Journal of Mathematics, Computations, and Statistics (hal. 62 – 77)
Vol. 1. No. 1, April 2018
http://www.ojs.unm.ac.id/jmathcos

Model Regresi Cox dan Aplikasinya dalam Menganalisis Ketahanan Hidup Pasien Penderita Diabetes Mellitus di Rumah Sakit Bhayangkara Makassar

Wahidah Sanusi[1], Alimuddin[1], dan Sukmawati[1, a)]

[1] *Jurusan Matematika, FMIPA Universitas Negeri Makassar, 90224*

[a)] e-mail: sukmawatishizukaru@gmail.com

Abstrak. *Analisis tahan hidup adalah salah satu prosedur statistik untuk melakukan analisa data berupa waktu tahan hidup dan variabel yang mempengaruhi waktu tahan hidup. Pada penelitian ini analisis tahan hidup diaplikasikan pada kasus diabetes mellitus di Rumah Sakit Bhayangkara Makassar pada tahun 2016. Salah satu metode analisis tahan hidup yang digunakan adalah model Regresi Cox Proporsional Hazard. Penggunaan model regresi cox proporsional hazard harus memenuhi asumsi proporsional hazard. Penelitian ini juga menggunakan distribusi eksponensial dua parameter untuk menentukan fungsi hazard dan metode Breslow dalam membentuk model cox terbaik. Dari hasil penelitian diperoleh faktor-faktor signifikan yang mempengaruhi waktu tahan hidup adalah umur dan kadar gula darah, namun faktor kadar gula darah tidak memenuhi asumsi proporsional hazard, sehingga digunakan Model Cox Extended untuk memperbaiki model cox proporsional hazard. Covariate yang tidak memenuhi asumsi proporsional hazard dalam model cox extended dinteraksikan dengan fungsi waktu* $g(t) = t$. *Model Cox Extended pada akhirnya memberikan informasi tentang faktor -faktor yang berpengaruh signifikan terhadap waktu tahan hidup yaitu umur dan kadar gula darah terikat waktu, dimana setiap individu yang berumur kurang dari 45 tahun memiliki resiko kegagalan 0,015 kali lebih kecil dibandingkan dengan pasien yang berumur lebih dari 45 tahun dan individu yang kadar gula darahnya tinggi memiliki resiko kegagalan sebesar 1,128 kali lebih besar dibandingkan dengan pasien yang memiliki kadar gula darah rendah dan normal.*
Kata Kunci: *Analisis Tahan Hidup, Regresi Cox Proporsional Hazard, Diabetes Mellitus, Model Cox Extended*

Abstract. *Survival analyze is one of the statistical procedures to analyze data survival time and variable that will affect the rate of recovery of patients. In this research, survival analyze was applicated by diabetes mellitus case in Bhayangkara Hospital Makassar 2016. One of the methods survival analyze used is cox regression model with proportional hazard. The use of cox regression model with proportional hazard must fulfill assumption of proportional hazard. This research also use 2-parameter exponential distribution to determine of hazard function and Breslow method to shaping the best of cox model. From the results of the research give conclusion that factors affecting of time recovery are age and blood sugar level. But the blood sugar level factor does not fulfill the proportional hazard assumptions. So that the extended cox model was used to improve the cox proportional hazard model. Variables that does not fulfill the proportional hazard assumption in the extended cox model are interacted with the time function* $g(t) = t$. *Finally, the extended cox model give information about the factors most affect the rate of recovery are age and time bound blood sugar level. Every individual less than 45 years old has a 0,015 times greater risk of failure than patients older than 45 years old and individuals with high blood sugar level had a risk of failure is 1,128 times greater than the low and normal blood sugar level*
Keywords: *Survival Analyze, Cox Proportional Hazard Regression, Diabetes Mellitus, Extended Cox Model*

PENDAHULUAN

Dalam kehidupan sehari-hari banyak ditemui masalah yang berkaitan dengan waktu, seperti kambuhnya suatu penyakit yang diderita seseorang, waktu menganggur setelah lulus kuliah sampai mendapatkan pekerjaan, waktu menyelesaikan disertasi doktor, dan lain sebagainya.Waktu-waktu tersebut dapat dipandang sebagai peubah respon (response variable) atau peubah tak bebas (dependent variable), sedangkan faktor-faktor lain yang mempengaruhi waktu tersebut dapat dipandang sebagai peubah penjelas (explanatory variable) atau peubah bebas (independent variable). Jangka waktu sampai terjadinya suatu kejadian dalam statistika dikenal dengan istilah waktu tahan hidup (survival time) (Saefuddun & Ratnaningsih, 2008). Analisis tahan hidup adalah salah satu prosedur statistik untuk melakukan analisa data berupa waktu tahan hidup dan variabel yang mempengaruhi waktu tahan hidup, yaitu data waktu tahan hidup mulai dari waktu awal penelitian yang sudah ditentukan sampai waktu terjadinya suatu kejadian. Terdapat dua cara yang dapat dilakukan dalam pengambilan sampel pada analisis data tahan hidup, yaitu pengamatan tersensor dan pengamatan tidak tersensor (Mandini, 2015).

Salah satu analisis survival yang digunakan adalah regresi Cox. 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). Pada dasarnya model regresi cox terdiri dari dua, yaitu regresi cox proportional hazard dan regresi cox nonproportional hazard (Ridwan, 2016).

Penggunaan model regresi Cox proportional hazard harus memenuhi asumsi proportional hazard, Jika asumsi ini tidak terpenuhi, maka model dikatakan nonproportional hazard. Salah satu perluasan model Cox yang memperhatikan pelanggaran asumsi proportional hazard adalah model Cox extended. Model Cox extended merupakan perluasan model dari model Cox proportional hazard, yaitu mengandung variabel terikat oleh waktu atau perkalian dari variabel bebas dengan fungsi waktu. Fungsi waktu yang dapat digunakan dalam model Cox extended antara lain, $g(t) = 0, g(t) = t, g(t) = \log t$ dan fungsi Heaviside (Vitriana & R, 2016).

