



TUGAS AKHIR - KS184822

**DETERMINAN USIA PERTAMA MELAHIRKAN PADA
WANITA USIA SUBUR DI JAWA TIMUR
MENGUNAKAN REGRESI *COX EXTENDED***

**FITRIA NUR AIDA
NRP 062116 4000 0070**

**Dosen Pembimbing
Dr. Santi Wulan Purnami, S.Si.,M.Si.
Dr. Sutikno, S.Si.,M.Si.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS SAINS DAN ANALITIKA DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2020**



TUGAS AKHIR - KS184822

**DETERMINAN USIA PERTAMA MELAHIRKAN PADA
WANITA USIA SUBUR DI JAWA TIMUR
MENGUNAKAN REGRESI *COX EXTENDED***

**FITRIA NUR AIDA
NRP 062116 4000 0070**

**Dosen Pembimbing
Dr. Santi Wulan Purnami, S.Si.,M.Si.
Dr. Sutikno, S.Si.,M.Si.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS SAINS DAN ANALITIKA DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2020**



FINAL PROJECT - KS184822

**DETERMINANT FACTORS OF TIMING OF FIRST
CHILDBIRTH AMONG WOMEN OF CHILDBEARING
AGE IN EAST JAVA USING *EXTENDED COX
REGRESSION***

**FITRIA NUR AIDA
SN 062116 4000 0070**

Supervisors

Dr. Santi Wulan Purnami, S.Si.,M.Si.

Dr. Sutikno, S.Si.,M.Si.

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF SCIENCE AND DATA ANALYTICS
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2020**

(Halaman ini sengaja dikosongkan)

LEMBAR PENGESAHAN

DETERMINAN USIA PERTAMA MELAHIRKAN PADA WANITA USIA SUBUR DI JAWA TIMUR MENGUNAKAN REGRESI COX EXTENDED

TUGAS AKHIR

Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Statistika
pada
Program Studi Sarjana Departemen Statistika
Fakultas Sains dan Analitika Data
Institut Teknologi Sepuluh Nopember

Oleh :

Fitria Nur Aida

NRP. 062116 4000 0070

Disetujui oleh Pembimbing:

Dr. Santi Wulan Purnami, S.Si., M.Si.

NIP. 19720923 199803 2 001

Dr. Sutikno S.Si., M.Si.

NIP. 19710313 199702 1 001

Mengetahui,
Kepala Departemen Statistika

DEPARTEMEN
STATISTIKA

Dr. Kartika Fithriasari, M. Si

NIP. 19691212 199303 2 002

SURABAYA, JANUARI 2020

(Halaman ini sengaja dikosongkan)

**DETERMINAN USIA PERTAMA MELAHIRKAN PADA
WANITA USIA SUBUR DI JAWA TIMUR
MENGUNAKAN REGRESI COX EXTENDED**

Nama Mahasiswa : Fitria Nur Aida
NRP : 062116 4000 0070
Departemen : Statistika-FSAD-ITS
Dosen Pembimbing : Dr. Santi Wilan Purnami, S.Si., M.Si.
Dr. Sutikno, S.Si., M.Si.

ABSTRAK

Usia pertama melahirkan signifikan berpengaruh terhadap hidup seorang wanita karena keibuan melibatkan komitmen besar dengan waktu dan sumber daya serta cenderung mengatur tahap dimana peran lain. Usia pertama kali melahirkan juga mempengaruhi lingkungan tempat anak-anak tumbuh dan perkembangan intelektual mereka. Setiap wanita yang pertama kali melahirkan memiliki kemungkinan mengalami gangguan psikologis seperti kecemasan dan ketakutan berlebihan selama proses persalinan bahkan depresi pasca melahirkan. Mengingat pentingnya dan kemungkinan luasnya konsekuensi dari usia pertama melahirkan bagi seorang wanita, penelitian ini yaitu determinan usia pertama melahirkan yang dapat diketahui dari analisis faktor-faktor yang diduga berpengaruh terhadap usia pertama melahirkan dengan pendekatan analisis ketahanan hidup (survival). Analisis survival yang dapat dilakukan adalah Regresi Cox Proportional Hazard, tetapi dalam penelitian ini asumsi Proportional Hazard tidak terpenuhi sehingga penelitian ini menggunakan model Cox Extended. Pemodelan menggunakan Regresi Cox Extended menghasilkan kombinasi fungsi waktu $\ln(T)$ dan Heavyside sebagai kombinasi waktu terbaik. Faktor usia kawin pertama dan pendidikan merupakan faktor yang berpengaruh secara signifikan terhadap usia pertama melahirkan dari wanita usia subur di Jawa Timur.

