Model Survival Nonparametrik Pada Data Rawat Inap Pasien Diare di Puskesmas Indralaya

Ali Amran
Jurusan Matematika, Fakultas MIPA, Universitas Sriwijaya
e-mail: aliamran42@yahoo.com

Alfensi Faruk
Jurusan Matematika, Fakultas MIPA, Universitas Sriwijaya
e-mail: alfensifaruk@unsri.ac.id

Abstract: This research aimed to (1) estimate the survival functions, and (2) investigate the effect of some covariates toward the survival time from hospitalization time data in Public Health Center of Indralaya. The research subjects consisted of all the diarrhea patients who were hospitalized within the period from January 1, 2014 to June 30, 2014. The patient's characteristics were sex, age, job, and disease status. The result showed that the lowest probability of a patient will be out of hospitalization was on the fourth day and the highest probability was on the fifteenth day. In order to find the best model, some statistical tests were also conducted in this research.

Keywords: *Survival Model, Time of Hospitalization, Cox Regression*

1. Pendahuluan

Salah satu tujuan utama dalam analisis regresi adalah mempelajari bagaimana pengaruh dari satu atau lebih variabel bebas (kovariat) terhadap variabel terikat (variabel respon). Variabel respon dimodelkan sebagai suatu fungsi dari kovariat, koefisien (parameter regresi), dan *random error*. Apabila hubungan antara variabel respon dengan parameter adalah linear (dalam parameter), maka disebut sebagai model regresi linear. Akan tetapi, seringkali diperoleh bahwa hubungan antara variabel respon dengan parameter adalah nonlinear (dalam parameter), sehingga disebut sebagai model regresi nonlinear. Salah satu contoh model regresi nonlinear adalah model proporsional *hazard* Cox atau model regresi Cox [1], yang merupakan model yang digunakan untuk mengetahui bagaimana pengaruh dari satu atau lebih kovariat terhadap waktu *survival*.

Studi tentang penggunaan model regresi Cox pada berbagai masalah nyata sangat banyak ditemukan. Sebagai contoh, Farewell [2] menerapkan regresi Cox pada data multi infeksi, kemudian Jiayi [3] mengaplikasikan regresi Cox pada data pasar saham, sedangkan Hidayat [4] menggunakan regresi Cox untuk mengetahui pengaruh beberapa karakteristik terhadap waktu lahirnya anak pertama. Penerapan analisis *survival* pada data waktu rawat inap pasien juga pernah dilakukan oleh Quesenberry *et. al.* [5], yang mengestimasi fungsi *survival* dari waktu rawat inap pasien yang terdiagnosa virus AIDS.

Amran dan Faruk [6] membandingkan kecocokan beberapa distribusi terhadap 30 data waktu rawat inap pasien diare di Puskesmas Indralaya, serta mengestimasi fungsi-fungsi *survival* berdasarkan distribusi tersebut. Namun, dalam penelitian tersebut belum dibahas pengaruh karakteristik-karakteristik pasien terhadap waktu rawat inap tersebut dan belum dibahas juga estimasi fungsi *survival* secara nonparametrik, khususnya metode Kaplan-Meier. Oleh karena itu, penelitian ini bertujuan untuk (1) mengestimasi fungsi *survival* berdasarkan pendekatan nonparametrik, dan (2) menginvestigasi pengaruh dari beberapa karakteristik terhadap waktu rawat inap pasien diare di Puskesmas Indralaya. Hasil penelitian ini diharapkan bermanfaat untuk menambah literatur tentang penerapan analisis *survival* pada data waktu rawat inap pasien, dan bagi pihak terkait dapat dijadikan sebagai salah satu landasan dalam membuat kebijakan terkait pengelolaan rawat inap pasien.

