

Aplikasi Regresi Cox Pada Selang Kelahiran Anak Pertama di Provinsi Sumatera Selatan

Alfansi Faruk

Jurusan Matematika, FMIPA, Universitas Sriwijaya

Email: alfensifaruk@unsri.ac.id

Abstract: The birth interval of first child is one example of survival analysis. The objective of this research is to determine the best Cox regression model that describes the impact of several factors to the birth interval of first child in South Sumatra Province. The data were based on the Indonesia Demographic and Health Survey (IDHS) 2012. In this work, we included four factors, i.e. the age of the mother, mother's highest education, place of residence, and wealth index. Based on several tests that have been conducted, the factors which significantly affect the survival time were the age of the mother, mother's highest education (junior high school), and wealth index (middle and very rich).

Keywords: Cox regression, IDHS, The birth interval

1. Pendahuluan

Fertilitas (*fertility*) merupakan salah satu faktor utama yang memengaruhi kepadatan penduduk selain mortalitas (*mortality*) dan migrasi (*migration*). Mantra [1] mendefinisikan bahwa fertilitas adalah suatu peristiwa terlepasnya bayi dari rahim seorang perempuan dengan adanya tanda-tanda kehidupan, misalnya menangis, jantung berdenyut, bernapas dan sebagainya. Tingkat fertilitas (*fertility rate*) penduduk di suatu daerah yang tinggi mengindikasikan bahwa penduduk di daerah tersebut memiliki angka kelahiran penduduk yang juga tinggi.

Selang kelahiran anak pertama didefinisikan sebagai waktu yang diperlukan oleh pasangan suami istri untuk mendapatkan anak pertama terhitung sejak hari pernikahan, dengan catatan bahwa pasangan tersebut melakukan hubungan suami istri secara rutin dan tidak melakukan berbagai tindakan yang bertujuan untuk menunda terjadinya kehamilan (kontrasepsi). Waktu hingga lahirnya anak pertama dari pasangan suami istri dapat menjadi salah satu indikator tingkat fertilitas [2], yang dalam hal ini semakin cepat pasangan tersebut dikaruniai anak pertama (semakin pendek selang kelahiran) maka tingkat fertilitas pasangan tersebut dapat dikatakan juga menjadi semakin tinggi. Dalam analisis *survival*, waktu hingga lahirnya anak pertama merupakan salah satu contoh dari data *survival* karena syarat-syarat agar dapat dikategorikan sebagai waktu *survival* telah terpenuhi [3].

Model proporsional *hazard* Cox atau model regresi Cox merupakan model *survival* yang banyak digunakan oleh para praktisi dalam menganalisis faktor yang

berpengaruh terhadap waktu *survival* [4], [5], [6]. Model regresi Cox memiliki keunggulan dalam hal interpretasi yang relatif mudah serta tidak memerlukan asumsi bahwa data *survival* mengikuti suatu distribusi tertentu seperti dalam model *survival* parametrik [7], [8]. Tujuan utama dari penelitian ini adalah menentukan faktor-faktor yang berpengaruh signifikan pada permasalahan selang kelahiran anak pertama di Provinsi Sumatera Selatan menggunakan model regresi Cox. Subjek penelitian adalah responden yang memberikan informasi mengenai kelahiran anak pertamanya dan bertempat tinggal di Provinsi Sumatera Selatan berdasarkan data hasil survei *Indonesia Demographic and Health Survey* (IDHS) tahun 2012.

2. Data dan Metode

Data yang digunakan pada penelitian ini adalah data hasil survei yang dilaksanakan atas kerjasama antara *United States Agency for International Development* (USAID) dengan pemerintah Indonesia melalui program IDHS pada tahun 2012. Tujuan IDHS ini adalah menyediakan informasi bagi para pembuat kebijakan tentang keadaan kependudukan, keluarga berencana, dan kesehatan dari penduduk Indonesia. Data selang kelahiran anak pertama di Provinsi Sumatera Selatan yang digunakan dalam penelitian ini berjumlah 97 responden, sedangkan karakteristik responden yang diduga memengaruhi selang kelahiran tersebut adalah usia, tempat tinggal, tingkat pendidikan, dan tingkat ekonomi (Tabel 1).

