami oleh makhluk hidup saja, melainkan dapat pula diterapkan pada keadaan suatu benda sampai benda tersebut mengalami *event* tertentu, seperti kerusakan dari suatu alat atau benda tertentu.

Menurut Collett (2003), analisis ketahanan hidup menggambarkan analisis data waktu tahan hidup dari awal waktu penelitian sampai *event* tertentu terjadi. Salah satu metode analisis ketahanan hidup adalah regresi Cox. Regresi Cox merupakan salah satu metode statistika yang digunakan untuk mengetahui hubungan antara variabel dependen dengan variabel independen. Regresi Cox pertama kali dikembangkan oleh Cox pada tahun 1972. Regresi ini lebih populer digunakan dalam penelitian tentang data kesehatan, data ekonomi, yang variabel responnya berupa waktu (hari, bulan, tahun).

Menurut Allison (2010), dalam analisis *survival* kadang-kadang ditemukan adanya kejadian bersama atau yang lebih sering disebut *ties*. *Ties* adalah keadaan dimana terdapat dua individu atau lebih yang mengalami *event* pada waktu yang bersamaan. Jika suatu data terdapat *ties*, maka akan menimbulkan permasalahan dalam membentuk *partial likelihood* yaitu saat menetukkan anggota dari himpunan risikonya. Terdapat tiga metode yang biasa digunakan untuk mengatasi kejadian bersama dalam analisis *survival* yaitu metode *Exact*, pendekatan Breslow, dan pendekatan Efron.

Aplikasi dari distribusi Weibull dan kerabatnya sangat luas dan mencakup hampir semua disiplin ilmu. Sebagian besar aplikasi dari distribusi Weibull berkaitan dengan data kehidupan. Proses penarikan kesimpulan untuk data kehidupan sangat bergantung pada bagaimana data yang telah diambil dan diolah sedemikian rupa (Rinne, 2009).

Penelitian sebelumnya yang terkait dengan model regresi Cox Weibull dilakukan oleh Bastyan dan Latra (2013) dengan studi kasus nya adalah penderita Demam Berdarah Dengue (DBD) di Rumah Sakit Haji Sukolilo Surabaya. Penelitian lain dilakukan oleh Prabawati (2017) yang membahas analisis *survival* pada kejadian bersama dengan pendekatan Efron *Partial Likelihood* yang diterapkan pada lama waktu studi yang dilaksanakan oleh mahasiswa MIPA sampai waktu kelulusan.

Menurut Subuh dan Priohutomo (2014), tuberkulosis sampai saat ini masih merupakan salah satu masalah kesehatan masyarakat di dunia. Upaya pencegahan dan pengendaliannya telah diterapkan di banyak negara sejak tahun 1995, akan tetapi angka kematian akibat penyakit ini masih saja tinggi di beberapa negara. Seperti yang dimuat dalam CNN Indonesia (2018), di Indonesia sendiri, masih tercatat sebagai salah satu dari negara dengan beban Tuberkulosis yang tinggi. WHO Global TB Report 2017 memperkirakan jumlah kasus Tuberkulosis sebanyak 1.020.000 kasus. Sedangkan di Kabupaten Kutai Kartanegara sendiri, Tuberkulosis masuk dalam 20 besar penyakit yang diderita oleh penduduk Kukar, sebanyak 274 orang yang dilaporkan menderita penyakit Tuberkulosis pada tahun 2016 saja (Bps Kukar, 2017).

Tujuan penelitian ini adalah untuk mengetahui distribusi dari data lama pengobatan pasien penyakit tuberkulosis telah berdistibusi Weibull. Selain itu, penelitian ini juga bertujuan untuk menentukan model regresi Cox dan mengetahui faktor-faktor yang mempengaruhinya.

**Fungsi-fungsi Waktu *Survival***

Waktu survival T merupakan variabel random kontinu non-negatif. Misalkan variabel random T menunjukan durasi waktu sampai terjadinya suatu *event*. Waktu survival T merupakan variabel random kontinu dan non-negatif dalam interval $[0,\infty]$ (Lawles, 2003). Fungsi-fungsi yang saling berhubungan dalam analisis *survival* adalah, fungsi kepadatan peluang, fungsi *survival*, fungsi *hazard*.