2. Tinjauan Pustaka

Fungsi-Fungsi dalam Analisis *Survival*

Misalkan T adalah suatu variabel acak positif kontinu yang melambangkan waktu hingga terjadinya suatu kejadian tertentu atau dapat disebut sebagai waktu *survival*. Jika realisasi dari T adalah waktu amatan (*observed time*) yang dilambangkan dengan $t \geq 0$, maka fungsi kepadatan peluang (fkp) dari T, $f(t)$, dapat dituliskan sebagai

$$f(t) = \lim_{\Delta t \to 0} \frac{\Pr(t < T < t + \Delta t)}{\Delta t}. \qquad (1)$$

Selanjutnya, dapat diperoleh fungsi distribusi kumulatif (fdk) dari variabel acak T, $F(t)$, yang didefinisikan sebagai peluang terjadinya suatu kejadian hingga atau pada waktu t yang dituliskan sebagai

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

sedangkan fungsi *survival*, $S(t)$, yang didefinisikan sebagai peluang terjadinya peristiwa tertentu lebih dari waktu t dengan syarat bahwa peristiwa tersebut belum terjadi hingga waktu t adalah

$$S(t) = P(T > t) = 1 - F(t), \qquad (3)$$

dalam hal ini fungsi *survival* $S(t)$ adalah fungsi kontinu yang monoton turun dengan sifat

$$S(t) = \begin{cases} 1, & untuk\ t = 0 \\ 0, & untuk\ t = \infty \end{cases}.$$

Selain ketiga fungsi-fungsi di atas, terdapat satu fungsi penting lainnya, yaitu fungsi *hazard* yang dilambangkan dengan $h(t)$. Fungsi *hazard* didefinisikan sebagai peluang terjadinya suatu peristiwa sepanjang interval waktu yang sangat kecil, dengan diasumsikan bahwa peristiwa tersebut tidak pernah terjadi sejak awal pengamatan hingga waktu awal interval tersebut. Secara matematis, fungsi *hazard* didefinisikan sebagai berikut

$$h(t) = \frac{\lim_{\Delta t \to 0} P\left[\frac{\Pr(t \le T < t + \Delta t | T \ge t)}{\Delta t}\right]}{\Delta t}$$
$$= \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t(1 - F(t))} = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)}, \quad (4)$$

kemudian berdasarkan persamaan (2) diperoleh

$$f(t) = \frac{d}{dt} F(t) = \frac{d}{dt} [1 - S(t)] = -\frac{d}{dt} S(t),$$

sehingga fungsi *hazard* (4) dapat juga dituliskan dalam bentuk

$$h(t) = \frac{f(t)}{S(t)} = \frac{-\frac{d}{dt}S(t)}{S(t)} = -\frac{d}{dt}[\ln S(t)]. \quad (5)$$

Bentuk Umum Regresi Cox

Jika **X** adalah vektor berukuran $p \times 1$ yang elemen-elemennya adalah kovariat-kovariat $X_1, X_2, \dots, X_p$, maka bentuk umum model regresi Cox adalah

$$h(t|\mathbf{X}) = h_0(t|\mathbf{X}) \exp(\boldsymbol{\beta}^t \mathbf{X}) , \quad (6)$$

dimana $h(t|\mathbf{X})$ adalah fungsi *hazard* dengan diberikan p buah kovariat, $h_0(t|\mathbf{X})$ adalah fungsi *hazard* dasar yang menentukan bentuk dari fungsi *survival*, dan $\boldsymbol{\beta}$ adalah vektor berukuran $p \times 1$ yang elemennya adalah parameter atau koefisien-koefisien regresi.

3. Metode Penelitian

Data *survival* yang digunakan dalam penelitian ini diperoleh dari Puskesmas Indralaya, Kabupaten Ogan Ilir, Provinsi Sumatera Selatan. Periode pengamatan dimulai pada tanggal 1 Januari 2014 hingga 30 Juni 2014. Terdapat 30 orang pasien diare rawat inap dalam periode tersebut, yang semuanya dijadikan sebagai subjek penelitian. Tidak terdapat data tersensor dalam penelitian ini, karena hingga batas akhir observasi seluruh pasien sudah keluar dari rawat inap, sehingga studi ini juga dapat dikategorikan sebagai analisis *survival* pada data lengkap. Karakteristik-karakteristik pasien yang dianalisa adalah jenis kelamin (X_1), usia (X_2), pekerjaan (X_3), dan status penyakit (X_4). Untuk mengestimasi fungsi *survival* dan fungsi *hazard*, digunakan metode Kaplan-Meier, sedangkan paramater-paremeter dalam regresi Cox diestimasi menggunakan metode maksimum *likelihood* parsial. Pengujian signifikansi dari estimator-estimator yang diperoleh dilakukan baik secara simultan maupun parsial, sedangkan penentuan model terbaik menggunakan metode eliminasi mundur.