Secara global diperkirakan 346 juta orang menderita diabetes, dan diperkirakan akan menjadi penyakit terbesar ke tujuh yang menyumbang kematian pada tahun 2030. WHO pada September 2012 menjelaskan bahwa jumlah penderita DM di dunia mencapai 347 juta orang dan lebih dari 80% kematian akibat DM terjadi pada negara miskin dan berkembang. Berdasarkan data Dinas Kesehatan Provinsi Sulawesi Selatan, Makassar merupakan kota dengan penderita DM terbanyak. Pada tahun 2010 terdapat 3827 kasus baru dari 17245 atau sekitar 22,19% .Di Rumah Sakit Bayangkara sendiri prevalensi DM juga terus mengalami peningkatan yang ditandai dengan jumlah kasus yang terus cenderung meningkat dari tahun ke tahun. Studi pendahuluan yang dilakukan oleh peneliti dari data rekam medik Rumah Sakit Bayangkara terdapat 356 kasus baru dari jumlah 2741 atau sekitar 12,99% (Masfufah, Hadju, & Djafar, 2014).

Dasar Teori Analysis Survival (Analisis Ketahanan Hidup)

Analisis *survival* merupakan suatu metode yang berkaitan dengan waktu, mulai dari *time origin* atau *start point* sampai dengan terjadinya suatu kejadian khusus (*failure event/end point*). Analisis *survival* memerlukan data yang merupakan waktu *survival*. Penentuan waktu *survival*, ada tiga faktor yang dibutuhkan yaitu: 1) awal pencatatan (*time origin* atau *start-point*) harus didefinisikan dengan tepat pada setiap individu, misalkan awal mula pengamatan berupa tanggal perawatan pasien, 2) akhir pencatatan (*failure time* atau *end-point*) didefinisikan jelas untuk mengetahui status tersensor atau tidak tersensor, meninggal atau sembuh seorang pasien, 3) skala pengukuran sebagai batas dari waktu kejadian dari awal sampai akhir kejadian, misalnya skala tahunan, bulanan, harian, mingguan, harian.

Distribusi Survival

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

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

Misalkan T adalah variabel random bukan negatif pada interval $[0,\infty)$ yang menunjukkan waktu hidup pada suatu populasi dan $f(t)$ merupakan fungsi kepadatan peluang dari s maka fungsi distribusi kumulatif $F(t)$ adalah (Iskandar, 2015).

$$\begin{aligned}F(T)&=P(T\le t)\\&=\int_0^t f(x)dx\end{aligned} \quad (2)$$

Fumgsi Survival

Dari definisi fungsi distribusi kumulatif dari T, fungsi *survival* dapat dinyatakan sebagai berikut

$$\begin{aligned}S(t)&=P(T\ge t)\\&=1-P(T\le t)\\&=1-F(t)\end{aligned} \quad (3)$$

Fungsi Hazard

Misalkan T variabel random non negatif pada interval $[0,\infty)$ yang menunjukkan waktu individu sampai mengalami kejadian pada suatu populasi, maka peluang bahwa individu mengalami kejadian pada interval $(t,t+\Delta t)$ dinyatakan dengan fungsi *hazard* $h(t)$ (Iskandar, 2015)

$$\begin{aligned}h(t)&=\lim_{\Delta t\to 0}\frac{P(t\le T<t+\Delta t|T\ge t)}{\Delta t}\\&=\lim_{\Delta t\to 0}\frac{P(t\le T<t+\Delta t,T\ge t)}{\Delta t.P(T\ge t)}\\&=\frac{F'(t)}{S(t)}\\&=\frac{f(t)}{S(t)}\end{aligned} \quad (4)$$

Hazard Kumulatif

Fungsi *hazard* maka fungsi kumulatif *hazard* dinyatakan dengan $H(t)$ (Iskandar, 2015)

$$H(t)=\int_0^t h(x)dx \quad (5)$$

Analisis Distribusi

Pendugaan distribusi dilakukan dengan statistik uji *Anderson-Darling* untuk mengetahui distribusi data *survival* yang paling sesuai. Persamaan statistik uji *Anderson-Darling* dapat dituliskan pada persamaan sebagai berikut

$$A^2=-n-\frac{1}{2}\sum_{i=1}^{n}(2i-1)\left[\left(\ln F(X_i)+\ln\left(1-F(X_{n+1})\right)\right)\right] \quad (6)$$

Dimana :

F = fungsi distribusi kumulatif dari distribusi tertentu

X_i = data waktu survival

n = banyaknya data atau individu

Distribusi Eksponensial Dua Parameter

Suatu distribusi peluang dikatakan berdistribusi eksponensial dengan dua parameter $X \sim Exp(\theta, \eta)$, jika distribusi tersebut mempunyai fungsi kepadatan peluang

$$f(x) = \frac{1}{\theta} e^{-(x-\eta)/\theta} \quad , \eta < x \text{ dan } \theta > 0 \tag{7}$$

Model Cox Proporsional Hazard

Model *cox proportional hazard* disebut dengan model *cox* karena asumsi *proportional hazardnya* yaitu fungsi *hazard* dari individu yang berbeda adalah *proportional* atau rasio dari fungsi *hazard* dua individu yang berbeda adalah konstan (Iskandar, 2015). Melalui model *Cox* dapat dilihat hubungan antara variabel bebas (variabel independen) terhadap variabel terikat (variabel dependen) yaitu waktu *survival* melalui fungsi *hazardnya*. untuk variabel X yang ber-Covariate, maka persamaan yang digunakan adalah (Rahayu, Setiawan, & Mahatma, 2012).