Fungsi kepadatan peluang (FKP) adalah peluang suatu individu mengalami *event* dalam interval waktu dari $t$ sampai $t+\Delta t$, dengan waktu T merupakan variabel random. Fungsi kepadatan peluang dinotasikan $f(t)$ dan dirumuskan dengan

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

maka fungsi distribusi kumulatif $F(t)$ adalah

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

Menurut Lee dan Wang (2003) fungsi *survival* $S(t)$ didefinisikan sebagai peluang suatu individu *survive* (belum mengalami event) lebih lama daripada *t* dapat dinyatakan sebagai berikut

$$S(t)=1-F(t) \qquad (3)$$

Menurut Lawless (2003) fungsi hazard adalah peluang suatu individu mengalami *event* dalam waktu interval t sampai $t+\Delta t$, jika diketahui individu tersebut masih dapat bertahan hidup hingga waktu t, yang dinyatakan sebagai berikut :

$$\mathrm{h}(t)=\lim_{\Delta t\to 0}\frac{P(t\le T<(t+\Delta t)\mid T>t)}{\Delta t} \qquad (4)$$

Hubungan fungsi *survival* dan fungsi *hazard* dinyatakan dalam persamaan sebagai berikut

$$\mathrm{h}(t)=\frac{f(\mathrm{t})}{S(t)} \qquad (5)$$

**Distribusi Weibull**

Distribusi Weibull, diperkenalkan oleh Wallodi Weibull (1939), adalah salah satu distribusi yang paling populer di displin ilmu statistika modern. Distribusi Weibull menjadi populer, dikarenakan distribusi ini dapat diterapkan di berbagai macam bidang, mulai dari data waktu sampai dengan data cuaca atau bahkan data pengamatan dalam bidang ekonomi dan administrasi (Rinne, 2008). Fungsi peubah waktu T distribusi Weibull versi skala bentuk adalah

$$f(t)=\frac{\gamma t^{\gamma-1}}{\lambda^{\gamma}}\exp\left[-\left(\frac{t}{\lambda}\right)^{\gamma}\right] \qquad (6)$$

dimana $t>0,\lambda>0,\gamma>0$.

Fungsi *survival* dari distribusi Weibull dengan FKP yang diberikan oleh Persamaan (6) adalah

$$S(t)=-\int f(t)dt == \exp\left[-\left(\frac{t}{\lambda}\right)^{\gamma}\right] \qquad (7)$$

fungsi hazard dari distribusi Weibull Persamaan (6) adalah

$$h(t)=\frac{f(t)}{S(t)}=\frac{\gamma}{\lambda}\left(\frac{t}{\lambda}\right)^{\gamma\text{-}1} \qquad (8)$$

**Penaksiran Parameter Distribusi Weibull**

Misalkan $T_1,T_2,...T_n$ adalah sampel random berukuran *n* yang berasal dari distribusi Weibull dengan fungsi densitas pada Persamaan (7). Berdasarkan sampel random $T_1,T_2,...T_n$ ditaksir parameter dengan menggunakan metode *maximum likelihood Estimator* (MLE). Misalkan vektor $\mathbf{t}=\begin{pmatrix}t_1 & t_2 & \dots & t_n\end{pmatrix}^T$ hasil pengamatan dan diasumsikan sampel random $T_1,T_2,...T_n$ adalah saling bebas dan berdistribusi identik Weibull, maka fungsi *likelihood* adalah

$$L(\boldsymbol{\theta},\mathbf{t})=L(t;\lambda,\gamma)=\prod_{i=1}^{n} f(t_i;\lambda,\gamma)$$

$$=\prod_{i=1}^{n}\left(\frac{\gamma}{\lambda^{\gamma}}\right)t_i^{\gamma-1}\exp\left[-\left(\frac{t_i}{\lambda}\right)^{\gamma}\right] \qquad (9)$$

dimana $\boldsymbol{\theta}^T=[\alpha \quad \gamma]$. Fungsi *log likelihood* dari fungsi *likelihood* (9) adalah

$$\ln L(\boldsymbol{\theta},\mathbf{t})=\sum_{i=1}^{n}\ln\left[\left(\frac{\gamma}{\lambda^{\gamma}}\right)t_i^{\gamma-1}\exp\left[-\left(\frac{t_i}{\lambda}\right)^{\gamma}\right]\right] \qquad (10)$$

(Suyitno, 2017)

**Pengujian Distribusi Weibull**

Menurut Corder dan Foreman (2014), uji distribusi data dapat dilakukan dengan uji Kolmogorov-Smirnov sebagai berikut. Misalkan ingin diuji apakah suatu data terdistribusi Weibull dengan fungsi distribusi $F^*(x)$. Hipotesis yang digunakan adalah

$H_0$: $F(x) = F^*(x)$ (Data berdistribusi Weibull)

$H_1$: $F(x) \neq F^*(x)$ (Data tidak berdistribusi Weibull)

Statistik uji yang digunakan adalah

$$D = \max | F^*(x) - A(x) | \qquad (11)$$

Kriteria keputusan:

Daerah penolakan $H_0$ adalah jika nilai $D_{hitung} > D_{tabel}$ yang berarti bahwa data tidak mengikuti distribusi Weibull.