Kata kunci: Analisis Survival, Cox Extended, Cox Proportional Hazard, Usia Pertama Melahirkan, Wanita Usia Subur.

(Halaman ini sengaja dikosongkan)

**DETERMINANT FACTORS OF TIMING OF FIRST
CHILDBIRTH AMONG WOMEN OF CHILDBEARING
AGE IN EAST JAVA USING *EXTENDED COX
REGRESSION***

Name : **Fitria Nur Aida**
Student Number : **062116 4000 0070**
Department : **Statistics**
Supervisors : **Dr. Santi Wulan Purnami, S.Si., M.Si.**
Dr. Sutikno, S.Si., M.Si.

ABSTRACT

Timing of first childbirth significantly influences a woman's life because motherhood involves a large commitment with time, resources, and tends to regulate the stage at which other. The timing of first childbirth birth also influences the environment where children grow and their intellectual develop. Every woman whose first childbirth has the possibility of experiencing psychological disorders such as anxiety and excessive fear during labor and even postpartum depression. Given the importance and possible extent of the consequences of the timing of childbirth for a woman, this study is a determinant factors of the timing of the childbirth which can be known from the analysis of some factors thought to influence the timing childbirth using a survival analysis approach. Survival analysis that can be used is Cox Proportional Hazard Regression, but in this study the assumptions of Proportional Hazard are not fulfilled so this study uses the Cox Extended model. Modeling using Extended Cox Regression produces a combination of the Ln(T) and Heavyside time functions as the best combination of time functions. Factors of age at first marriage and education are factors that significantly influence the timing of first childbirth among woman of childbearing age in East Java.

Keywords: Survival Analysis, Extended Cox, Cox Proportional Hazard, Timing of First Childbirth, Woman Of Childbearing Age.

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji syukur penulis panjatkan atas rahmat dan hidayah yang diberikan Allah SWT sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul “**Determinan Usia Pertama Melahirkan Pada Wanita Usia Subur Di Jawa Timur Menggunakan Regresi Cox Extended**” dengan lancar.

Penulis menyadari bahwa Tugas Akhir ini dapat terselesaikan tidak terlepas dari bantuan dan dukungan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Kedua orang tua, Bapak Mansur dan Ibu Siti Asiyah, dan keluarga tercinta yang sudah mendoakan tiada henti dan memberi dukungan baik secara moril mau-pun materil kepada penulis.
2. Dr. Kartika Fithriasari, M.Si. selaku Kepala Departemen Statistika yang telah memberikan fasilitas, sarana, dan prasarana sehingga membantu kelancaran penyelesaian Tugas Akhir.
3. Dr. Kartika Fithriarsari, M.Si. selaku dosen wali selama masa studi yang telah banyak memberikan saran, arahan, dan semangat dalam menjalani perkuliahan.
4. Dr. Santi Wulan Purnami, S.Si., M.Si. dan Dr. Sutikno, S.Si., M.Si. selaku dosen pembimbing yang telah meluangkan waktu dan dengan sabar memberi bimbingan, saran, dan motivasi selama pengerjaan dan penyusunan Tugas Akhir.
5. Dr. Puhadi, M.Sc. dan Bu Diyah Herowati, S.Sos., M.PHR selaku dosen penguji yang telah memberikan koreksi dan masukan dengan sangat sabar sehingga Tugas Akhir ini dapat lebih baik.
6. Seluruh dosen Statistika ITS yang telah memberikan ilmu dan pengetahuan yang sangat berharga, serta segenap karyawan Departemen Statistika ITS.
7. Partner dan teman-teman seperjuangan Wisuda 121 khususnya Kinanthi Sukma Wening dan Ni Luh Putu Ika Candrawengi

yang saling memberikan banyak motivasi, bantuan, semangat hingga dapat menyelesaikan Tugas Akhir bersama-sama.

8. Teman-teman Statistika ITS angkatan 2016, “TR16GER”, khususnya Mohammad Naufal Abdullah yang selalu membantu menyelesaikan persoalan akademis maupun *me-review* materi-materi perkuliahan untuk menyelesaikan Tugas Akhir penulis ini dan Finola Trisnisa yang selalu memberikan dukungan dan bersedia mendengarkan keluh kesah hiruk pikuk masa perkuliahan maupun curhatan pribadi penulis selama ini.
9. Teman-teman dan sahabat yang sudah memberikan semangat dan doa hingga saat ini.