Tabel 1. Karakteristik Responden

Variabel	Penjelasan	Tipe Variabel	Keterangan
X_1	Usia Responden Ketika Melahirkan	Kontinu	Dinyatakan dalam tahun
X_2	Tempat Tinggal Responden Pada Saat Melahirkan	Kategorik	1: Perkotaan 2: Pedesaan
X_3	Pendidikan Terakhir Responden	Kategorik	1: Tidak berpendidikan 2: SD 3: SMP 4: SMA
X_4	Tingkat Ekonomi Responden	Kategorik	1: Sangat Miskin 2: Miskin 3: Menengah 4: Kaya 5: Sangat Kaya

Langkah-langkah dari metode penelitian yang dilaksanakan dalam penelitian ini adalah:

1. Mendeskripsikan data selang kelahiran anak pertama di Provinsi Sumatera Selatan berdasarkan data IDHS tahun 2012
2. Mengasumsikan bahwa sebagian data *survival* sebagai data tersensor kanan tipe I

3. Memilih faktor yang dilibatkan dalam studi dan membentuk variabel *dummy* untuk faktor bertipe kategorik
4. Melakukan uji simultan dan uji parsial pada semua faktor
5. Mengimplementasikan metode eliminasi mundur untuk memperoleh model terbaik
6. Menginterpretasikan hasil-hasil yang diperoleh.

3. Hasil dan Pembahasan

Analisis Survival

Analisis *survival* adalah salah satu cabang ilmu statistika yang membahas waktu terjadinya peristiwa tertentu dari suatu waktu awal (titik awal) ke suatu waktu akhir (titik akhir). Lee dan Wang [3] menyatakan bahwa waktu *survival* adalah waktu dari awal observasi hingga waktu terjadinya suatu peristiwa tertentu. Namun, pada kenyataannya tidak semua subjek penelitian dapat mengalami peristiwa yang diperhatikan dalam masa observasi, sehingga informasi yang diperoleh dari subjek tersebut menjadi tidak utuh. Data-data *survival* yang tidak memberikan informasi secara utuh tersebut disebut sebagai data tersensor [9].

Beberapa alasan terjadinya penyensoran pada data *survival* antara lain adalah subjek mengalami peristiwa di luar masa observasi, subjek menghilang pada masa observasi, subjek mengalami peristiwa lain sedemikian sehingga peristiwa yang menjadi perhatian tidak terjadi, dan subjek mengalami peristiwa yang menjadi perhatian namun disebabkan oleh alasan lain yang bukan menjadi perhatian peneliti. Karena di dalam data tersensor masih terdapat sebagian informasi yang dapat digunakan, yaitu waktu terjadinya penyensoran, maka data tersensor berbeda dengan data hilang (*missing data*) yang mana informasinya benar-benar tidak diketahui. Oleh karena itu, untuk menganalisis data *survival* dengan data tersensor diperlukan metode statistika khusus dan banyak prosedur standar dalam statistika yang tidak lagi tepat digunakan pada data *survival* [10].

Kepadatan Penduduk di Provinsi Sumatera Selatan

Sumatera Selatan adalah salah satu Provinsi di Indonesia yang terletak di bagian selatan Pulau Sumatra. Provinsi ini berbatasan langsung dengan Provinsi Lampung di sebelah selatan, Provinsi Jambi di sebelah utara, Provinsi Bengkulu di sebelah barat, serta Provinsi Kepulauan Bangka-Belitung di sebelah timur. Provinsi ini kaya dengan sumber daya alam, seperti minyak bumi, gas alam, dan batu bara. Provinsi Sumatera Selatan secara administratif terdiri atas tiga belas kabupaten dan empat kota, dengan Palembang sebagai ibu kota Provinsi.