**Estimasi Fungsi *Survival* Kaplan-Meier**

Estimasi Kaplan-Meier disebut juga estimasi *product limit*. Estimasi Fungsi *survival* ini pertama kali dikemukakan oleh Kaplan dan Meier pada tahun 1958. Misalkan *T* variabel random non negatif pada interval $[t_i, t_{i+1})$. Estimasi fungsi *survival* didefinisikan sebagai berikut

$$\widehat{S}(t) = \prod_{i|t_i \le t} \frac{n_i - d_i}{n_i} \qquad (12)$$

dimana

$\widehat{S}(t)$ = estimasi fungsi *survival*

$n_i$ = banyaknya data yang lebih dari sama dengan $t_i$

$d_i$ = jumlah individu yang mengalami *event* pada saat $t_i$

(Hosmer dan Lemeshow, 2008)

**Model Regresi Cox *Proportional Hazard***

Model regresi Cox *proportional hazard* (PH) adalah salah satu cara yang digunakan untuk mengetahui hubungan antara waktu *survival* dengan variabel-variabel yang diduga mempengaruhi waktu survival. Bentuk dari model regresi Cox PH adalah

$$h(t, x) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \ldots \beta_p x_p) \quad (13)$$

dengan $h_0(t)$ adalah fungsi *baseline hazard*, $\beta_1, \beta_2, \ldots, \beta_p$ adalah parameter regresi, dan $x_1, x_2, \ldots, x_p$ adalah variabel bebasnya (Lee dan Wang, 2003)

**Estimasi Parameter dengan Pendekatan Efron *Partial Likelihood***

Pendekatan Efron dianggap sebagai pendekatan yang lebih baik daripada pendekatan Breslow, walaupun pendekatan Breslow merupakan pendekatan yang lebih sering digunakan. Semakin besar jumlah *ties* pada data, maka hasil estimasi parameter semakin tidak akurat karena fungsi *partial likelihood* menjadi semakin berbeda dari fungsi *partial likelihood* yang sebenarnya. Pendekatan Efron keduanya memiliki perhitungan yang akurat dibandingkan pendekatan Breslow, terutama ketika data mengandung banyak *ties* Fungsi *partial likelihood* dengan pendekatan Efron adalah sebagai berikut:

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{r} \frac{\exp\left(\sum_{\ell \in D(t_i)} \boldsymbol{\beta}^T \mathbf{x}_\ell\right)}{\prod_{k=1}^{d_i}\left[\sum_{q \in R(t_i)} \exp\left(\boldsymbol{\beta}^T \mathbf{x}_q\right) - \frac{k-1}{d_i} \sum_{\ell \in D(t_i)} \exp\left(\boldsymbol{\beta}^T \mathbf{x}_\ell\right)\right]} \qquad (14)$$

dengan

$\boldsymbol{\beta}^T = \begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_p \end{bmatrix}$

$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \ldots & x_p \end{bmatrix}$

$D(t_i)$ = himpunan individu yang mengalami *event* pada saat $t_i$

$R(t_i)$ = himpunan individu yang berisiko mengalami *event* pada saat $t_i$

$k$ = individu yang mengalami *ties*

$d_i$ = banyaknya *ties* yang teramati pada saat $t_i$

(Allison, 2010)

**Pengujian Asumsi *Proportional Hazard***

Kleinbum dan Klein (2005) menyebutkan bahwa pendekatan yang dapat dilakukan untuk pemeriksaan asumsi *proportional hazard* yaitu uji *goodness of fit* dengan residual Schoefeld. Pengujian asumsi dengan pendekatan *goodness of fit* merupakan variasi dari tes Schoenfeld, yaitu menggunakan residual Schoenfeld. Residual Schoenfeld didefinisikan hanya pada waktu survival yang tidak tersensor. Residual Schoenfeld untuk individu yang mengalami *event* pada saat $t_1$ pada variabel bebas ke-j adalah sebagai berikut :

$$R_{ij} = \delta_i \left( x_{ij} - \frac{\sum_{q \in R(t_i)} x_{qj} \exp\left(\hat{\boldsymbol{\beta}}^T \mathbf{x}_q\right)}{\sum_{q \in R(t_i)} \exp\left(\hat{\boldsymbol{\beta}}^T \mathbf{x}_q\right)} \right) \qquad (15)$$

dengan $j=1,2,\ldots,p; i=1,2,\ldots,n$ dan $R(t_i)$ merupakan himpunan dari individu-individu yang berisiko pada waktu $t_i$ lalu $\delta_i$ adalah indikator penyensoran, $\delta_i =$ 1 untuk data tidak tersensor, $\delta_i =$ 0 untuk data tersensor dan $\hat{\boldsymbol{\beta}}$ merupakan estimator *partial likelihood* maksimum dari $\boldsymbol{\beta}$.