Besar harapan penulis untuk mendapatkan kritik dan saran yang membangun sehingga Tugas Akhir ini dapat memberikan manfaat bagi semua pihak yang terkait.

Surabaya, Januari 2020

Penulis

DAFTAR ISI

Halaman

LEMBAR PENGESAHAN.....	Error! Bookmark not defined.
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI.....	xiii
DAFTAR GAMBAR	xvii
DAFTAR TABEL.....	xix
DAFTAR LAMPIRAN	xxi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian	6
1.4 Manfaat Penelitian	6
1.5 Batasan Masalah	6
BAB II TINJAUAN PUSTAKA	9
2.1 Statistika Deskriptif.....	9
2.1.1 Mean	9
2.1.2 Deviasi Standar	10
2.2 Analisis <i>Survival</i>	10
2.2.1 Fungsi <i>Survival</i> dan Fungsi <i>Hazard</i>	12
2.2.2 Kurva <i>Kaplan Meier</i>	14
2.2.3 Uji <i>Log-Rank</i>	15
2.2.4 Regresi <i>Cox Proportional Hazard</i>	16
2.2.5 Asumsi <i>Proportional Hazard</i>	19
2.2.6 Model <i>Cox Extended</i>	20
2.2.7 Pengujian Signifikansi Parameter	24
2.2.8 Pemilihan Model Terbaik.....	26

2.2.9	<i>Hazard Ratio</i>	27
2.3	Usia Pertama Melahirkan Beserta Faktor-faktor yang Diduga Berpengaruh	28
2.3.1	Usia Kawin Pertama.....	28
2.3.2	Pendidikan	29
2.3.3	Region.....	29
BAB III	METODOLOGI PENELITIAN	31
3.1	Sumber Data.....	31
3.2	Metodologi SDKI 2017.....	31
3.2.1	Rancangan Sampel dan Hasil Kunjungan.....	31
3.2.2	Kuesioner	33
3.3	Kerangka Konsep	33
3.4	Variabel Penelitian	39
3.5	Struktur Data	40
3.6	Penulisan Model <i>Cox Extended</i>	41
3.7	Langkah Analisis.....	42
3.8	Diagram Alir Langkah-langkah Analisis.....	44
BAB IV	ANALISIS DAN PEMBAHASAN	47
4.1	Karakteristik Usia Pertama Melahirkan Pada Wanita Usia Subur di Jawa Timur.....	47
4.1.1	Karakteristik Usia Pertama Melahirkan Berdasarkan Faktor Usia Kawin Pertama.....	48
4.1.2	Karakteristik Usia Pertama Melahirkan Berdasarkan Faktor Pendidikan.....	48
4.1.3	Karakteristik Usia Pertama Melahirkan Berdasarkan Faktor Region	50
4.2	Pemodelan Regresi <i>Cox Extended</i> Pada Usia Pertama Melahirkan	52
4.2.1	Kurva <i>Survival Kaplan Meier</i> dan Uji <i>Log-Rank</i> ...	52
4.2.2	Pengujian Asumsi <i>Proportional Hazard</i>	56

4.2.3	Pemodelan Regresi <i>Cox Extended</i> Menggunakan Kombinasi Fungsi Waktu <i>Heavyside</i>	57
4.2.4	Pemodelan Regresi <i>Cox Extended</i> Menggunakan Kombinasi Fungsi Waktu $\ln(T)$ dan <i>Heavyside</i>	60
4.2.5	Pemilihan Kombinasi Fungsi Waktu Terbaik Berdasarkan Kriteria AICc.....	62
4.2.6	Pemilihan Model Terbaik Menggunakan <i>Backward Elimination</i>	63
BAB V KESIMPULAN DAN SARAN		69
5.1	Kesimpulan	69
5.2	Saran dan Rekomendasi	70
DAFTAR PUSTAKA		
LAMPIRAN		
BIODATA PENULIS		

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

	Halaman
Gambar 2. 1 Ilustrasi Kurva Kaplan Meier.	15
Gambar 3. 1 Kerangka Konsep Penelitian.	38
Gambar 3. 2 Diagram Alir Penelitian.	45
Gambar 4. 1 Rata-rata Usia Pertama Melahirkan dan Rata-rata Usia Kawin Pertama Berdasarkan Pendidikan.	50
Gambar 4. 2 Rata-rata Usia Pertama Melahirkan dan Rata-rata Usia Kawin Pertama Berdasarkan Region.	51
Gambar 4. 3 Kurva Survival Usia Pertama Melahirkan.	52
Gambar 4. 4 Kurva Kaplan Meier Berdasarkan Faktor Pendidikan.	53
Gambar 4. 5 Kurva Kaplan Meier Berdasarkan Faktor Region.	55
Gambar 4. 6 Grafik Hazard Ratio Faktor Usia Kawin Pertama.	65
Gambar 4. 7 Grafik Hazard Ratio Faktor Pendidikan.	65