$$h(t, X) = h_0(t)\exp(\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p) \tag{8}$$

dengan memisalkan,

t = Waktu Survival

$h_0(t)$ = Fungsi dasar *hazard*,

$\beta_1, \beta_2, \ldots, \beta_p$ = Parameter regresi,

$X_1, X_2, \ldots, X_p$ = Variabel bebas $i = 1,2, \ldots, p$

Estimasi Parameter

Parameter β_i pada model *Cox* proporsional hazard dapat diestimasi dengan menggunakan metode *Maximum Partial Likelihood Estimation* (MPLE). Pendugaan β_i dengan metode MPLE adalah nilai ketika fungsi *partial likelihood*-nya maksimum. Jika terdapat *n* waktu survival yang diobservasi, dinotasikan oleh $t_1, t_2, \ldots, t_n$ dan δ_i adalah *value indicator* maka fungsi *likelihood*nya dinyatakan dalam fungsi *parsial likelihood* pada persamaan sebagai berikut (Hanni & Wryandari, 2013).

$$L(\beta) = \prod_{i=1}^{r} \left[\frac{\exp(\beta X_{(i)})}{\sum_{i \epsilon R(t_i)} \exp(\beta X_{(i)})} \right]^{\delta_i} \tag{9}$$

Dengan $\delta_i = \begin{cases} 0, \text{individu yang tersensor} \\ 1, \text{ individu tidak tersensor} \end{cases}$

Terdapat tiga cara untuk menguji signifikansi parameter yaitu dengan uji *partial likelihood ratio*, uji *Wald*, dan uji *score*. Pengujian signifikasi parameter bertujuan untuk memeriksa apakah variabel bebas memiliki pengaruh nyata dalam model (Iskandar, 2015).

Uji partial likelihood rasio

Untuk menguji hipotesis bahwa satu atau beberapa parameter regresi β_j adalah nol dapat menggunakan uji *partial likelihood* rasio dinotasikan dengan G. Statistik uji ini mengikuti distribusi *chi-square* dengan derajat bebas p . Berikut langkah-langkah uji *partial likelihood* rasio:

1. Hipotesis
   $H_0: \beta_j = 0$
   $H_1: \beta_j \neq 0$
2. Daerah penolakan:
   H_0 ditolak jika $G \geq \chi^2_{\alpha:db=p-1}$ atau p-value $\leq \alpha$, dimana $G = -2\left[\ln L(0) - \ln L\left(\hat{\beta}_j\right)\right]$
   p : banyaknya variabel bebas
3. Kesimpulan:
   Jika H_0ditolak maka $\beta_i \neq 0$, mengindikasikan bahwa variabel bebas berpengaruh terhadap waktu *survival* (variabel dependen).

Uji Wald

Uji *Wald* digunakan untuk menguji pengaruh parameter secara terpisah, dinotasikan dengan W. Statistik uji ini mengikuti distribusi *chi-square* dengan derajat bebas p . Berikut langkah-langkah uji *Wald*:

1. Hipotesis:
   $H_0: \beta_j = 0$
   $H_1: \beta_j \neq 0$
2. Daerah penolakan:
   H_0 ditolak jika $W > \chi^2_{(0,05;1)}$ atau *p-value* < $\alpha = 0{,}05$, dimana $W = \left(\frac{\beta_j}{SE\beta_j}\right)^2$
   p : banyaknya variabel bebas
3. Kesimpulan:
   Jika H_0 ditolak maka $\beta_j \neq 0$, mengindikasikan bahwa variabel bebas berpengaruh terhadap waktu *survival* (variabel dependen).

Pengujian Asumsi Proporsional Hazard

Asumsi *proportional hazard* adalah suatu keadaan dimana *hazard ratio* bersifat konstan terhadap waktu Terdapat tiga pendekatan yang dapat digunakan untuk menguji asumsi *proportional hazard* yaitu pendekatan grafik, pendekatan *goodness of fit* dan pendekatan variabel *time dependent* (Afifah & Purnami, 2016).

Odss Ratio

Odds Ratio (Hazard Ratio) merupakan suatu ukuran yang digunakan untuk mengetahui tingkat risiko(kecenderungan) yaitu perbandingan fungsi *hazard* antar dua kategori, yaitu kategori sukses dengan kategori gagal pada satu peubah bebas X. Odds ratio dapat didefinisikan sebagai berikut

$$\text{ODSS RATIO} = \frac{h_0(t|x=1)}{h_0(t|x=0)} = \frac{h_0(t)e^{\beta}}{h_0(t)} = e^{\beta} \tag{10}$$

Model Cox Extended

Model Cox *Extended* merupakan perluasan dari model Cox *proportional hazard* yaitu mengandung variabel yang bergantung terhadap waktu (*time-dependent variable*). Model Cox *Extended* digunakan untuk mengatasi asumsi *proportional hazard* yang tidak terpenuhi, yaitu dengan mengalikan variabel yang tidak memenuhi asumsi dengan suatu fungsi waktu dengan bentuk umum sebagai berikut:

$$h(t, X(t)) = h_0(t)\exp\left[\sum_{i=1}^{p_1} \beta_i X_i + \sum_{i=1}^{p_2} \delta_i X_i g_i(t)\right] \tag{11}$$

dengan,

$h_0(t)$ = fungsi baseline *hazard*

X_i = variabel bebas oleh waktu ke- i dengan $i = 1,2, \dots, p_1$

δ_i = koefisien variabel bebas yang tidak memenuhi asumsi proporsional *hazard*

$g_i(t)$ = fungsi waktu untuk variabel bergantung waktu ke- i dengan $i = 1,2, \dots, p_2$