4. Hasil dan Pembahasan

Deskripsi Data

Berdasarkan tabel 1, diperoleh sebanyak 30 data pasien dengan empat karakteristik. Sebanyak 18 pasien berjenis kelamin laki-laki dan 12 pasien adalah perempuan. Usia pasien cukup beragam mulai dari yang paling muda dengan usia 0,25 tahun hingga yang paling tua yang berusia 72 tahun. Untuk pekerjaan pasien, terdapat 2 pasien yang bekerja di sektor swasta, 2 pasien bekerja sebagai pegawai negeri sipil (PNS), 4 pasien bekerja sebagai petani, 15 pasien masih turut orang tua (TOT), 4 pasien ibu rumah tangga (IRT), dan 3 pasien tidak bekerja. Ada beberapa status dari penyakit diare yang diderita pasien pada saat pertama kali dirawat inap, yaitu *gead* sebanyak 18 pasien, *gead* sedang sebanyak 2 pasien, *gead* sedang-berat dan vomitas+*gead* masing-masing sebanyak 4 pasien, serta *gead*+morbivi dan *gead* (disentri) yang masing–masing sebanyak 1 pasien.

Berdasarkan definisi dari waktu *survival*, waktu rawat inap pasien diare dalam tabel 1 merupakan waktu *survival* karena memiliki waktu awal pengamatan (1 Januari 2014) dan waktu akhir pengamatan (30 Juni 2014), serta mempunyai peristiwa yang diamati yaitu pasien keluar dari rawat inap di Puskesmas tersebut. Dalam periode pengamatan ini tidak terdapat data tersensor karena semua pasien mengalami kejadian sampai batas akhir waktu pengamatan, walaupun pada awalnya desain studi yang digunakan adalah penyensoran kanan campuran yang memperbolehkan semua subjek penelitian masuk ke dalam studi ketika pengamatan sedang berlangsung (Gijbels, [7]).

Estimasi Fungsi *Survival* Menggunakan Metode Kaplan-Meier

Dalam metode Kaplan-Meier, waktu amatan t_i dengan $i = 1,2,\ldots,k$ terlebih dahulu diurutkan dari waktu terkecil ke yang terbesar, sehingga diperoleh barisan waktu terurut $0 < t_1 < t_2 < \cdots < t_k$. Apabila banyaknya individu dengan waktu amatan t_i adalah d_i dan Y_i melambangkan banyaknya individu yang mengalami kejadian pada waktu t_i atau setelahnya, maka estimasi Kaplan-Meier dari fungsi *survival* $S(t)$ diberikan oleh

$$\hat{S}(t) = \begin{cases} 1, \text{ jika } t < t_1 \\ \prod_{t_i \le t} \left[1 - \frac{d_i}{Y_i}\right], \text{ jika } t_1 \le t\ . \end{cases} \tag{7}$$

Perhitungan estimasi fungsi *survival* $S(t)$ berdasarkan data *survival* yang terdapat di dalam tabel 1 dibantu dengan *software* Minitab 16. Hasil perhitungan tersebut dituliskan dalam tabel 2, sedangkan hasil estimasi dari kurva *survival* dan kurva *hazard* berturut-turut ditampilkan oleh gambar 1 dan gambar 2.