(Halaman ini sengaja dikosongkan)

DAFTAR TABEL

	Halaman
Tabel 2. 1 Hazard Ratio Masing-masing $g_m(t)$	24
Tabel 3. 1 Variabel Penelitian.....	39
Tabel 3. 2 Struktur Data Penelitian	40
Tabel 4. 1 Statistika Deskriptif Usia Pertama Melahirkan	47
Tabel 4. 2 Statistika Deskriptif Usia Kawin Pertama	48
Tabel 4. 3 Statistika Deskriptif Usia Pertama Melahirkan	49
Tabel 4. 4 Statistika Deskriptif Usia Pertama Melahirkan	50
Tabel 4. 5 Hasil Uji Goodness of Fit	56
Tabel 4. 6 Hasil estimasi Parameter Model Cox Extended Kombinasi Fungsi Waktu Heavyside.....	58
Tabel 4. 7 Hasil estimasi Parameter Model Cox Extended Kombinasi Fungsi Waktu Ln(T) dan Heavyside.....	60
Tabel 4. 8 Perbandingan Nilai AICc Dari Kombinasi Fungsi Waktu yang Berbeda	62
Tabel 4. 9 Hasil estimasi Parameter Model Cox Extended Terbaik	63
Tabel 4. 10 Hazard Ratio Model Terbaik.....	66

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data Wanita Usia Subur di Jawa Timur yang Pernah Menikah.....	75
Lampiran 2. Syntax R Statistika Deskriptif	76
Lampiran 3. Syntax R Kaplan Meier	76
Lampiran 4. Syntax R Uji Log-Rank.....	77
Lampiran 5. Syntax R Uji Asumsi PH Menggunakan Kurva -ln[-lnS(t)]	77
Lampiran 6. Syntax R Pemodelan Cox PH dan Uji Asumsi PH Menggunakan Uji Goodness of Fit.....	79
Lampiran 7. Syntax R Regresi Cox Extended Fungsi Waktu Heavyside	80
Lampiran 8. Syntax R Regresi Cox Extended Fungsi Waktu Ln(t) dan Heavyside	83
Lampiran 9. Syntax R Regresi Cox Extended Fungsi Waktu T dan Heavyside	84
Lampiran 10. Syntax R Kurva Estimated Hazard Ratio Variabel Usia Kawin Pertama	85
Lampiran 11. Output Statistika Deskriptif	85
Lampiran 12. Output R Uji Log-Rank.....	86
Lampiran 13. Output R Regresi Cox Proportional Hazard	87
Lampiran 14. Output R Uji Asumsi PH Menggunakan Uji Goodness of Fit.....	87
Lampiran 15. Grafik Martingale Residual	88
Lampiran 16. Grafik -ln[-lnS(t)].....	88
Lampiran 17. <i>Output R Regresi Cox Extended Fungsi Waktu Heavyside</i>	89
Lampiran 18. Output R Regresi Cox Extended Fungsi Waktu Ln(t) dan Heavyside.....	91

Lampiran 19. Output R Regresi Cox Extended Fungsi Waktu T dan Heavyside	93
Lampiran 20. Surat Pernyataan Data	95

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara dengan jumlah penduduk terbesar keempat di dunia pada tahun 2018. Berdasarkan proyeksi Badan Perencanaan Pembangunan Nasional (Bappenas) 2013 bahwa jumlah penduduk Indonesia pada tahun 2018 mencapai 265 juta jiwa yang terdiri dari sekitar 133,17 juta jiwa laki-laki dan 131,88 juta jiwa perempuan. Sekitar 57% dari jumlah tersebut merupakan penduduk Pulau Jawa. Besarnya jumlah tersebut sejalan dengan tingkat kelahiran bayi pada tahun 2018 sebesar 13.370 kelahiran bayi yang menempatkan Indonesia sebagai negara terbesar kelima yang menyumbang jumlah kelahiran bayi berdasarkan data UNICEF.