Hipotesis yang diuji adalah sebagai berikut :

$H_0$: Tidak terdapat korelasi antara residual Schoenfeld dengan peringkat waktu survival

$H_1$: Terdapat korelasi antara residual Schoenfeld dengan peringkat waktu survival

Statistik uji yang digunakan adalah :

$$z = r\sqrt{n-1} \tag{16}$$

dimana

$$r = \frac{\sum_{i=1}^{r}\left(R_{ij}-\bar{R}_{ij}\right)\left(RT_i - \overline{RT_i}\right)}{\sqrt{\sum_{i=1}^{r}\left(R_{ij}-\bar{R}_{ij}\right)}\sqrt{\sum_{i=1}^{r}\left(RT_i - \overline{RT_i}\right)}}$$

$R_{ij}$ = residual Schoenfeld untuk individu yang mengalami *event* pada saat $t_i$

$\bar{R}_{ij}$ = rata-rata residual Schoenfeld untuk individu yang mengalami *event* pada saat $t_i$ untuk variabel bebas ke-j

$RT_i$ = peringkat waktu survival untuk waktu kejadian ke-i

$\overline{RT_i}$ rata-rata peringkat waktu survival untuk waktu kejadian ke-i

Apabila nilai $|z| > z_{\alpha/2}$ atau $p-value \le \alpha$ akan dapat disimpulkan bahwa terdapat nilai korelasi antara residual Schoenfeld dengan peringkat waktu survival sehingga asumsi proportional hazard tidak terpenuhi.

**Pengujian Sigifikansi Parameter**

Pengujian signifikansi parameter bertujuan untuk memeriksa apakah variabel bebas memiliki pengaruh nyata dalam model regresi. Pengujian signifikansi parameter regresi Cox proportional hazard dilakukan secara serentak menggunakan uji *partial likelihood ratio* dan secara individu menggunakan uji Wald.

1. Pengujian parameter secara serentak

   Pengujian parameter secara serentak, digunakan untuk menguji pengaruh variabel bebas dalam model regresi secara serentak atau bersamaan. Hipotesis dari pengujian ini adalah

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

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

   Statistik uji ditentukan dengan *likelihood ratio test*, yang disimbolkan dengan $G$, rumusnya sendiri adalah sebagai berikut :

   $$G = 2\left[\ln L\left(\hat{\boldsymbol{\beta}}\right) - \ln L(\mathbf{0})\right] \tag{17}$$

   dimana

   $\ln L(\mathbf{0})$ = *log partial likelihood* dari model regresi Cox *proportional hazard* tanpa variabel bebas, yaitu $\boldsymbol{\beta} = \mathbf{0}$

   $\ln L(\hat{\boldsymbol{\beta}})$ = *log partial likelihood* dari model regresi Cox *proportional hazard* yang memuat $p$ variabel bebas.

2. Uji Parsial

   Uji parsial digunakan untuk menguji pengaruh parameter $(\beta_j)$ secara individu. Hipotesis dalam pengujian untuk variabel *j* dengan $j=1,2,\ldots,p$ adalah

   $H_0 : \beta_j = 0$

   $H_1 : \beta_j \ne 0$

   Statistik uji ditentukan dengan metode pengujian Wald. Rumus umum statistik uji Wald adalah sebagai berikut :

   $$W^2 = \left(\frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}\right)^2 \tag{18}$$

   dimana

   $SE(\hat{\beta}_j)$ = $\sqrt{a_{ij}}$ , dengan $a_{ij}$ adalah elemen diagonal utama ke-*j* dari matriks $-\left[I(\hat{\boldsymbol{\beta}})\right]^{-1}$

**Interpretasi *Hazard Ratio***

Secara umum, *hazard ratio* (HR) didefinisikan sebagai perbandingan antara fungsi *hazard* individu satu dengan fungsi *hazard* individu yang lain. Misalkan individu pertama mempunyai nilai *hazard* $h(t, x_j)$ dimana $x_j = 1$, dan individu kedua mempunyai *hazard* $h(t, x_j^*)$ dimana $x_j^* = 0$, bentuk *hazard ratio* sebagai berikut :

$$HR = \frac{h(t,x_j)}{h(t,x_j^*)} = \frac{h_0(t)\exp(\beta_j x_j)}{h_0(t)\exp(\beta_j x_j^*)} = \exp\left(\beta_j(1-0)\right)$$

$$= \exp\left(\beta_j\right) \tag{19}$$

(Kleinbaum dan Klein, 2005)