Berdasarkan hasil sensus penduduk pada tahun 2010, jumlah penduduk Provinsi Sumatera Selatan adalah sebanyak 7.450.394 jiwa, yang menempatkan Sumatera

Selatan sebagai Provinsi dengan jumlah penduduk terpadat ke-10 dari total 33 Provinsi yang ada di Indonesia [11]. Diketahui juga bahwa laju pertumbuhan penduduk Provinsi Sumatera Selatan pada tahun 2010 tersebut sebesar 1,45% per tahun. Oleh karena itu, jika diproyeksikan dengan nilai laju pertumbuhan tersebut, maka diperkirakan bahwa pada akhir tahun 2016 jumlah penduduk di Provinsi Sumatera Selatan kurang lebih sebanyak 8.240.311 jiwa. Hal ini berarti bahwa dalam kurun waktu tahun 2010 hingga tahun 2016 penduduk Sumatera Selatan telah bertambah sebanyak 789.917 jiwa, sehingga dapat dikatakan bahwa Provinsi Sumatera Selatan adalah daerah dengan pertumbuhan penduduk yang tinggi.

Data Survival Selang Kelahiran Anak Pertama

Pada awalnya, data IDHS 2012 yang digunakan dalam penelitian ini terdiri dari 98 responden yang tinggal di Provinsi Sumatera Selatan yang memberikan informasi mengenai kelahiran anak pertama. Akan tetapi, karena terdapat satu orang responden yang telah melahirkan sebelum hari pernikahan, maka responden tersebut dikeluarkan dari studi, sehingga jumlah responden yang dilibatkan menjadi sebanyak 97 orang. Penghapusan data ini dilakukan karena waktu *survival* harus bernilai positif dan lebih besar dari nol. Hari pernikahan dijadikan sebagai waktu awal dan saat melahirkan anak pertama sebagai waktu akhir (dalam bulan).

Pasangan suami istri dikatakan mengalami infertilitas apabila sang istri belum hamil hingga satu tahun terhitung sejak hari pernikahan, dengan catatan bahwa pasangan suami istri tersebut tidak melakukan usaha penundaan kehamilan selama pernikahan. Oleh karena itu, dalam penelitian ini waktu observasi ditetapkan selama 21 bulan, yang diperoleh dari 1 tahun (12 bulan) ditambah dengan 9 bulan (lama kehamilan normal). Karena hampir semua responden menikah pada waktu yang berbeda, maka waktu *survival* dari setiap responden diasumsikan sebagai waktu *follow up* yang dimulai dari titik awal (awal bulan pertama) dan berakhir pada akhir bulan ke-21. Apabila terdapat responden yang melahirkan anak pertamanya lebih dari 21 bulan, maka data *survival* dari responden tersebut diasumsikan sebagai data tersensor kanan tipe I [3], [9].

Beberapa Fungsi Penting dalam Analisis Survival

Analisis *survival* secara garis besar dapat dibagi menjadi dua pokok bahasan, yaitu analisis *survival* univariat dan analisis *survival* multivariat [12], [13]. Apabila dalam analisis hanya difokuskan pada satu variabel saja, maka hal tersebut dikatakan sebagai analisis *survival* univariat. Sebaliknya, jika pembahasan dalam analisis *survival* tentang bagaimana hubungan antara dua variabel atau lebih, maka hal ini disebut sebagai analisis *survival* multivariat. Beberapa contoh pokok bahasan dalam analisis *survival*

univariat adalah estimasi fungsi *survival* dan fungsi *hazard*, sedangkan contoh pokok bahasan dari analisis *survival* multivariat adalah model regresi Cox [14].