Salah satu sasaran program Badan Kependudukan dan Keluarga Berencana Nasional (BKKBN, 2018) Provinsi Jawa Timur yaitu menurunkan angka kelahiran pada wanita remaja khususnya berusia 15-19 tahun (ASFR 15-19 tahun) sebagaimana tercantum dalam Rencana Strategis Perwakilan BKKBN Provinsi Jawa Timur tahun 2015-2019. ASFR 15-19 tahun di Provinsi Jawa Timur tahun 2012 sebanyak 53 kelahiran per 1000 wanita dan menurun menjadi 29 kelahiran per 1000 wanita pada tahun 2017. ASFR 20-24 tahun sebesar 138 kelahiran per 1000 wanita pada tahun 2012 dan mengalami penurunan menjadi 104 kelahiran per 1000 wanita pada tahun 2017.

Penyebab utama terjadinya kelahiran pada usia remaja yaitu mengalami kehamilan pra-nikah yang menyebabkan terjadinya pernikahan usia dini sehingga usia pertama melahirkan semakin muda. Wanita yang menikah di usia muda cenderung hamil pada usia remaja sehingga memiliki resiko yang lebih besar dari wanita yang hamil pada usia yang lebih dewasa. Remaja wanita yang melahirkan sebelum berusia 15 tahun memiliki resiko kematian 5 kali lebih tinggi saat melahirkan dibandingkan wanita yang berusia di atas 20 tahun (UNICEF, 2008). Pernikahan dini memiliki

dampak yang kompleks pada seluruh aspek kehidupan seperti aspek sosial dan ekonomi, aspek demografi dan kesehatan, dll. Semakin muda usia menikah maka peluang seorang wanita melahirkan di usia muda semakin tinggi. Kehamilan remaja memiliki resiko bagi wanita dan keturunannya seperti resiko lahir mati, keguguran, dan kehamilan ektopik, serta beberapa kelahiran dan kelainan bawaan (Baird, et al., 2005). Tidak hanya dialami oleh wanita yang melahirkan di usia muda, tetapi setiap wanita yang pertama kali melahirkan memiliki kemungkinan mengalami gangguan psikologis seperti kecemasan dan ketakutan berlebihan selama proses persalinan bahkan depresi pasca melahirkan (Nevid & Greene, 2005).

Wanita Usia Subur (WUS) adalah wanita yang keadaan organ reproduksinya berfungsi dengan baik antara usia 15-49 tahun dan berpotensi mempunyai keturunan. Kelahiran anak pertama merupakan peristiwa penting dalam kehidupan wanita. Sebuah transisi dari seorang wanita menjadi seorang ibu dengan peran dan tanggung jawab yang terkait, sering kali dengan mengesampingkan pendidikan lebih lanjut dan peran membangun karir (Rindfuss & John, 1983). Awal kelahiran pertama mungkin memiliki efek buruk pada kesejahteraan sosial ekonomi wanita di tahun kemudian karena adanya pergantian peran seperti pengembangan karir dan pendidikan untuk peran lainnya (Fagbamigbe & Idemudia, 2016). Tidak hanya berpengaruh terhadap kesejahteraan sosial ekonomi wanita tetapi juga berpengaruh secara tidak langsung terhadap tumbuh kembang anak pertama.

Usia merupakan faktor yang sangat penting dalam hal kesuburan, dan bahkan dengan semua kemajuan dalam bantuan reproduksi, usia tetap merupakan hambatan yang tidak dapat diatasi (George & Kamath, 2010). Usia pertama melahirkan signifikan berpengaruh terhadap hidup seorang wanita karena keibuan melibatkan komitmen besar dengan waktu dan sumber daya dan cenderung mengatur tahap dimana peran lain (Rindfuss & John, 1983). Pada tingkat makro, tren usia pertama melahirkann memiliki efek penting pada laju perubahan sosial, tren masa

kesuburan, dan keadaan ekonomi. Usia pertama kali melahirkan juga mempengaruhi lingkungan tempat anak-anak tumbuh dan perkembangan intelektual mereka (Rindfuss & John, 1983). Oleh karena itu sangat perlu diperhatikan usia pertama melahirkan bagi seorang wanita.

Fagbamigbe & Idemudia (2016) menyebutkan faktor-faktor yang diduga mempengaruhi usia pertama melahirkan beragam, beberapa diantaranya seperti faktor individu, faktor keluarga, faktor ekonomi, faktor sosial. Faktor individu meliputi pendidikan dan penggunaan kontrasepsi. Faktor keluarga meliputi karakteristik latar belakang sosial ekonomi seperti pendidikan orang tua, tempat tinggal, wilayah kelahiran secara geografis, status sosial, status pekerjaan, dll. Faktor-faktor sosial termasuk praktek norma-norma, dan lingkungan sekitar.