Tabel 1. Data 30 Pasien Diare Rawat Inap di Puskesmas Indralaya

Pasien ke-	Jenis Kelamin *)	Usia (Tahun)	Pekerjaan **)	Status	Durasi Rawat Inap (Hari)
1	L	31	Swasta	Gead	4
2	P	35	Tani	Gead	7
3	P	1,83	TOT	Vomitas + Gead	9
4	P	1,58	TOT	Vomitas + Gead	8
5	L	29	Tani	Gead	6
6	L	72	Tidak Bekerja	Gead Sedang-Berat	10
7	L	0,25	TOT	Gead Sedang	9
8	L	71	Tidak Bekerja	Gead	5
9	P	65	Tidak Bekerja	Gead Sedang-Berat	9
10	L	4	TOT	Vomitas + Gead	7
11	P	50	IRT	Gead	4
12	L	8	TOT	Gead	4
13	L	19	TOT	Gead Sedang-Berat	13
14	P	38	IRT	Gead Sedang	8
15	L	19	TOT	Vomitasi + Gead	7
16	P	1,08	TOT	Gead + Morbivi	5
17	P	55	IRT	Gead	7
18	L	18	TOT	Gead	6
19	P	39	IRT	Gead	5
20	L	38	Tani	Gead (Disentri)	11
21	L	54	PNS	Gead	7
22	P	0,67	TOT	Gead	5
23	L	12	TOT	Gead	9
24	L	1,17	TOT	Gead	7
25	L	0,67	TOT	Gead	12
26	L	63	Tani	Gead	7
27	P	18	TOT	Gead	6
28	P	4	TOT	Gead	13
29	L	42	PNS	Gead Sedang-Berat	15
30	L	70	Swasta	Gead	13

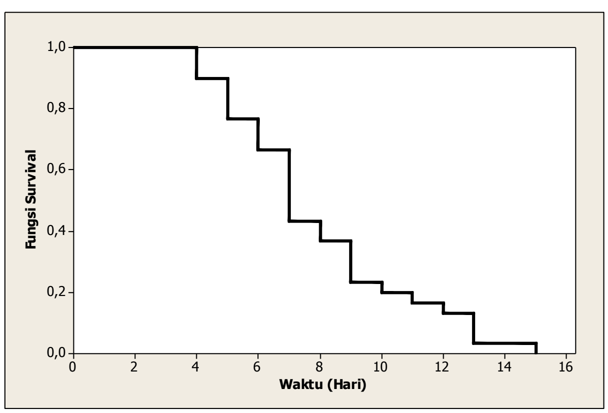
*) L: laki-laki; P: perempuan

**) TOT: Turut Orang Tua; IRT: Ibu Rumah Tangga; PNS: Pegawai Negeri Sipil

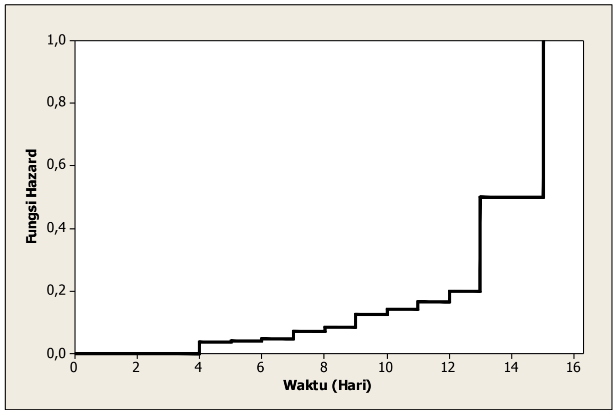
Tabel 2. Hasil Estimasi Fungsi *Survival* dan Fungsi *Hazard*

i	Waktu (t_i)	Y_i	d_i	Fungsi *Survival* $\hat{S}(t_i)$	Fungsi *Hazard* $\hat{h}(t_i)$
1	4	30	3	0.9000	0.0357
2	5	27	4	0.7667	0,0417
3	6	23	3	0.6667	0,0476
4	7	20	7	0.4333	0,0714
5	8	13	2	0.3667	0,08333
6	9	11	4	0.2333	0,125
7	10	7	1	0.2000	0,1429
8	11	6	1	0.1667	0,1667
9	12	5	1	0.1333	0,2
10	13	4	3	0.0333	0,5
11	15	1	1	0.0000	1