Jika dimisalkan $T > 0$ adalah suatu variabel acak yang merepresentasikan waktu *survival*, maka fungsi *survival* yang didefinisikan sebagai peluang individu mengalami peristiwa pada waktu lebih dari t dapat diformulasikan sebagai

$$S(t) = \Pr[T > t] = 1 - \Pr[T \leq t] = 1 - F(t),$$

dimana waktu t adalah realisasi dari variabel acak T dan $F(t)$ adalah fungsi distribusi kumulatif pada waktu t . Fungsi penting lainnya adalah fungsi *hazard*, yang didefinisikan sebagai peluang terjadinya peristiwa yang diperhatikan dalam selang waktu yang sangat kecil dengan syarat bahwa peristiwa tersebut belum terjadi hingga waktu t . Formula matematis dari fungsi *hazard* pada waktu t adalah

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{\Pr[t \leq T < t + \Delta t | T > t]}{\Delta t},$$

selanjutnya berdasarkan definisi fungsi distribusi kumulatif, $F(t) = \Pr[T \leq t]$, dan definisi dari peluang bersyarat, $\Pr[A|B] = \Pr[A \cap B]/\Pr[B]$, maka $h(t)$ dapat ditulis menjadi

$$\begin{aligned} h(t) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \cdot \frac{\Pr[t \leq T < t + \Delta t] \cap \Pr[T > t]}{\Pr[T > t]} = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \cdot \frac{\Pr[t \leq T < t + \Delta t]}{\Pr[T > t]} \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \cdot \frac{F(t + \Delta t) - F(t)}{(1 - F(t))} = \frac{1}{S(t)} \cdot \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} = \frac{f(t)}{S(t)}, \end{aligned}$$

dengan $S(t) = 1 - F(t)$ adalah fungsi *survival* dan $f(t)$ adalah fungsi kepadatan peluang pada waktu t yang berbentuk

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

Dalam model regresi Cox, fungsi *hazard* $h(t)$ merupakan variabel terikat, sedangkan faktor-faktor yang diduga memengaruhi waktu *survival* adalah variabel bebas. Asumsi utama yang digunakan dalam model regresi Cox adalah adanya proporsionalitas fungsi *hazard* antar individu dalam populasi, artinya rasio fungsi *hazard* untuk setiap dua individu besarnya selalu konstan sepanjang waktu t . Bentuk umum model regresi Cox adalah

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

dengan \mathbf{X} adalah vektor dari faktor, $\boldsymbol{\beta}$ adalah vektor dari koefisien-koefisien regresi yang tidak diketahui, dan $h_0(t)$ adalah fungsi *hazard* dasar yang didefinisikan sebagai fungsi yang nilainya tidak diketahui yang memberikan nilai *hazard* ketika semua faktor bernilai nol.

Model Regresi Cox pada Data Selang Kelahiran Anak Pertama

Seperti yang diperlihatkan dalam Tabel 1, terdapat satu faktor bertipe kontinu dan tiga faktor bertipe kategorik. Dari ketiga faktor bertipe kategorik tersebut selanjutnya dibentuk variabel-variabel *dummy* (Tabel 2). Dalam penelitian ini, ditetapkan kategori acuan untuk karakteristik tempat tinggal (X_2) adalah perkotaan, sedangkan kategori acuan untuk pendidikan ibu (X_3) dan tingkat perekonomian ibu (X_4) berturut-turut adalah tidak berpendidikan dan tingkat perekonomian sangat miskin.

Langkah awal yang dilakukan dalam proses penentuan model regresi terbaik adalah dengan melakukan uji simultan, yaitu suatu uji statistika yang bertujuan untuk mengetahui bagaimana kontribusi variabel bebas secara bersama-sama dalam model. Statistik uji yang digunakan adalah statistik uji *Chi-Square* (χ^2) yang berbentuk

$$\chi^2 = -2(\ln L_{m-r} - \ln L_m),$$

dengan L_m dan L_{m-r} berturut-turut adalah fungsi *likelihood* dari m variabel dan $(m - r)$ variabel. Hipotesis yang digunakan dalam pengujian ini adalah $H_0: \beta_i = 0, \forall i, i = 1, 2, \dots, m$ dan H_a : paling sedikit terdapat satu β_i yang nilainya tidak sama dengan nol, untuk $i = 1, 2, \dots, m$, dengan β_i adalah koefisien regresi dari variabel ke- i dari total sebanyak m variabel. Hipotesis H_0 ditolak apabila $\chi^2_{hitung} > \chi^2_{tabel}$ atau jika $p\ value < \alpha$.