$X_i g_i(t)$ = perkalian dari variabel bebas X_i dengan waktu t

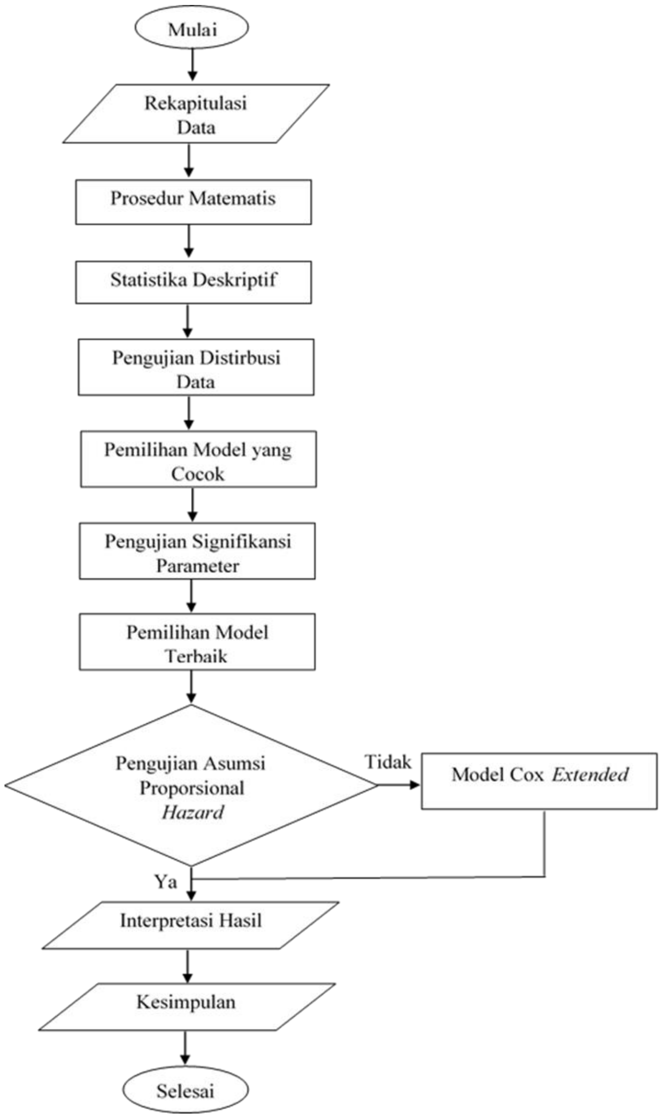
Diabetes Mellitus

Menurut WHO, Diabetes Mellitus (DM) didefinisikan sebagai suatu penyakit atau gangguan metabolisme kronis dengan multi etiologi yang ditandai dengan tingginya kadar gula darah disertai dengan gannguan metabolismekarbohidrat, lipid dan protein sebagai akibat insufisiensi fungsi insulin. Adapun faktor-faktor penyebab diabetes mellitus adalah umur, pendidikan, pekerjaan, tekanan darah, kadar gula darah, dan riwayat diabetes mellitus keluarga.

METODE PENELITIAN

Penelitian ini menggunakan data sekunder yaitu data rekam medis mengenai waktu ketahanan hidup yang diperoleh dari lama perawatan pasien penderita diabetes mellitus di Rumah Sakit Bhayangkara Makassar periode Januari – Desember 2016. Diperoleh 35 pasien , 3 data tersensor dan 32 data tidak tersensor. Variabel dependen didalam penelitian ini adalah waktu ketahanan hidup pasien diabetes mellitus, sejak awal masuk rumah sakit atau awal pengamatan sampai akhir pengamatan pada Desember 2016. Adapun variabel independen yang menjadi faktor-faktor dari ketahanan hidup pasien Diabetes Mellitus adalah umur, jenis kelamin, pendidikan, pekerjaan, tekanan darah, kadar gula darah dan riwayat DM keluarga.

Alur atau skema penelitian sebagai berikut:



GAMBAR 1. Skema Penelitian

HASIL DAN PEMBAHASAN

Prosedur matematis model distribusi

Estimasi Parameter Distribusi Eksponensial Dua Parameter

Bentuk umum fungsi kepadatan peluang distribusi eksponensial dua parameter yaitu

$$f(t) = \frac{1}{\theta}\exp\left(-\frac{t-\eta}{\theta}\right) \quad ,\eta < t, \theta > 0 \tag{12}$$

Dimana:

θ : parameter lokasi (scale parameter)

η : parameter ambang batas (threshold parameter)

Berdasarkan nilai rata-rata waktu ketahanan hidup dan jumlah data waktu ketahanan hidup, diperoleh nilai estimasi parameter $\hat{\theta}$ dan $\hat{\eta}$ sebagai berikut

$$\hat{\theta} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} t_i^2 - \bar{t}^2} = \sqrt{\frac{2312}{35} - 53{,}50} = \sqrt{12{,}557} = 3{,}543$$

$$\hat{\eta} = \bar{t} - \sqrt{\frac{1}{n}\sum_{i=1}^{n} t_i^2 - \bar{t}^2} = 7{,}314 - 3{,}543 = 3.771$$

Jadi, fungsi kepadatan peluang dari distribusi eksponensial dua parameter sebagai berikut

$$f(t) = \frac{1}{\hat{\theta}}\exp\left(-\frac{t-\hat{\eta}}{\hat{\theta}}\right)$$

$$f(t) = \frac{1}{3{,}543}\exp\left(-\frac{t-3{,}771}{3{,}543}\right) = \frac{1}{3{,}543}\exp\left(\frac{3{,}771-t}{3{,}543}\right)$$

$$F(t) = \int_0^t \frac{1}{3{,}543}\exp\left(-\frac{t-3{,}771}{3{,}543}\right)dt$$

$$F(t) = 3{,}543\exp\left(\frac{3{,}771}{3{,}543}\right) + \exp\left(\frac{3{,}771-t}{3{,}543}\right)$$

Berdasarkan persamaan (16), maka diperoleh persamaan fungsi *hazard* kumulatif sebagai berikut:

$$H(t) = h_0(t) = \frac{\frac{1}{3{,}543}\exp\left(\frac{3{,}771-t}{3{,}543}\right)}{3{,}543\exp\left(\frac{3{,}771}{3{,}543}\right)+\exp\left(\frac{3{,}771-t}{3{,}543}\right)}$$

Estimasi Parameter Model Cox Pada Kejadian Bersama

Pendekatan yang akan digunakan dalam estimasi ini adalah pendekatan metode *Breslow*. Metode *Breslow* mengasumsikan bahwa ukuran dari himpunan risiko adalah sama. Terdapat dua kasus yang memiliki waktu yang sama yaitu pada saat $t = 5$ dan $t = 6$ yang dapat dilihat pada Tabel 1 berikut:

TABEL 1. Data *survival* dengan terdapat *ties*

Individu ke-*i*	t_i	**Riwayat DM** (X_7)
1	5	2
2	5	1
3	5	2
4	5	2
5	6	1
6	6	2
7	6	2

Berdasarkan tabel 1, diperoleh fungsi *hazard* dasar pada persamaan berikut:

$$P(A|B) = P(A_1|B) \times P(A_2|B) \times P(A_3|B) \times P(A_4|B)$$

$$= \frac{\exp(\beta X_{7:1})}{\exp(\beta X_{7:1}) + \exp(\beta X_{7:2}) + \exp(\beta X_{7:3}) + \exp(\beta X_{7:4})}$$
$$\times \frac{\exp(\beta X_{7:2})}{\exp(\beta X_{7:1}) + \exp(\beta X_{7:2}) + \exp(\beta X_{7:3}) + \exp(\beta X_{7:4})}$$
$$\times \frac{\exp(\beta X_{7:3})}{\exp(\beta X_{7:1}) + \exp(\beta X_{7:2}) + \exp(\beta X_{7:3}) + \exp(\beta X_{7:4})}$$
$$\times \frac{\exp(\beta X_{7:4})}{\exp(\beta X_{7:1}) + \exp(\beta X_{7:2}) + \exp(\beta X_{7:3}) + \exp(\beta X_{7:4})}$$
$$= \frac{\exp(\beta X_{7:1} + \beta X_{7:2} + \beta X_{7:3} + \beta X_{7:4})}{[\exp(\beta X_{7:1}) + \exp(\beta X_{7:2}) + \exp(\beta X_{7:3}) + \exp(\beta X_{7:4})]^4}$$
$$= \frac{\exp(\beta X_{7:1}+\beta X_{7:2}+\beta X_{7:3}+\beta X_{7:4})}{\sum_{j=1}^{4}[\exp(\beta X_{7:j})]^4} \quad (13)$$

Sehingga, bentuk umum dari fungsi hazard sebagai berikut.

$$P(A|B) = \frac{\exp\left(\sum_{j=1}^{p} \beta_j s_j\right)}{\sum_{\iota\epsilon R(t_j)} \exp(\sum_{j=1}^{p} \beta_j X_{\iota j}))^{d_i}} \quad (14)$$

Dengan S_j adalah jumlah kovariat pada kasus ties dan d_i adalah banyaknya kasus ties pada waktu t_i. Dengan mengambil fungsi hazard pada persamaan (17), diperoleh fungsi partial likelihood sebagai berikut

$$L(\beta)_{breslow} = \prod_{i=1}^{r} \frac{\exp(\beta S_i)}{(\sum_{\iota\epsilon R(t_j)} \exp(\beta X_i))^{d_i}}, \text{dimana } \beta = (\beta_1, \beta_2, \dots, \beta_p)^t$$

$$\ln L(\beta) = \sum_{i=1}^{r} \left[\ln(\exp(\beta S_i) - \ln((\sum_{\iota\epsilon R(t_j)} \exp(\beta X_i))^{d_i}\right]$$
$$= \sum_{i=1}^{r} \left[\beta S_i - d_i \ln \sum_{\iota\epsilon R(t_j)} (\exp(\beta X_i))\right] \quad (15)$$

Statistika deskriptif data pasien diabetes mellitus

TABEL 2. Statistika Deskriptif Data Pasien Diabetes Mellitus

Variabel Independen	Kategori	Persentase
Jenis Kelamin	Laki-Laki	57,14%
	Perempuan	42,86%
Umur	Rentan	97,14%
	Tidak Rentan	2,86%
Pendidikan	Tinggi	62,86%
	Rendah	37,14%
Pekerjaan	Bekerja	62,86%
	Tidak Bekerja	37,14%
Tekanan Darah	Hipertensi	40 %
	Normal	60%

Kadar gula Darah	Tinggi	77,14%.
	Rendah	8,57%
	Normal	14,29%.
Riwayat DM Keluarga	Ada	25,71
	Tidak Ada	74,29%

TABEL 3. Analisis deskriptif terhadap variabel data kontinu

Variabel	N	Minimum	Maksimum	Rata-rata	Simpangan Baku	Varians
Survival	35	3,00	15,00	7,3143	3,59552	12,928

Berdasarkan Tabel 3 dapat dilihat bahwa rata-rata lama rawat inap pasien penderita diabetes mellitus di Rumah Sakit Bhayangkara Makassar adalah 7 hari.

TABEL 4. Hasil uji kesesuaian distribusi pada data waktu survival

Distribusi	Anderson Darling	*p-value*
Normal	2.105	<0.005
Exponential	5.298	<0.003
2-Parameter Exponential	**1.157**	**0.05**
Weibull	1.502	<0.010
Smallest Extreme Value	3.068	<0.010
Largest Extreme Value	1.052	<0.010
Gamma	1.135	0.006
Logistic	1.734	<0.005

Ada beberapa distribusi penting dalam uji survival, seperti yang diberikan pada Tabel 4. Berdasarkan Distribusi *IDplot* pada Tabel 4 untuk lama perawatan diperolehlah nilai *Anderson-Darling* dan nilai *p-value* yang ditunjukkan pada Tabel 4. Untuk menentukan apakah data mengikuti distribusi tertentu dapat dilakukan dengan membandingkan nilai Anderson-Darling atau *p-value* untuk distribusi yang diuji. Distribusi yang sesuai merupakan distribusi yang memiliki nilai Anderson Darling terkecil atau *p-value* terbesar. Sehingga dapat disimpulkan bahwa data mengikuti distribusi eksponensial dua parameter karena nilai Anderson-Darling terkecil yaitu 1,157 dan *p-value* sebesar 0,05. Berdasarkan *p-value* tersebut maka distribusi yang sesuai adalah distribusi Eksponensial Dua Parameter

Pemilihan Model yang Cocok

Pemilihan model yang cocok pada Tabel 4 di bawah diperoleh model dengan *p-value* terbesar pada variabel bebas dari setiap langkah. Proses pengeluaran variabel bebas berhenti pada langkah ke enam karena $G_6 \geq \chi^2_{(0,1;1)}$ dan $p - value < 0{,}1$ untuk semua signifikansi variabel. Berikut langkah-langkah pemilihan model terbaik dengan seleksi *backward*.