Menurut Hosmer dan Lemeshow (2008), jika model regresi Cox PH mengandung variabel bebas kontinu, maka interpretasi *hazard ratio* tergantung pada seberapa besar perubahan $c$ unit variabel bebas yang akan diamati. Misalkan individu pertama mempunyai nilai *hazard* $h\left(t, x_j + c\right)$ dan individu kedua mempunyai *hazard* $h\left(t, x_j\right)$ maka berdasarkan Persamaan diperoleh bentuk *hazard ratio* sebagai berikut :

$$HR = \frac{h(t, x_j)}{h(t, x_j)} = \frac{h_0(t)\exp(\beta_j(x_j + c))}{h_0(t)\exp(\beta_j x_j)}$$
$$= \exp(\beta_j(x_j + c) - \beta_j x_j)$$
$$= \exp(c\beta_j) \qquad (20)$$

**Pemilihan Model Terbaik**

Analisis regresi seringkali digunakan untuk mengkaji hubungan antara beberapa variabel. Agar diperoleh hasil analisis yang optimal, maka diperlukan model regresi terbaik. Beberapa metode dapat digunakan untuk memilih model regresi terbaik, diantaranya adalah dengan kriteria *Akaike's Information Criterion* (AIC). Kriteria AIC adalah metode yang dapat digunakan untuk memilih model regresi terbaik yang ditemukan oleh Akaike. Kriteria ini didasarkan pada metode *maximum likelihood estimation* (MLE). Untuk menghitung nilai AIC digunakan rumus sebagai berikut :

$$AIC = e^{\frac{2m}{o}} \frac{\sum_{i=1}^{o} \hat{u}_i^2}{o} \qquad (21)$$

dengan :

$m =$ jumlah parameter yang diestimasi dalam model regresi

$o =$ jumlah observasi

$\hat{u} =$ Residual $R_{ij}$

(Fathurahman, 2009)

**Metode Penelitian**

Penelitian ini dilakukan dengan cara mengumpulkan data pasien Tuberkulosis pada tahun 2017. Rancangan penelitian ini adalah rancangan yang bersifat *ex post facto* dimana data dikumpulkan setelah semua kejadian yang dipersoalkan berlangsung atau lewat. Penelitian ini merupakan penelitian non eksperimen karena data yang dikumpulkan bukan dari pengamatan langsung. Penelitian ini merupakan studi literatif dengan rancangan model Regresi Cox Weibull. Teknik pengumpulan data yang digunakan adalah pengambilan data sekunder yang berasal dari rekam medis pasien penderita Tuberkulosis dari Puskesmas Loa Ipuh Tenggarong

**Teknik Analisis Data**

Teknik analisis data terdiri dari analisis deskriptif dan analisis model regresi Cox. Analisis deskriptif dan analisis model regresi Cox menggunakan aplikasi R 3.2.2

1. Analisis deskriptif
   Analisis deskriptif dimaksudkan untuk memberikan gambaran angka-angka dari objek penelitian. Data pasien yang mengalami *ties*, disajikan dalam bentuk tabel.

2. Analisis model Regresi Cox
   a. Penentuan data tersensor materi *survival*
      Jenis penyensoran data yang digunakan adalah sensor kanan, dengan ketentuan, jika pasien rawat jalan atau dinyatakan sembuh, selama masih dalam waktu penelitian, maka dikategorikan sebagai data *survival* tidak tersensor yang dinyatakan dengan nilai 1 dan jika pasien dinyatakan sembuh melebihi waktu penelitian, maka data *survival* diberi nilai 0.
   b. Pengujian distribusi Weibull terhadap data waktu *survival*
      Pengujian distribusi dilakukan untuk mengecek data yang telah diperoleh, terdistribusi secara Weibull atau tidak dengan menggunakan uji Kolmogorov-Smirnov.
   c. Pengujian asumsi PH
      Pengujian asumsi PH dengan uji *goodness of fit*
   d. Penentuan model regresi Cox
      Membuat model Cox PH pada kejadian bersama dilakukan dengan membentuk model regresi berdasarkan variabel bebas yang parameternya akan diuji signifikansi. Konstruksi model regresi Cox dibentuk berdasarkan Persamaan (13)
   e. Penaksiran parameter model regresi Cox PH
      Penaksiran parameter model regresi Cox dilakukan dengan pendekatan Efron *partial likelihood*. Penaksiran parameter Efron dilakukan berdasarkan Persamaan (14).
   f. Uji sigifikansi parameter
      Parameter-parameter akan diuji signifikansi terhadapa model secara serentak *partial likelihood ratio* dan secara parsial dengan uji Wald. Pengujian signifikansi parameter menggunakan statistik uji 5%.
   g. Uji kebaikan model.
   h. Interpretasi model Cox PH pada kasus kejadian bersama.