Tabel 2. Variabel Dummy yang Dibentuk

Faktor	Kategori Acuan	Variabel Dummy
X_2	Perkotaan	X_{2a} : Pedesaan
X_3	Tidak Berpendidikan	X_{3a} : SD X_{3b} : SMP X_{3c} : SMA
X_4	Sangat Miskin	X_{4a} : Miskin X_{4b} : Menengah X_{4c} : Kaya X_{4d} : Sangat Kaya

Langkah selanjutnya adalah melakukan uji parsial yang bertujuan untuk mengetahui bagaimana pengaruh dari masing-masing variabel bebas (faktor) dalam model regresi yang dibentuk [4]. Statistik uji yang digunakan dalam penelitian ini adalah statistik uji Wald yang dapat dituliskan sebagai

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

dengan $\hat{\beta}$ adalah koefisien penduga parameter dan $SE(\hat{\beta})$ adalah kesalahan standar (*standard error*) dari penduga parameter $\hat{\beta}$. Hipotesis dalam uji ini adalah $H_0: \beta_i = 0$

dan $H_a: \beta_i \neq 0$, untuk $i = 1, 2, \dots, m$. Hipotesis H_0 ditolak jika pada taraf signifikansi α , nilai W lebih besar dari nilai $Z_{\alpha/2}$ atau $p \text{ value} < \alpha$.

Setelah uji simultan dan uji parsial dilakukan, proses selanjutnya adalah seleksi variabel (*variable selection*) yang bertujuan untuk mendapatkan persamaan regresi terbaik serta memuat faktor-faktor yang berpengaruh nyata (signifikan) terhadap variabel respons. Dalam penelitian ini, metode seleksi variabel yang digunakan adalah metode eliminasi mundur Wald. Konsep dasar pada metode eliminasi mundur adalah dengan membuang variabel satu per satu dari model lengkap sedemikian sehingga terbentuk kombinasi variabel yang memberikan pengaruh secara nyata pada model.

Berdasarkan perhitungan yang telah dilakukan dan dibantu dengan perangkat lunak SPSS 19, $p \text{ value}$ pada uji simultan menggunakan metode *enter* adalah 0,009. Hal ini berarti bahwa dengan taraf signifikansi sebesar 10% pengaruh dari semua faktor terhadap waktu *survival* jika dimasukkan ke dalam model secara bersama-sama adalah signifikan. Hasil uji parsial, diperlihatkan dalam Tabel 3, sedangkan hasil dari seleksi variabel berdasarkan metode eliminasi mundur dituliskan dalam Tabel 4.

Tabel 3. Hasil Uji Parsial

Faktor	$\hat{\beta}$	$Exp(\hat{\beta})$	$p \text{ value}$
Usia (X_1)	-0,056	0,945	0,099
Pedesaan (X_{2a})	0,087	1,091	0,793
SD (X_{3a})	0,601	1,823	0,046
SMP(X_{3b})	0,413	1,512	0,489
SMA (X_{3c})	-0,212	0,809	0,839
Miskin (X_{4a})	-0,036	0,965	0,927
Menengah (X_{4b})	-0,630	0,533	0,079
Kaya (X_{4c})	-0,042	0,959	0,928
Sangat Kaya (X_{4d})	0,615	1,850	0,182

Tabel 4. Hasil Metode Eliminasi Mundur

Langkah	X_1	X_{2a}	X_{3a}	X_{3b}	X_{3c}	X_{4a}	X_{4b}	X_{4c}	X_{4d}	$p \text{ value}$	Ket.
1	√	√	√	√	√	√	√	√	√	0,009	Model 1
2	√	√	√	√	√	√	√		√	0,005	Model 2
3	√	√	√	√	√		√		√	0,003	Model 3
4	√	√	√	√			√		√	0,001	Model 4
5	√		√	√			√		√	0,001	Model 5
6	√		√				√		√	0,000	Model 6