TABEL 5. Prosedur seleksi *backward* dalam pemilihan model terbaik

		Koefisien	Wald	*p-value*	Exp(B)	-2 Log Likelihood	G_i
Langkah 0	*Null*					173,545	
Langkah 1	Jenis_Kelamin	0,137	0,095	0,757	1,146	165,998	15,094
	Umur	-2,516	3,592	0,060	0,081		
	Pendidikan	0,337	0,361	0,548	1,401		
	Pekerjaan	-0,213	0,222	0,637	0,808		
	Tekanan_Darah	0,209	0,255	0,614	1,232		
	Kadar_Gula_Darah	0,564	3,325	0,068	1,757		
	Riwayat_DM Keluarga	-0,670	1,715	0,190	0,512		
Langkah 2	Umur	-2.451	3.423	0,064	0,086	166,092	0,188
	Pendidikan	0,342	0,373	0,542	1,408		
	Pekerjaan	-0.236	0,281	0,596	0,790		
	Tekanan_Darah	0,181	0,199	0,655	1,198		
	Kadar_Gula_Darah	0,541	3,275	0,070	1,717		
	Riwayat_DM Keluarga	-0,636	1,639	0,201	0,529		
Langkah 3	Umur	-2,577	4,009	0,045	0,076	166,294	0,404
	Pendidikan	0,318	0,324	0,569	1,375		
	Pekerjaan	-0,260	0,351	0,553	0,771		
	Kadar_Gula_Darah	0,558	3,605	0,058	1,747		
	Riwayat_DM_Keluarga	-0,561	1,448	0,229	0,571		
Langkah 4	Umur	-2,304	3,725	0,054	0,100	166,597	0,606
	Pekerjaan	-0,259	0,350	0,554	0,711		
	Kadar_Gula_Darah	0,570	3,858	0,050	1,768		
	Riwayat_DM_Keluarga	-0,521	1,266	0,261	0,594		
Langkah 5	Umur	-2,191	3,456	0,063	0,112	166,949	0,704
	Kadar_Gula_Darah	0,535	3,590	0,058	1,707		
	Riwayat_DM_Keluarga	-0,405	0,962	0,327	0,667		
Langkah 6	Umur	-2,332	3,993	0,046	0,097	167,863	1,828
	Kadar_Gula_Darah	0,577	4,225	0,040	1,781		

Dengan memisalkan

X_1 : Jenis Kelamin

X_2 : Umur

X_3 : Pendidikan

X_4 : Pekerjaan

X_5 : Tekanan Darah

X_6 : Kadar Gula Darah

X_7 : Riwayat DM Keluarga

Diasumsikan semua variabel berpengaruh terhadap model, selanjutnya semua variabel dimasukkan ke dalam persamaan umum model *Cox*, sehingga diperoleh estimasi model Regresi *Cox* sebagai berikut:

$$h(t,X) = h_0(t)\exp\begin{pmatrix} 0{,}137X_1 - 2{,}516X_2 + 0{,}337X_3 - 0{,}213X_4 \\ +0{,}209X_5 + 0{,}564X_6 - 0{,}670X_7 \end{pmatrix} \tag{16}$$

Untuk mengetahui apakah model pada persamaan (19) sudah tepat, maka dilakukan uji *partial likelihood ratio* sebagai berikut:

1. Hipotesis:
1. $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$ (variabel $X_1, X_2, X_3, X_4, X_5, X_6, X_7$ tidak berpengaruh dalam model)
2. $H_1: \exists \beta_i \neq 0, i = 1,2,3,4,5,6,7$ (variabel $X_1, X_2, X_3, X_4, X_5, X_6, X_7$ berpengaruh dalam model).

2. Daerah penolakan :H_0 ditolak jika $G \geq \chi^2_{(0,05;6)}$ atau *p-value* < 0,05, dimana $G = -2(ln\, L_R - ln\, L_F)$
3. Perhitungan
   Dari hasil perhitungan menggunakan software SPSS, diperoleh nilai *log likelihood* untuk model *Cox* tanpa variabel bebas (model *null*) yaitu $\ln L_R = -173{,}545$ dan nilai *log likelihood* model *Cox* yaitu $ln\, L_F = -165{,}998$, sehingga diperoleh perhitungan sebagai berikut:

$$
\begin{aligned}
G &= -2(\ln L_R - \ln L_F)\\
&= -2(-173{,}545 - (-165{,}998))\\
&= 15{,}094
\end{aligned}
$$

Karena $G^1(0) = 15{,}094 \geq \chi^2_{(0,05:6)} = 12{,}6$ dan *p-value* = 0,295 $> 0{,}05$, sehingga H_0 ditolak dan dapat disimpulkan bahwa variabel $X_1, X_2, X_3, X_4, X_5, X_6, X_7$ berpengaruh dalam model, mengindikasikan bahwa pada persamaaan (16) lebih baik daripada model tanpa variabel bebas (model *null*).

Model Regresi *Cox* berdasarkan hasil seleksi *backward* sebagai berikut:

$$h(t, X) = h_0(t) \exp(-2{,}332\, X_2 + 0{,}577\ X_6) \tag{17}$$

Pengujian Signifikansi Parameter

Tabel 6. Hasil pengujian parameter secara parsial dengan uji wald

Variabel	Koefisien	SE	$Wald$	$\chi^2_{(0,5:1)}$	*p-value*	Keputusan
Umur	-1,228	1,036	1,404	3,84	0,236	H_0 diterima
Kadar Gula Darah	0,119	0,234	0,257	3,84	0,612	H_0 diterima

Berdasarkan tabel 6 dapat disimpulkan sebagai berikut.