Nilai estimasi fungsi *survival* $\hat{S}(t_i)$ dalam tabel 2 dapat didefinisikan sebagai besarnya peluang seorang pasien diare rawat inap untuk tetap dirawat inap di Puskemas Indralaya lebih dari hari ke-t_i. Sebagai contoh, peluang seorang pasien diare untuk tetap dirawat inap setelah hari ke-4 dilambangkan dengan $\hat{S}(t_1) = \hat{S}(4)$ yang nilainya adalah 0,9. Terlihat bahwa nilai estimasi peluang *survival* yang ditampilkan dalam tabel 2 semakin menurun dan bernilai nol pada saat $t_{11} = 15$. Hal ini diperlihatkan juga oleh gambar 1, dimana estimasi kurva *survival* yang diperoleh berupa fungsi tangga yang monoton turun. Estimasi fungsi *hazard* $\hat{h}(t_i)$ dapat diinterpretasikan sebagai peluang keluarnya pasien diare dari rawat inap di Puskesmas Indralaya pada hari ke-t_i. Berdasarkan tabel 2, terlihat bahwa peluang terendah berada pada hari ke-4, yaitu $\hat{h}(4) = 0.0357$, sedangkan peluang tertinggi berada pada hari ke-15, yaitu $\hat{h}(15) = 1$.



Gambar 1. Grafik Estimasi Fungsi *Survival*



Gambar 2. Grafik Estimasi Fungsi *Hazard*

Estimasi Koefisien Regresi Cox Menggunakan Metode *Likelihood* Parsial

Dalam model regresi, terdapat dua metode standar yang biasanya digunakan untuk mengestimasi koefisien regresinya, yaitu metode kudrat terkecil (*least square*) dan metode *maximum likelihood estimation* (MLE). Sedangkan, untuk mengestimasi koefisien regresi Cox dapat dilakukan menggunakan metode *partial likelihood* (Cox, [1]). Metode ini merupakan pengembangan dari metode MLE, yang mana persamaan *likelihood* dalam MLE diganti dengan persamaan *likelihood* parsial. Namun, tujuan yang ingin dicapai masih tetap sama yaitu mendapatkan estimator yang memaksimumkan persamaan *likelihood*.

Sama seperti metode Kaplan-Meier, langkah awal dalam metode *likelihood* parsial adalah mengurutkan waktu amatan t_i, $i = 1,2,\dots,k$, sedemikian sehingga diperoleh barisan waktu terurut $0 < t_1 < t_2 < \cdots < t_k$. Dimisalkan $\mathbf{X}_i$ adalah vektor dari jumlahan setiap p kovariat untuk subjek-subjek yang mengalami peristiwa pada waktu ke-i, t_i. Jika terdapat d_i kejadian pada waktu t_i, maka elemen ke-h dari $\mathbf{X}_i$ adalah $X_{hi} = \sum_{r=1}^{d_i} x_{hir}$, dimana x_{hir} adalah nilai dari kovariat ke-h , $h = 1,2,\dots,p$, untuk subjek ke-r dari d_i, $r = 1,2,\dots,d_i$, yang mengalami peristiwa pada waktu ke-i, $i = 1,2,\dots,k$. Menurut Breslow [8], aproksimasi fungsi *likelihood* parsial dalam kasus yang memungkinkan terdapat lebih dari satu subjek yang mengalami peristiwa pada waktu amatan t_i diberikan oleh persamaan berikut

$$L_p(\boldsymbol{\beta}) = \prod_{i=1}^{k} \left(\frac{exp(\boldsymbol{\beta}^{\mathbf{t}} \mathbf{X}_i)}{\left\{ \sum_{j \in Y_i} exp(\boldsymbol{\beta}^{\mathbf{t}} \mathbf{X}_j) \right\}^{d_j}} \right)^{\delta_i}, \qquad (8)$$

dimana δ_i adalah variabel indikator yang bernilai 1 jika peristiwa terjadi dan bernilai 0 jika peristiwa tersebut tersensor. Collet [9] menyatakan bahwa komputasi dari fungsi *likelihood* (8) relatif mudah dan merupakan pendekatan yang cukup baik ketika jumlah kejadian yang sama pada setiap waktu amatan t_i tidak terlalu besar. Berdasarkan tabel 2, nilai dari setiap d_i tidak terlalu besar sehingga dalam penelitian ini fungsi *likelihood parsial* (8) cocok digunakan untuk mengestimasi model regresi Cox dari data dalam tabel 1.