Tabel 5. Hasil Uji Parsial Langkah Ke-6 Metode Eliminasi Mundur

Faktor	$\hat{\beta}$	$Exp(\hat{\beta})$	$p\ value$
Usia (X_1)	-0,048	0,953	0,078
SD (X_{3a})	0,501	1,650	0,045
Menengah (X_{4b})	-0,592	0,553	0,052
Sangat Kaya (X_{4d})	0,728	2,072	0,024

Dalam Tabel 3, diperlihatkan hasil uji parsial dari masing-masing faktor. Dari total tujuh faktor yang diuji, hanya tiga faktor yang signifikan pada tingkat signifikansi 10%, yaitu usia (X_1), tingkat pendidikan SD (X_{3a}), dan tingkat perekonomian menengah (X_{4b}). Berdasarkan *output* metode eliminasi mundur dalam Tabel 4, model 2 diperoleh dengan mengeliminasi variabel X_{4c} dari model dikarenakan memiliki $p\ value = 0,928$ yang paling besar (Tabel 3). Setelah dilakukan uji simultan pada model 2, diperoleh $p\ value$ sebesar 0,005 yang lebih kecil dari $p\ value$ model 1, yang berarti model 2 lebih baik daripada model 1. Proses ini dilanjutkan dengan cara mengeliminasi satu persatu faktor yang paling tidak signifikan hingga berhenti di langkah ke-6, karena pada langkah terakhir ini semua $p\ value$ pada uji parsial telah signifikan pada tingkat signifikansi 10% (Tabel 5). Model 6 juga merupakan model terbaik dibandingkan dengan model-model yang terbentuk pada langkah-langkah sebelumnya, yang dalam hal ini $p\ value = 0,000$ untuk uji simultan pada langkah ke-6 adalah yang terkecil dibandingkan dengan $p\ value$ uji simultan pada langkah-langkah sebelumnya. Jadi, model regresi Cox terbaik yang diperoleh adalah

$$h(t) = h_0(t) \exp(-0,048 X_1 + 0,501 X_{3a} - 0,592 X_{4b} + 0,728 X_{4d}), \quad (1)$$

dengan X_1 adalah usia, X_{3a} adalah tingkat pendidikan SD, X_{4b} adalah tingkat ekonomi menengah, dan X_{4d} adalah tingkat ekonomi sangat kaya.

Interpretasi dari Model Terbaik yang Diperoleh

Tanda negatif atau positif dari penduga koefisien regresi dalam (1) memperlihatkan arah dari tingkat fertilitas individu. Koefisien dari X_1 bertanda negatif berarti bahwa semakin tinggi usia seorang individu maka waktu tunggu hingga melahirkan anak pertama akan semakin lama. Dengan kata lain, individu yang berusia lebih tua memiliki selang kelahiran anak pertama yang lebih panjang dibandingkan individu yang berusia lebih muda atau semakin tinggi usia ibu maka semakin rendah pula tingkat fertilitas yang dimiliki ibu tersebut. Tanda negatif juga dimiliki oleh koefisien penduga dari faktor X_{4b} dalam (1) yang bertipe kategorik. Hal ini berarti bahwa individu yang dikategorikan sebagai penduduk miskin memiliki selang kelahiran anak pertama yang lebih pendek jika dibandingkan dengan individu dengan tingkat

perekonomian pada kategori acuan (tingkat perekonomian sangat miskin). Koefisien penduga dari faktor X_{3a} bertanda positif, yang dapat diinterpretasikan bahwa ibu yang berpendidikan SD memiliki selang kelahiran anak pertama yang lebih panjang atau memiliki tingkat fertilitas yang lebih rendah jika dibandingkan individu dengan tingkat pendidikan pada kategori acuan (tidak berpendidikan). Interpretasi yang hampir sama juga berlaku pada tingkat perekonomian sangat kaya (X_{4d}) yang koefisiennya bertanda positif, yakni individu yang memiliki tingkat perekonomian sangat kaya memiliki tingkat fertilitas yang rendah jika dibandingkan dengan individu yang memiliki tingkat perekonomian sangat miskin sebagai kategori acuan.