1. Variabel umur tidak berpengaruh secara individu terhadap waktu *survival*, hal ini dapat dilihat dari nilai $W = 1{,}404 < \chi^2_{(0,05;1)} = 3{,}84$ dan *p-value* pada tablel 4.6 yaitu 0,236 > $\alpha = 0{,}05$ maka H_0 diterima.
2. Variabel kadar gula darah tidak berpengaruh secara individu terhadap waktu *survival*, hal ini dapat dilihat dari nilai $W = 0{,}257 < \chi^2_{(0,05;1)} = 3{,}84$ dan *p-value* pada tablel 4.6 yaitu 0,612 > $\alpha = 0{,}05$ maka H_0 diterima.

Pemilihan Model Cox Terbaik

Tabel 7. Estimasi parameter model *Cox* terbaik dengan seleksi *backward*

Variabel	Koefsien	SE	$p - value$	$Exp(\beta_j)$
Umur (X_2)	-2,332	1,167	0,046	0,097
Kadar Gula Darah (X_6)	0,577	0,281	0,040	1,781

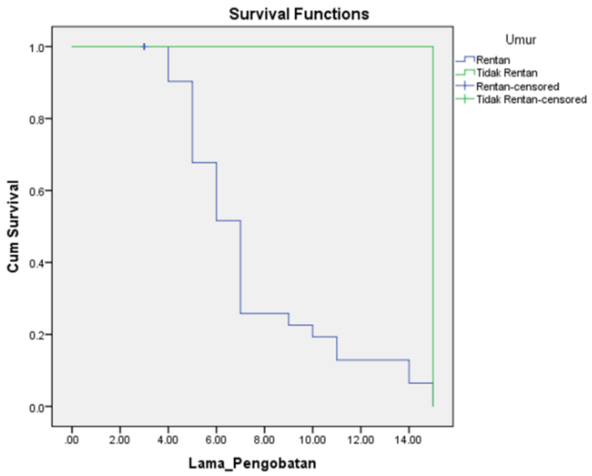
Berdasarkan hasil dari seleksi *backward* didapatkan dua variabel terpilih yang masuk dalam model terbaik *Cox* yaitu umur dan kadar gula darah. Tabel 5 menampilkan hasil estimasi parameter model terbaik *Cox* berdasarkan hasil seleksi *backward*

Model Regresi *Cox* berdasarkan hasil seleksi *backward* sebagai berikut:

$$(t, X) = h_0(t)\exp(-2{,}332 \ \text{Umur} + 0{,}577 \ \text{Kadar Gula Darah}) \tag{18}$$

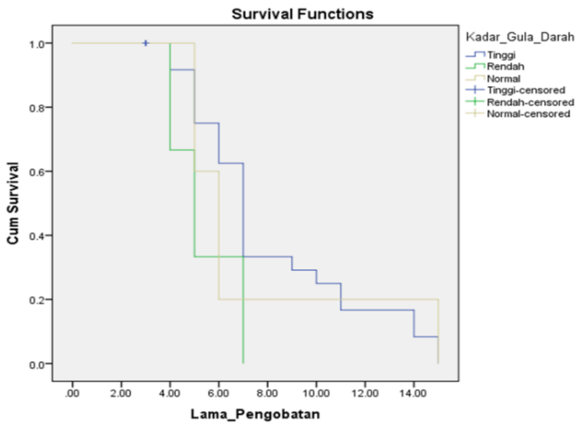
Pengujian Asumsi Proporsional *Hazard*

1. Kurva Kaplan Meier untuk variabel Umur



GAMBAR 2. Kurva *Kaplan Meier* untuk variabel Umur

2. Kurva Kaplan meier untuk variabel kadar gula darah



GAMBAR 3. Kurva *Kaplan Meier* untuk variabel Kadar Gula Darah

Berdasarkan pengujian asumsi *proportional hazard* dengan kurva *Kaplan Meier* terhadap waktu *survival* diperoleh kesimpulan bahwa variabel umur memenuhi asumsi *proportional hazard* sedangkan variabel kadar gula darah tidak memenuhi asumsi *proporsional hazard*

Model Cox *Extended*

Langkah-langkah pembentukan model cox *Extended* untuk mengatasi nonproporsional *hazard* pada kejadian bersama sebagai berikut:

1. Penambahan fungsi waktu $(t) = t$
3. Persamaan model Cox *Extended* dengan fungsi waktu $g(t) = t$ pada variabel umur (X_2) dan kadar gula darah (X_6), yaitu sebagai berikut:
4.

$$h(t, X) = h_0(t) exp[\beta_2 X_2 + \beta_6 X_6 + \delta_6 X_6 g(t)] \quad (19)$$

5.
2. Estimasi Parameter Model Cox *Extended*

TABEL 8. Estimasi Parameter Model Cox *Extended*

Variabel	**Koefisien**	$\boldsymbol{Exp(\beta_j)}$	**SE**	$\boldsymbol{p-value}$
Umur (X_2)	-8,655	0,000	5,886	0,141
Kadar Gula Darah (X_6)	-1,411	0,244	1,693	0,405
Kadar Gula Darah $g(t)$	0,373	1,452	0,309	0,227

Berdasarkan hasil estimasi parameter model cox *extended* dalam tabel 4.8

$$h_i(t_i) = h_0(t)\exp\left(-8{,}655X_2 - 1{,}411X_6 + 0{,}373X_6 g(t)\right) \quad (20)$$

Untuk mengetahui variabel yang signifikan, maka dilakukan uji *partial likelihood ratio* dengan menggunakan seleksi backward sebagai berikut:

TABEL 9. Prosedur Seleksi Backward dalam Pemilihan Model Signifikan

		Koefisien	**Wald**	***p-value***	**Exp(B)**	**-2 Log Likelihood**	$\boldsymbol{G_i}$
Langkah 0	*Null*					173,545	
Langkah 1	Umur	-8,655	2,163	0.141	0,000	165,437	16,216
	Kadar Gula Darah	-1,411	0,695	0,405	0,244		
	Kadar Gula Darah $g(t)$	0,373	1,457	0,227	1,452		
Langkah 2	Umur	-4,212	5,881	0,015	0,015	166,337	1,8
	Kadar Gula Darah $g(t)$	0,120	5,809	0,016	1,128		