Salah satu analisis statistika yang sesuai untuk digunakan adalah analisis ketahanan hidup (*survival*). Analisis ketahanan hidup adalah teknik statistika yang digunakan untuk menganalisis data yang diperoleh dari catatan waktu yang dicapai suatu objek sampai terjadinya peristiwa tertentu (*failure event*) dan bertujuan mengetahui variabel yang berpengaruh mulai dari awal kejadian hingga akhir kejadian (Kleinbaum & Klein, 2012). Salah satu metode dalam analisis *survival* yang sering digunakan dalam analisis ketahanan hidup adalah Regresi *Cox Proportional Hazard*. Regresi *Cox Proportional Hazard* merupakan metode semiparametrik yang dapat digunakan untuk memodelkan hubungan antara variabel respon yang berupa waktu *survival* dengan satu atau lebih variabel prediktor (Kleinbaum & Klein, 2012). Metode Regresi *Cox Proportional Hazard* juga merupakan metode yang dapat digunakan apabila keluaran yang diobservasi berupa panjang waktu suatu kejadian.

Penelitian yang telah berkembang dilakukan oleh Fagbamigbe & Idemudia (2016) mengenai analisis *survival* dan penentuan faktor-faktor yang berpengaruh terhadap waktu pertama melahirkan pada wanita di Nigeria. Penelitian tersebut menginisiasi dengan hipotesis bahwa karakteristik latar belakang

serorang wanita akan mempengaruhi perkembangan dirinya untuk kelahiran pertama. Faktor-faktor yang digunakan dalam penelitian tersebut adalah region, wilayah bagian, pendidikan, agama, usia pertama menikah, etnisitas, pernah tidaknya merokok, pernah tidaknya mengakhiri kehamilan, dan pernah tidaknya menggunakan alat kontrasepsi. Metode yang digunakan dalam penelitian tersebut adalah Regresi *Cox Proportional Hazard* dan Kurva *Kaplan Meier* dengan sembilan variabel prediktor, data tersensor, dan usia pertama melahirkan sebagai *survival time*. Hasil dari penelitian tersebut adalah faktor pendidikan dan waktu menikah menjadi faktor yang paling berpengaruh signifikan terhadap usia pertama melahirkan.

Penelitian yang serupa juga dilakukan oleh Rao K. V. & Balakhrisan (1988) analisis model *Hazard* pada usia pertama melahirkan di Canada. Penelitian tersebut bertujuan untuk mendapatkan pola usia pertama melahirkan dengan menggunakan dua metode yaitu median untuk mendapatkan pola kelompok atau kohor usia pertama melahirkan berdasarkan setiap variabel dan *Proportional Hazard* model untuk mendapatkan nilai *Hazard Ratio* dan besarnya pengaruh dari setiap variabel prediktornya. Penelitian ini menggunakan *Proportional Hazard* model karena keunggulannya yang melebihi metode regresi pada umumnya: (i) dapat menangani pengamatan yang tersensor, (ii) memungkinkan untuk mengambil informasi tentang jumlah, waktu, dan urutan kelahiran pertama, (iii) menghasilkan koefisien untuk berbagai kovariat usia pada kelahiran pertama yang mirip dengan hasil regresi sehingga hasil mudah diinterpretasi. Dalam penelitian ini tidak dijelaskan mengenai pemenuhan asumsi *Proportional Hazard*.

Namun pada model Regresi *Cox Proportional Hazard* diperlukan asumsi *Proportional Hazard* yang mana *hazard ratio* adalah konstan. Waktu seringkali menyebabkan perubahan terhadap *hazard ratio* sehingga diperlukan pemenuhan asumsi ini. Apabila asumsi *Proportional Hazard* tidak terpenuhi maka diperlukan model yang melibatkan variabel yang bergantung pada

waktu, yaitu model Regresi *Cox Extended* dan *Cox Stratified*. Metode Regresi *Cox Extended* memiliki kelebihan dibandingkan Regresi *Cox Stratified* yaitu lebih mudah dalam interpretasi model, model memperhatikan fungsi waktu, dan interpretasi *Hazard Ratio* untuk interval waktu yang berbeda lebih dibutuhkan dalam penelitian ini. Variabel yang bergantung terhadap waktu memiliki arti bahwa nilai variabel tersebut dapat berubah setiap saat (Kleinbaum & Klein, 2012).