Nilai $exp(\beta)$ disebut juga sebagai *hazard ratio* (HR), yang didefinisikan sebagai rasio peluang terjadinya peristiwa yang diperhatikan antara suatu individu dengan individu lainnya. Berdasarkan hasil dari pemilihan model terbaik menggunakan metode eliminasi mundur yang diperlihatkan dalam Tabel 5, telah didapatkan nilai-nilai HR untuk setiap faktor yang termuat dalam model regresi terbaik (1). Nilai HR untuk faktor X_1 adalah 0,953 yang berarti untuk setiap kenaikan usia individu sebesar satu unit, maka nilai tingkat *hazard* dari individu tersebut akan turun sebesar $(1 - 0,953) \times 100\% = 4,7\%$. Sebagai contoh, seorang ibu yang baru menikah dan berusia 25 tahun memiliki peluang yang lebih kecil sebesar 4,7% untuk melahirkan anak pertama dibandingkan dengan ibu yang baru menikah dan berusia 24 tahun. Interpretasi yang hampir sama juga pada nilai HR untuk faktor X_{4b} , namun menurunnya tingkat *hazard* dibandingkan dengan kategori acuan (kategori sangat miskin). Misalkan ada seorang ibu yang pada saat menikah berasal dari keluarga dengan tingkat perekonomian menengah, maka ibu tersebut memiliki peluang yang lebih rendah sebesar $(1 - 0,553) \times 100\% = 44,7\%$ untuk mendapatkan anak pertama jika dibandingkan dengan ibu yang berasal dari keluarga dengan tingkat perekonomian sangat miskin.

Dalam (1) juga diperoleh nilai HR yang nilainya lebih besar dari satu, yaitu untuk faktor X_{3a} dan X_{4d} yang nilainya berurut-turut adalah 1,65 dan 2,072. Nilai HR = 1,65 dari faktor X_{3a} berarti bahwa seorang ibu yang baru menikah dan berpendidikan SD memiliki peluang untuk mendapatkan anak pertama $(1,65 - 1) \times 100\% = 65\%$ lebih besar dibandingkan dengan seorang ibu yang baru menikah dan tidak berpendidikan (kategori acuan). Nilai HR = 2,072 dari faktor X_{4d} , dapat diinterpretasikan bahwa seorang ibu yang baru menikah dan berasal dari keluarga yang sangat kaya memiliki peluang yang lebih besar sebesar $(2,072 - 1) \times 100\% = 107,2\%$ untuk mendapatkan anak pertama dibandingkan dengan ibu yang baru menikah dan berasal dari keluarga dengan tingkat perekonomian sangat miskin.

4. Simpulan dan Saran

Berdasarkan model regresi terbaik yang diperoleh, tempat tinggal bukanlah faktor yang signifikan memengaruhi tingkat fertilitas penduduk, sedangkan usia ibu, tingkat pendidikan SD, tingkat ekonomi menengah, dan tingkat ekonomi sangat kaya adalah faktor-faktor yang berpengaruh signifikan. Semakin tinggi usia ibu ketika menikah, maka semakin rendah peluangnya untuk melahirkan anak pertama. Ibu yang tidak berpendidikan memiliki tingkat fertilitas yang lebih tinggi dibandingkan dengan ibu yang berpendidikan SD, namun untuk tingkat pendidikan lainnya belum dapat dibandingkan karena dalam model terbaik yang diperoleh faktor tersebut tidak berpengaruh signifikan. Ibu dengan tingkat ekonomi menengah memiliki tingkat fertilitas yang lebih rendah dibandingkan dengan ibu dengan tingkat ekonomi sangat miskin. Sementara itu, ibu yang tingkat perekonomiannya sangat kaya memiliki tingkat fertilitas yang lebih tinggi jika dibandingkan dengan ibu yang tingkat perekonomiannya dikategorikan sangat miskin.