**Hasil Penelitian dan Pembahasan**

**1. Statistika Deskriptif**

Data penelitian terdiri dari data tidak tersensor dan data tersensor, dari 47 data pasien yang diperoleh, berdasarkan lama perawatan, terdapat 16 data yang dikategorikan tidak teramati atau tersensor. Data *ties* pada penelitian ini terjadi apabila terdapat dua atau lebih pasien penderita tuberkulosis yang menjalani perawatan dari Puskesmas Loa Ipuh Tenggarong pada tahun 2017 secara bersamaan. Data lama perawatan pasien tuberkulosis yang mengalami kejadian bersama (*ties*) ditunjukan dalam Tabel 1.

Tabel 1. Data Lama Perawatan Pasien Tuberkulosis Berdasarkan *Ties*

| Lama Perawatan (Hari) | Banyaknya Pasien |
|---|---|
| 118 | 2 |
| 151 | 2 |
| 165 | 3 |
| 167 | 9 |
| 168 | 5 |
| 169 | 2 |
| 170 | 3 |
| 171 | 3 |

Tabel 1 menunjukkan bahwa dari 47 pasien, diantaranya terdapat 29 pasien yang mengalami *ties*, dan diantara 29 pasien yang mengalami *ties*, terdapat 4 pasien yang masuk dalam kategori tersensor.

**2. Pengujian Distribusi**

Pengujian distribusi bertujuan untuk melihat apakah data dari lama perawatan pasien penderita Tuberkulosis telah distribusi Weibull, dengan hipotesis yang digunakan adalah

$H_0$: $F(x) = F^*(x)$ (Data berdistribusi Weibull)

$H_1$: $F(x) \neq F^*(x)$ (Data tidak berdistribusi Weibull)

Statistik uji dalam pengujian distribusi seperti yang diberikan oleh Persamaan (11), dengan taraf signifikansi $\alpha = 0,05$ dan kriteria $H_0$ ditolak jika $D_{hitung} > 0,264$ atau $P_{value} \leq 0,05$. Hasil perhitungan statistik uji $D_{hitung}$ dan $P_{value}$ disajikan pada Tabel 2.

Tabel 2 Nilai $D_{hitung}$ dan $P_{value}$

| Nilai $D_{hitung}$ | Nilai $P_{value}$ |
|---|---|
| 0,204 | 0,132 |

Berdasarkan hasil pengujian distribusi pada Tabel 2, dengan taraf signifikansi 0,05 diperoleh nilai $D_{hitung} = 0,204$ lebih kecil daripada nilai $D_{tabel} = 0,264$, maka $H_0$ gagal ditolak, sehingga dapat disimpulkan bahwa data dari lama perawatan pasien penderita Tuberkulosis berdistribusi Weibull.

**3. Pengujian Asumsi PH**

Pemeriksaan asumsi PH dilakukan sebelum penentuan model. Asumsi ini dapat terpenuhi dengan uji *goodness of fit* untuk tiap variabel penjelas yang berupa kategorik, yaitu variabel jenis kelamin, klasifikasi penyakit, pemeriksaan dahak, dan asal tempat tinggal. Pengujian asumsi PH dengan pendekatan dengan uji *goodness of fit* ditunjukkan oleh Tabel 3.

Tabel 3 Hasil Pengujian Asumsi PH

| Variabel | $P_{value}$ | Keputusan |
|---|---|---|
| Jenis Kelamin | 0,616 | $H_0$ Gagal Ditolak |
| Klasifikasi Penyakit | 0,218 | $H_0$ Gagal Ditolak |
| Pemeriksaan Dahak | 0,123 | $H_0$ Gagal Ditolak |
| Asal Tempat Tinggal | 0,008 | $H_0$ Ditolak |

Berdasarkan hasil pengujian asumsi PH pada Tabel 3, diperoleh $P_{value}$ lebih besar daripada $\alpha$ maka $H_0$ gagal ditolak untuk variabel jenis kelamin, klasifikasi penyakit, dan pemeriksaan dahak. Sehingga dapat disimpulkan bahwa variabel Jenis Kelamin, Klasifikasi Penyakit, dan Pemeriksaan Dahak telah memenuhi asumsi PH.

**4. Analisis Regresi Cox**

Langkah pertama dalam analisis regresi Cox adalah penaksiran parameter model regresi Cox PH. Penghitungan penaksir parameter model Cox PH menggunakan *software* R 3.2.2. Hasil penaksiran model Cox PH dengan semua variabel bebas dapat dilihat pada Tabel 4.