Tujuan dari metode *likelihood* parsial adalah mengestimasi nilai-nilai parameter $\boldsymbol{\beta}$ yang memaksimumkan fungsi *likelihood* parsial. Agar lebih mempermudah proses komputasi, persamaan (8) terlebih dahulu dilinearisasi dengan mengambil logaritmanya sehingga diperoleh

$$\log L_p(\boldsymbol{\beta}) = \sum_{\boldsymbol{i=1}}^{\boldsymbol{k}} \delta_i \left[\boldsymbol{\beta}^{\mathbf{t}}\mathbf{X}_i - d_j \log\left(\sum_{j\in Y_i} exp\left(\boldsymbol{\beta}^{\mathbf{t}}\mathbf{X}_j\right)\right)\right], \tag{9}$$

dalam hal ini estimator dari parameter-parameter $\boldsymbol{\beta}$ yang memaksimalkan persamaan *loglikelihood* (9), juga akan memaksimalkan persamaan *likelihood* (8).

Seperti yang telah dijelaskan sebelumnya, terdapat empat karakteristik pasien yang dianalisa yaitu jenis kelamin, usia, pekerjaan, dan status penyakit, dimana untuk selanjutnya karakteristik-karakteristik tersebut secara berturut-turut dinotasikan dengan variabel X_1, X_2, X_3 dan X_4. Kovariat X_1, X_3, dan X_4 adalah variabel bebas bertipe kategorik, sedangkan X_2 adalah variabel bebas bertipe kontinu. Kovariat X_1 memiliki dua kategori, yaitu dilambangkan perempuan dan laki-laki. Kovariat X_3 terdiri atas 6 kategori, yang selanjutnya dibentuk 5 variabel *dummy* yaitu X_{3a}= swasta, X_{3b}= tani, X_{3c}= turut orang tua, X_{3d} = ibu rumah tangga, dan X_{3e} = pegawai negeri sipil, yang mana kategori acuannya adalah tidak bekerja. Selanjutnya, kovariat X_4 juga terdiri dari 6 kategori, dan dibentuk 5 variabel *dummy* yaitu X_{4a}= vomitas + *gead*, X_{4b}= *gead* sedang, X_{4c}= *gead* sedang-berat, dan X_{4d}= *gead* + morbivi, dan X_{4e}= *gead* disentri, dengan kategori acuannya adalah *gead*. Jadi, dalam perhitungannya kovariat dalam penelitian ini terdiri atas 13 buah kovariat.

Selanjutnya, untuk mengestimasi koefisien-keofisien dari model regresi Cox digunakan fungsi *likelihood* parsial (9) dan pendekatan Breslow. Pengujian kontribusi simultan dan parsial, berurut-turut dilakukan menggunakan statistik uji χ^2 dan statistik uji Wald, sedangkan dalam pemilihan model terbaik digunakan metode *backward elimination*.

Pengujian Kontribusi Variabel Bebas Secara Simultan

Pengujian secara simultan adalah pengujian dimana semua variabel bebas (kovariat) dianggap berpengaruh secara signifikan. Statistik ujinya, adalah χ^2 yang berbentuk

$$\chi^2 = -2(\ln L_F - \ln L_R), \tag{10}$$

dimana L_F adalah fungsi *likelihood* dari model kovariat lengkap, dan L_R adalah fungsi *likelihood* dari model kovariat yang tidak lengkap. Hipotesis uji yang digunakan adalah

$H_0: \beta_i = 0, \forall_i; i = 1,2, \dots, p$
$H_0: \exists \beta_i \neq 0\ ; i = 1,2, \dots, p$,

dengan β_i adalah koefisien regresi dari kovariat ke-i dari sejumlah p buah kovariat. Dalam hal ini, hipotesis awal ditolak apabila nilai $\chi^2 > \chi^2_{db;(1-\alpha)}$ atau p-*value* < α. Dalam penelitian ini, nilai taraf signifikansi (α) yang digunakan adalah 0,05 atau 5%. Berdasarkan perhitungan menggunakan *software* SPSS 19, diperoleh nilai p$-value =$ $0{,}224$ > $\alpha = 0{,}05$ sehingga dapat disimpulkan bahwa hipotesis awal diterima , atau dengan kata lain jika semua semua kovariat dimasukkan secara bersama-sama ke dalam model, maka pengaruh dari semua kovariat tersebut tidak signifikan terhadap waktu *survival*.