Hasil dari penelitian ini dapat dijadikan sebagai salah satu landasan bagi para praktisi maupun peneliti yang ingin menerapkan teori dan metode dalam analisis *survival* pada data selang kelahiran anak pertama. Namun, jumlah subjek dalam penelitian ini masih belum cukup untuk digunakan sebagai sampel yang representatif dalam mengestimasi faktor-faktor utama yang berpengaruh terhadap tingkat fertilitas penduduk di Provinsi Sumatera Selatan. Oleh karena itu, bagi peneliti yang tertarik untuk melanjutkan penelitian ini pada masa yang akan datang, disarankan untuk menambah lagi jumlah subjek penelitiannya sehingga hasil yang diharapkan akan lebih baik.

Daftar Pustaka

- [1] Ida B. Mantra, *Pengantar Studi Demografi*. Jakarta: Nur Cahaya, 2003.
- [2] R. Hidayat, H. Sumarno, and Endar H. Nugrahani, "Survival Analysis in Modeling the Birth Interval of the First Child in Indonesia," *Open Journal of Statistics*, vol. 4, pp. 198-206, 2014.
- [3] Elisa T. Lee and Wengyu J. Wang, *Statistical Methods for Survival Data Analysis*, 3rd ed. Hoboken: John Wiley & Sons, Inc., 2003.
- [4] A. Faruk, A. Amran, and N. Nasir, "Aplikasi Model Proportional Hazard Cox pada Waktu Tunggu Kerja Lulusan Jurusan Matematika Fakultas MIPA Universitas Sriwijaya," *Jurnal Penelitian Sains*, vol. 17, no. 1, pp. 5-8, January 2014.
- [5] A. Faruk, "Analisis Data Tersensor Interval Dalam Pemodelan Waktu Mendapatkan Pekerjaan Pertama Alumni Universitas Sriwijaya," in *Seminar*

Nasional Matematika dan Pendidikan Matematika, Yogyakarta, 2015, pp. 123-130.

- [6] A. Amran and A. Faruk, "Model Survival Nonparametrik Pada Data Rawat Inap Pasien Diare di Puskesmas Indralaya," *Jurnal Matematika*, vol. 5, no. 2, pp. 105-116, 2015.
- [7] A. Faruk, "Estimasi Parameter Data Tersensor Tipe I Berdistribusi Loglogistik Menggunakan Maximum Likelihood Estimate dan Iterasi Newton-Rhapson," in *Seminar Nasional MIPA*, Palembang, 2014, pp. 21-25.
- [8] A. Faruk, "Analisis Survival Parametrik Pada DataTracer Study Universitas Sriwijaya," *Jurnal Matematika*, vol. 5, no. 2, pp. 68-78, 2015.
- [9] I. Gijbels, "Censored data," *WIREs Computational Statistics*, vol. 2, pp. 178–188, 2010.
- [10] J. Ni, "Application of Cox Proportional Hazard Model to the Stock Exchange Market," *B.S. Undergraduate Mathematics Exchange*, vol. 6, no. 1, pp. 12-18, 2009.
- [11] Badan Pusat Statistik, *Statistik Indonesia 2016*. Jakarta: Badan Pusat Statistik, 2016.
- [12] David G. Kleinbaum and M. Klein, *Survival Analysis A Self-Learning Text*, 2nd ed. New York: Springer Science+ Business Media, Inc., 2005.
- [13] John P. Klein et al., Eds., *Handbook of Survival Analysis*. Boca Raton: Taylor & Francis Group, LLC, 2014.
- [14] D. R. Cox, "Regression Models and Life-Tables," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 34, no. 2, pp. 187-220, 1972.