Tabel 4 Hasil estimasi Parameter Model Cox untuk Semua Variabel

| Variabel | Parameter | $\hat{\beta}$ | exp($\hat{\beta}$) |
|---|---|---|---|
| Jenis Kelamin | $\beta_1$ | -0,167 | 0,846 |
| Usia | $\beta_2$ | 0,032 | 1,032 |
| Berat Badan | $\beta_3$ | 0,021 | 1,022 |
| Klasifikasi Penyakit | $\beta_4$ | -0,646 | 0,524 |
| Pemeriksaan Dahak | $\beta_5$ | -0,064 | 0,938 |
| Asal Tempat Tinggal | $\beta_6$ | -0,201 | 0,818 |

Berdasarkan Tabel 4, Persamaan (13) dapat dituliskan kembali menjadi :

$$\hat{h}(t,\mathbf{x}) = \hat{h}_0(t)\exp(-0.167x_1 + 0.032x_2 + 0.021x_3 - 0.646x_4 - 0.064x_5 - 0.201x_6) \quad (22)$$

Tahapan selanjutnya setelah penaksiran parameter adalah pengujian parameter yang terdiri dari pengujian hipotesis parameter regresi secara serentak dan parsial. Statistik uji dalam pengujian serentak diberikan oleh Persamaan (17) dengan taraf signifikansi 0,05 kriteria $H_0$ ditolak jika $G_{hitung} \geq \chi^2_{(0,05;6)}$ atau $P_{value} \leq 0,05$. Perhitungan statistik uji menggunakan *software* R 3.2.2 dan hasil pengujian hipotesis parameter Cox PH secara serentak disajikan pada Tabel 5.

Tabel 5 Uji Serentak untuk Semua Variabel

| Variabel | Nilai G | Nilai $\chi^2_{tabel}$ | $P_{value}$ |
|---|---|---|---|
| $x_1, x_2, x_3, x_4, x_5, x_6$ | 7,78 | 12,6 | 0,255 |

Berdasarkan Tabel 5, diperoleh nilai $G_{hitung} = 7,78$ lebih kecil daripada nilai $\chi_{tabel}{}^2 = 12,6$ atau $P_{value} = 0,255$ lebih besar daripada nilai $\alpha = 0,05$, maka $H_0$ gagal ditolak dan dapat disimpulkan bahwa model regresi dengan 6 variabel bebas secara bersamaan tidak berpengaruh pada model Cox PH atau model PH belum layak untuk digunakan.

Karena model yang telah diperoleh belum layak untuk digunakan, langkah selanjutnya adalah Pemilihan model terbaik dengan eliminasi Backward dan didasarkan pada hasil pengujian parameter secara serentak yang ditunjukkan oleh nilai statistik uji *G* atau $P_{value}$, juga didasarkan pada nilai AIC. Diperoleh model PH terbaik adalah model PH yang memuat dua variabel bebas, yaitu variabel usia ($x_2$) dan klasifikasi penyakit ($x_4$) yang dinyatakan oleh nilai AIC yang paling kecil yaitu sebesar 160,42. Berdasarkan nilai *G*, didapat $G_{hitung} = 6,69$ lebih besar daripada $\chi_{tabel}{}^2 = 5,6$ atau $P_{value} = 0,035$ lebih kecil daripada nilai $\alpha = 0,05$, sehingga dapat disimpulkan bahwa model regresi Cox PH dengan variabel bebas Usia ($x_2$) dan Klasifikasi Penyakit ($x_4$) sudah layak digunakan.

Hasil penaksiran model Cox PH dengan variabel bebas Usia ($x_2$) dan Klasifikasi Penyakit ($x_4$) dapat dilihat pada Tabel 6.

Tabel 6 Hasil Estimasi Parameter Model Cox Untuk Variabel Berpengaruh

| Variabel | Parameter | $\hat{\beta}$ | exp( $\hat{\beta}$ ) |
|---|---|---|---|
| Usia | $\beta_2$ | 0,024 | 1,024 |
| Klasifikasi Penyakit | $\beta_4$ | -0,637 | 0,529 |

Berdasarkan Tabel 6, model regresi Cox PH pada data lama perawatan pasien penderita tuberkulosis adalah

$$\hat{h}(t, x_2, x_4) = \hat{h}_0(t)\exp(0,024x_2 - 0,637x_4) \quad (23)$$

Hasil perhitungan statistik uji parsial seperti yang disajikan pada Tabel 7.