Pengujian Kontribusi Variabel Bebas Secara Parsial

Pengujian secara parsial bertujuan untuk menguji pengaruh dari setiap kovariat dengan dianggap bahwa pengaruh dari kovariat-kovariat yang lain tidak ada. Statistik uji yang digunakan dalam pengujian parsial ini adalah statistik uji Wald yang berbentuk

$$W = \left(\frac{\widehat{\beta_i}}{se(\widehat{\beta_i})}\right)^2, \tag{11}$$

dimana $\widehat{\beta_i}$ = estimator dari parameter β_i , dan $se(\widehat{\beta_i})$ adalah *standar error* dari estimator $\widehat{\beta_i}$. Adapun, hipotesisnya berbentuk

$H_0: \beta_i = 0, \forall_i; i = 1,2, \dots, p$
$H_0: \beta_i \neq 0\ ; i = 1,2, \dots, p$,

dimana p adalah jumlah kovariat. H_0 ditolak apabila nilai $W > Z\alpha/_2$ atau p-*value* < α. Pengujian parsial dalam penelitian ini dibantu dengan *software* SPSS 19, yang mana hasil perhitungannya dituliskan seara ringkas dalam tabel 3.

Berdasarkan hasil yang terlihat dalam tabel 3, hampir semua kesimpulan dari pengujian adalah menerima H_0 pada taraf signifikansi sebesar 5%, kecuali pengujian untuk koefisien dari pekerjaan swasta (X_{3a}) dan status *gead* sedang-berat (X_{4c}), yang disimpulkan bahwa H_0 ditolak. Jadi, berdasarkan pengujian kontribusi secara parsial dengan taraf signifikansi sebesar 5%, hanya dua kovariat saja yang memiliki pengaruh signifikan terhadap waktu *survival*, yaitu pekerjaan swasta (X_{3a}) dan status penyakit *gead* sedang-berat (X_{4c}).

Tabel 3. Hasil Pengujian Kontribusi Kovariat Secara Parsial

Kovariat	$\widehat{\beta_i}$	Skor Wald	p-value	Kesimpulan H_0 ($\alpha = 0{,}05$)
Jenis Kelamin (X_1)	-0,674	1,550	0,213	Diterima
Usia (X_2)	-0,023	0,653	0,419	Diterima
Pekerjaan (X_3)		7,887	0,163	Diterima
Swasta (X_{3a})	-2,861	5,115	0,024	Ditolak
Tani (X_{3b})	-1,583	1,538	0,215	Diterima
Turut Orang Tua (X_{3c})	-3,272	2,573	0,109	Diterima
Ibu Rumah Tangga (X_{3d})	-0,517	0,196	0,658	Diterima
Pegawai Negeri Sipil (X_{3e})	-2,612	3,458	0,063	Diterima
Status Penyakit (X_4)		9,899	0,078	Diterima
Vomitas + gead (X_{4a})	0,393	0,349	0,555	Diterima
Gead sedang (X_{4b})	-1,013	1,297	0,255	Diterima
Gead sedang-berat (X_{4c})	-2,151	5,502	0,019	Ditolak
Gead + morbivi (X_{4d})	2,309	3,454	0,063	Diterima
Gead disentri (X_{4e})	-1,7	2,060	0,151	Diterima

Pemilihan Model Terbaik

Untuk memilih model regresi Cox terbaik, dapat dilakukan menggunakan metode eliminasi mundur *Likelihood Ratio* (LR). Prosedur dan perhitungan dalam pemilihan model terbaik ini dibantu oleh *software* SPSS 19, yang hasilnya ditampilkan dalam tabel 4.

Tabel 4. Prosedur Pemilihan Model Terbaik

Langkah	Jenis Kelamin	Usia	Pekerjaan	Status Penyakit	*p-value*	Ket.
1	√	√	√	√	0,224	Model 1
2	√		√	√	0,197	Model 2
3			√	√	0,197	Model 3
4				√	0,210	Model 4

Berdasarkan hasil yang diperlihatkan dalam tabel 4, tidak ada nilai *p-value* yang lebih lebih kecil dari nilai taraf signifikansi (α) sebesar 0,05. Oleh karena itu, semua kovariat dari model yang ditampilkan, yaitu dari model 1 sampai model 4, tidak ada yang signifikan secara simultan berpengaruh terhadap waktu *survival*, sehingga berdasarkan perhitungan terhadap data yang diperoleh dapat disimpulkan bahwa tidak ada model terbaik yang dapat dibentuk oleh kovariat-kovariat X_1, X_2, X_3 dan X_4.