Berdasarkan hasil estimasi parameter Model Cox *Extended* menggunakan metode Breslow dan eliminasi bakcward diperoleh dua variabel signifikan yaitu umur dan kadar gula darah terikat waktu. Sehingga model Cox *Extended* yang terbentuk dari dua variabel signifikan dapat dilihat pada persamaan berikut

$$h_i(t_i) = h_0(t)\exp(-4{,}212X_2 + 0{,}120X_6 g(t)) \quad (21)$$

Dimana

X_2 : Umur

$X_6 g(t)$: Kadar Gula Darah Terikat Waktu

KESIMPULAN

Berdasarkan hasil penelitian yang telah dipaparkan sebelumnya, dapat disimpulkan bahwa:

1. Model Estimasi parameter dalam data waktu survival atau data lama rawat inap (t) pasien diabetes mellitus merupakan data yang mengikuti distribusi eksponensial dua parameter yang ditentukan menggunakan metode momen. Diperoleh nilai estimasi parameter $\hat{\theta}$ = 3,543 dan $\hat{\eta}$ = 3,771
2. Pada pemodelan regresi Cox diperoleh model Cox Extended dengan variabel terikat adalah lama rawat inap di rumah sakit sebagai waktu survival, yang dipengaruhi oleh variabel bebas yaitu umur, pendidikan, pekerjaan, tekanan darah, kadar gula darah, kadar gula darah terikat waktu dan riwayat DM keluarga. Dari hasil analisis diperoleh model Cox Extended dengan g(t)=t sebagai berikut:

$$h_i(t_i) = \frac{\frac{1}{3{,}543}\exp\left(\frac{3{,}771-t}{3{,}543}\right)}{3{,}543\ \exp\left(\frac{3{,}771}{3{,}543}\right)+\exp\left(\frac{3{,}771-t}{3{,}543}\right)}\ \exp(-4{,}212X_2 + 0{,}120X_6 g(t))$$

   Berdasarkan model Cox *Extended* tersebut dapat disimpulkan bahwa model terbaik adalah model yang melibatkan dua variabel bebas yaitu umur dan kadar gula darah terikat waktu yang mengindikasikan bahwa:

   a. Koefisien Umur sebesar -4,212 bernilai negatif menunjukkan bahwa pasien yang berumur kurang dari 45 tahun memiliki resiko kegagalan 0,015 kali lebih kecil dibandingkan dengan pasien yang berumur lebih dari 45 tahun. Sehingga dapat disimpulkan bahwa semakin muda umur pasien maka waktu ketahanan hidupnya semakin lama
   b. Koefisien kadar gula darah terikat waktu 0,120 bernilai positif menunjukkan bahwa pasien yang kadar gula darahnya tinggi memiliki resiko kegagalan sebesar 1,128 kali lebih besar dibandingkan dengan pasien yang memiliki kadar gula darah rendah dan normal. Sehingga dapat disimpulkan bahwa pasien yang memiliki kadar gula rendah dan normal waktu ketahanan hidupnya semakin lama.
3. Berdasarkan model Cox Extended tersebut dapat disimpulkan bahwa model terbaik adalah model yang melibatkan dua variabel bebas yaitu umur dan kadar gula darah terikat waktu, sehingga faktor-faktor signifikan yang mempengaruhi waktu ketahanan hidup pasien penderita diabetes mellitus adalah
   a. Variabel umur (X_2) berpengaruh signifikan terhadap ketahanan hidup pasien Diabetes Mellitus yang ditunjukkan dengan nilai signifikansi *p-value* = 0,015 yang kurang dari 0,05
   b. Variabel kadar gula darah terikat waktu $(X_6)g(t)$ berpengaruh signifikan terhadap ketahanan hidup pasien Diabetes Mellitus yang ditunjukkan dengan nilai signifikansi *p-value* = 0,016 yang kurang dari 0,05.

DAFTAR PUSTAKA

Afifah, A., & Purnami, S. (2016). Uji Proportional Hazard pada Data Penderita Kanker Serviks di RSUD dr. Soetomo Surabaya. *Jurnal Sains dan Seni ITS Vol.5, No.1*, Institus Teknologi Sepuluh Nopember (ITS).

Hanni, T., & Wryandari, T. (2013). Model Regresi Cox Proporsional Hazard pada Data Ketahanan Hidup. *Media Statistika, Vol.6, No.1*, 11-20.

Iskandar, B. (2015). *Model Cox Proportional Hazard pada Kejadian Bersama.* Yogykarta: Universitas Negeri Yogyakarta.

Mandini, G. (2015). *Analisis Tahan Hidup Penderita Kanker Paru dengan Metode Kaplan-Maier.* Universitas Negeri Yogyakarta.

Masfufah, Hadju, V., & Djafar, N. (2014). *Pengetahuan Kadar Glukosa Darah dan Kualitas Hidup Penderita Diabetes Mellitus Tipe 2 Rawat Jalan di Wilayah Kerja Puskesmas Kota Makassar.* Makassar: Universitas Hasanuddin.

Rahayu, N., Setiawan, A., & Mahatma, T. (2012). Analisis Regresi Cox Proportional Hazard pada Ketahanan Hidup Pasien Diabetes Mellitus. *Seminar Nasional Matematika.* Salatiga: Universitas Kristen Satya Wacana.

Ridwan. (2016). *Analisis Survival dengan Penekatan Regresi Cox pada Kasus Demam Berdara Dengue (DBD)di Rumah Sakit Haji Labuang Baji Makassar.* Makassar: Universitas Negeri Makassar.

Saefuddun, A., & Ratnaningsih, D. (2008). Pemodelan Daya Tahan Mahasiswa Putus Kuliah paa Pendidikan Tinggi Jarak jau. *Statistika, vol.8 No.1*, 1-12.

Vitriana, A., & R, K. (2016). Model Cox Extended dengan g(t)=t untuk Mengatasi Nonproportional Hazard pada Kejadian Bersama. *Seminar Nasional Matematika dan Pendidikan Matematika.* UNY.