Pada penelitian ini akan dilakukan analisis *survival* pada usia pertama melahirkan dari wanita usia subur di Jawa Timur yang berusia 15-49 tahun dan pernah menikah berdasarkan hasil SDKI 2017 dengan faktor-faktor yang diduga mempengaruhinya. Dalam pemodelan *survival* ini dilakukan dengan menggunakan Regresi *Cox Extended*. Diharapkan dengan adanya penelitian ini dapat memberikan informasi kepada wanita usia subur dan atau masyarakat mengenai faktor-faktor yang signifikan berpengaruh terhadap usia pertama melahirkan pada wanita usia subur di Jawa Timur dan rekomendasi dalam penentuan kebijakan perencanaan program kesehatan reproduksi wanita.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijabarkan, mengingat pentingnya dan kemungkinan luasnya konsekuensi dari usia pertama melahirkan bagi seorang wanita, penelitian yang dilakukan adalah determinan usia pertama melahirkan yang dapat diketahui dari analisis faktor-faktor yang diduga berpengaruh terhadap usia pertama melahirkan dengan pendekatan analisis ketahanan hidup (*survival*). Analisis *survival* dapat dilakukan dengan menggunakan Regresi *Cox Proportional Hazard*. Tetapi dalam penelitian ini asumsi *Proportional Hazard* tidak terpenuhi. Oleh karena itu digunakan model *Cox Extended* untuk mengatasi permasalahan *Hazard Ratio* yang tidak konstan. Model *Cox Extended* merupakan salah satu metode yang dapat digunakan untuk analisis *survival* dengan satu atau lebih variabel tidak konstan terhadap waktu. Dari permasalahan yang telah diuraikan,

rumusan masalah dalam penelitian ini adalah karakteristik usia pertama melahirkan pada wanita usia subur di Jawa Timur berdasarkan data hasil SDKI 2017, faktor-faktor yang tidak memenuhi asumsi *Proportional Hazard*, dan pemodelan Regresi *Cox Extended* dari faktor-faktor yang berpengaruh terhadap usia pertama melahirkan.

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang akan diselesaikan maka tujuan penelitian yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Mendeskripsikan usia pertama melahirkan dari wanita usia subur di Jawa Timur yang pernah menikah berdasarkan data hasil SDKI 2017.
2. Mendapatkan model Regresi *Cox Extended* dari faktor-faktor yang berpengaruh terhadap usia pertama melahirkan pada wanita usia subur di Jawa Timur yang pernah menikah.

1.4 Manfaat Penelitian

Melalui penelitian ini diharapkan dapat menambah wawasan keilmuan dalam pengembangan dan penerapan analisis *survival* menggunakan beberapa metode baik semiparametrik maupun nonparametrik. Hasil penelitian ini diharapkan dapat menambah wawasan dan memberikan informasi kepada wanita usia subur dan atau masyarakat mengenai faktor-faktor yang signifikan berpengaruh terhadap waktu pertama kali melahirkan pada wanita usia subur di Jawa Timur, serta dapat bermanfaat sebagai bahan masukan bagi para penentu kebijakan dalam merencanakan program kesehatan reproduksi wanita usia subur pada umumnya.

1.5 Batasan Masalah

Batasan masalah yang digunakan dalam penelitian ini adalah sebagai berikut.

1. Data yang digunakan merupakan data hasil Survei Demografi dan Kesehatan Indonesia tahun 2017 pada periode 24 Juli 2017 hingga 30 September 2017 tentang

wanita usia subur di Provinsi Jawa Timur yang berusia 15-49 tahun, pernah menikah, dan bertempat tinggal di region yang sama dengan ketika saat pertama kali melahirkan.

2. *Event* atau kejadian dalam penelitian ini adalah wanita melahirkan pertama kali dalam rentang tahun 1968 hingga 30 September 2017.
3. Tipe data tersensor yang digunakan dalam penelitian ini adalah data sensor kanan (*right censored*).

(Halaman ini sengaja dikosongkan)

BAB II

TINJAUAN PUSTAKA

Bagian ini akan membahas beberapa kajian pustaka yang digunakan dalam penyelesaian masalah terkait analisis *survival* pada usia pertama melahirkan pada Wanita Usia Subur di Jawa Timur.