Tabel 7 Hasil Keputusan Uji Parsial Variabel Bebas

| Variabel | Parameter | $P_{value}$ | Keputusan |
|---|---|---|---|
| Usia | $\beta_2$ | 0,048 | $H_0$ ditolak |
| Klasifikasi Penyakit | $\beta_4$ | 0,107 | $H_0$ gagal ditolak |

Berdasarkan nilai $P_{value}$ yang disajikan pada Tabel 7, disimpulkan bahwa variabel bebas usia secara individual berpengaruh terhadap lama waktu perawatan pasien penderita tuberkulosis. Jadi faktor yang berpengaruh terhadap lama perawatan pasien penderita tuberkulosis adalah faktor usia. Berdasarkan variabel yang berpengaruh, didapat nilai rasio *hazard* untuk variabel usia adalah $e^{(0,024)} = 1,024$, artinya setiap peningkatan usia pasien 1 tahun, maka lama perawatan sampai pasien sembuh, meningkat menjadi 1,024 kali.

Berdasarkan *hazard* Weibull pada Persamaan (8) dan Persamaan (23), model regresi Cox PH Weibull pada data lama perawatan pasien penderita tuberkulosis dapat dituliskan kembali menjadi

$$\hat{h}(t, x_2, x_4) = 0,377\left(\frac{t}{169,7}\right)^{63,09}\exp(0,024x_2 - 0,637x_4) \quad (27)$$

## Kesimpulan

Berdasarkan hasil analisis data dan pembahasan, dapat ditarik kesimpulan sebagai berikut :

1. Berdasarkan hasil pengujian dengan metode Kolmogorov-Smirnov, dapat disimpulkan bahwa data dari lama pengobatan pasien penyakit tuberkulosis di Puskesmas Loa Ipuh, Tenggarong tahun 2017 berdistribusi Weibull.
2. Model regresi Cox PH Weibull pada data lama perawatan penderita tuberkulosis adalah
$$\hat{h}(t, x_2, x_4) = 0,377\left(\frac{t}{169,7}\right)^{63,09}\exp(0,024x_2 - 0,637x_4)$$
dengan *t* menyatakan waktu, $x_2$ adalah Usia, dan $x_4$ adalah Klasifikasi Penyakit.
3. Faktor-faktor yang berpengaruh signifikan terhadap lama perawatan pasien penyakit tuberkulosis di Puskesmas Loa Ipuh, Tenggarong adalah Usia.

4. Berdasarkan model regresi Cox PH Weibull yang didapat, bahwa setiap peningkatan usia pasien 1 tahun, maka lama perawatan sampai pasien sembuh, meningkat menjadi 1,024 kali.

**Daftar Pustaka**

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

Badan Pusat Statistik. (2017). Kutai Kartanegara Dalam Angka. Kukar:BPS.

Bastyan, E. dan Latra, I. N. (2013). Analisis Survival dengan Model Regresi Cox Weibull pada Penderita Demam berdarah Dengue (DBD) di Rumah Sakit Haji Sukolilo Surabaya. *Jurnal Sains dan Seni Pomits*, Vol 2, No,2 p. D165-D170.

CNN Indonesia. Beban Penyakit Tuberkulosis di Indonesia. Diakses pada tanggal 04 Agustus 2018, dari https://www.cnnindonesia.com.

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

Corder, G. W. dan Dale I. F. (2014). *Nonparametric Statistics*. New Jersey: John Wiley & Sons, Inc.

Fathurahman, M. (2009). Pemilihan Model Regresi Terbaik Menggunakan Metode Akaike's Information Criterion dan Schwarz Information Criterion. *Jurnal Informatika Mulawarman*, Vol 4, 37-41.

Hosmer, D. W. dan Lemeshow, S. (2008). *Applied Survival Analysis:Regression Modeling of Time to Event Data*. Hoboken: John Wiley& Sons, Inc.

Kleinbaum, D. G. dan Klein, M. (2005). *Survial Analysis: A Self-Learning text. Second Edition*. New York: Springer Science+Business Media, Inc.

Lee, T. E. dan Wang, J. W. (2003). *Statistical Method for Survival Analisis*. Third Edition. London: Jhon Wiley and Sons Interscience Publications.

Prabawati, S. (2017). Analisis *Survival* Data Kejadian Bersama dengan Pendekatan Efron *Partial Likelihood*. *Jurnal Eksponensial Universitas Mulawarman*, Samarinda.

Rinne, H. (2009). *The Weibull Distribution Handbook*. Boca Raton: Chapman & Hall/CRC.

Subuh, M. dan Priohutomo, S. (2014). *Pedoman Nasional Pengendalian Tuberkulosis.* Jakarta: Kementerian Kesehatan RI.

Suyitno. (2017). "Penaksiran Parameter dan Pengujian Hipotesis Model Regresi Weibull Univariat". *Jurnal EKSPONENSIAL Vol 8, No 2*,179-183, ISSN: 2085-7829. http://jurnal.fmipa.unmul.ac.id/index.php/exponensial/article/view/41/20 (diakses tanggal 23Januari 2019)