5. Kesimpulan

Berdasarkan perhitungan yang telah dilakukan, diperoleh estimasi fungsi *survival* $\hat{S}(t)$ dan fungsi *hazard* $\hat{h}(t)$ dari data rawat inap pasien diare di Puskesmas Indralaya. $\hat{S}(t)$ adalah peluang seorang pasien diare masih tetap dirawat inap lebih dari hari ke-t, sedangkan $\hat{h}(t)$ adalah peluang seorang pasien diare keluar dari rawat rawat inap pada hari ke-t. Peluang pasien diare keluar dari rawat inap terendah adalah pada hari ke-4, yaitu $\hat{h}(4) = 0.0357$, sedangkan peluang tertinggi berada pada hari ke-15, yaitu $\hat{h}(15) = 1$.

Terdapat dua jenis pengujian yang dilakukan dalam penelitian ini, yaitu pengujian kontribusi kovariat secara simultan dan pengujian secara parsial. Berdasarkan uji simultan, yaitu jika semua kovariat dimasukkan secara bersama-sama maka pengaruh terhadap waktu *survival* dari semua kovariat tidak signifikan. Sedangkan, pada uji parsial diperoleh bahwa ketika kovariat lainnya dianggap tidak berpengaruh maka hanya kovariat jenis pekerjaan swasta dan status penyakit *gead* sedang-berat yang berpengaruh signifikan terhadap waktu *survival*. Menggunakan metode eliminasi mundur LR, disimpulkan juga bahwa dalam kasus ini tidak diperoleh model terbaik yang dapat dibentuk dari kovariat X_1, X_2, X_3 dan X_4.

Ucapan Terimakasih

Terimakasih sebesar-besarnya diucapkan kepada Lembaga Penelitian (Lemlit) Universitas Sriwijaya, yang telah membiayai penelitian ini melalui skim penelitian Sains dan Teknologi untuk tahun anggaran 2015, sehingga penelitian ini dapat berjalan dengan baik dan diselesaikan tepat pada waktunya.

Daftar Pustaka

[1] Cox, D. R. 1972. Regression Models and Life Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2), 187-220

[2] Farewell, V. T. 1979. An Application of Cox's Proportional Hazard Model to Multiple Infection Data. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28 (2), 136-143

[3] Ni, J. 2009. Application of Cox Proportional Hazard Model to the Stock Exchange Market. *B.S. Undergraduate Mathematics Exchange*, 6(1), 12-18

[4] Hidayat, R., Sumarno, H., dan Nugrahani, E., H. 2014. Survival Analysis in Modeling the Birth Interval of the First Child in Indonesia. *Open Journal of Statistics*, 2014 (4) 198-206

[5] Quesenberry, P., C., Fireman, B., Hiatt, A., R., Selby, V., J. 1989. A survival analysis of hospitalization among patients with Acquired Immunodeficiency Syndrome. *AJPH* , 79 (12), 1643-1647

[6] Amran, A., Faruk, A. 2015. Perbandingan Distribusi Eksponensial, Weibull, dan Lognormal dalam Mengestimasi Model *Survival* Waktu Rawat Inap Pasien Diare di Puskesmas Indralaya. *Seminar Nasional Matematika dan Pendidikan Matematika*, IndoMS Wilayah Sumbagteng

[7] Gijbels, I. 2010. Censored Data. *WIREs Computational Statistics*, 2, 178-188

[8] Breslow, N. 1974. Covariance Analysis of Censored Survival Data. *Biometrics*, 30(1), 89-99

[9] Collet, D. 2003. *Modelling Survival Data in Medical Research Second Edition.* New York: Chapman & Hall/ CRC

[10] Kleinbaum, D. G., Klein, M. 2005. *Survival Analysis A Self-Learning Text Second Edition*. New York: Springer