2.1 Statistika Deskriptif

Statistika merupakan ilmu yang mempelajari tentang cara mengumpulkan, menganalisis, menginterpretasi, dan mengambil kesimpulan berdasarkan hasil analisis. Metode-metode statistika dibagi menjadi dua kelompok besar, yaitu statistika deskriptif dan statistika inferensia. Statistika deskriptif adalah salah satu cabang ilmu statistika yang berkaitan dengan metode-metode pengumpulan dan penyajian data sehingga diperoleh informasi yang berguna. Statistika deskriptif bertujuan untuk menguraikan tentang sifat-sifat atau karakteristik dari suatu fenomena yang diteliti atau dianalisis (Walpole, Myers, Myers, & Ye, 2012). Contoh teknik penyajian data dalam statistika deskriptif berupa tabel, diagram, dan grafik. Statistika deskriptif tidak menarik kesimpulan apapun, hanya memberikan informasi karakteristik data saja.

2.1.1 Mean

Mean atau rata-rata merupakan suatu ukuran khusus yang mewakili sifat tengah atau posisi pusat dari kumpulan data (ukuran pemusatan data). *Mean* atau rata-rata sering disebut dengan nilai tengah atau titik keseimbangan massa dari suatu data (Walpole, Myers, Myers, & Ye, 2012). Rumus untuk menghitung nilai *mean* atau rata-rata disajikan seperti pada persamaan (2.1) sebagai berikut.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (2.1)$$

Keterangan :

\bar{x} : nilai rata-rata

x_i : data ke- i

n : banyak data

2.1.2 Deviasi Standar

Deviasi standar merupakan ukuran penyebaran yang paling banyak digunakan. Deviasi standar tidak berubah apabila setiap unsur pada gugus datanya dikali atau dibagi dengan nilai konstan tertentu. Baik deviasi standar maupun varians mencerminkan konsep yang sama dalam mengukur variabilitas, tetapi deviasi standar mengukur variabilitas dalam satuan linier sedangkan varians diukur dalam satuan kuadrat (Walpole, Myers, Myers, & Ye, 2012). Nilai deviasi standar diperoleh dengan menggunakan (2.2) sebagai berikut.

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (2.2)$$

Keterangan :

s : deviasi standar untuk sampel

2.2 Analisis *Survival*

Analisis *survival* atau analisis ketahanan hidup adalah suatu metode statistik yang berkaitan dengan waktu dari suatu objek yang dimulai dari *time origin* atau *start point* sampai terjadinya *event* tertentu yang telah ditetapkan (*failure event* atau *end point*). Dalam hal ini *event* yang dimaksud adalah kematian, terjangkit penyakit, kambuh dari suatu penyakit, kesembuhan dan kejadian

lainnya yang bisa terjadi pada seseorang. Secara umum tujuan dari analisis *survival* yaitu mengestimasi dan menginterpretasikan fungsi *survival* dan/atau fungsi *hazard* dari data *survival*, membandingkan fungsi *survival* dan/atau fungsi *hazard*, mengetahui pengaruh dari variabel prediktor terhadap waktu *survival* (Kleinbaum & Klein, 2012).

Terdapat dua fokus dalam analisis *survival* yaitu waktu *survival* (T) dan status tersensor (d). Waktu *survival* (T) menunjukkan waktu seorang individu dapat *survive* dalam periode pengamatan tertentu. Suatu *event* dilambangkan dengan simbol d untuk mendefinisikan status *event* apakah *failure* atau tersensor. Nilai $d=1$ menunjukkan *failure* dan $d=0$ menunjukkan tersensor. Penyebab terjadinya data tersensor antara lain:

1. *Termination of study* yaitu berakhirnya penelitian sebelum kejadian *survival* terjadi (*failure*).
2. *Lost of following up* yaitu individu atau objek pengamatan meninggal atau pindah.
3. *Withdraw from the study* yaitu pemberhentian perlakuan karena memberikan efek yang merugikan.

Data tersensor dalam analisis *survival* terbagi menjadi tiga kategori yaitu data tersensor kanan, data tersensor kiri dan data tersensor interval yang dijelaskan seperti sebagai berikut (Kleinbaum & Klein, 2012).

1. Data tersensor kanan (*Right censored*)

Data tersensor kanan adalah data tersensor yang paling sering terjadi dalam analisis *survival*. Data tersensor kanan terjadi apabila tidak diketahui secara pasti waktu *survival* dari individu yang diamati setelah beberapa waktu dilakukan pengamatan. Sehingga pengamatan waktu *survival* terhenti di sebelah kanan periode pengamatan.

2. Data tersensor kiri (*Left censored*)

Data tersensor kiri terjadi apabila *event* yang ingin diamati dari individu terjadi saat waktu pengamatan dimulai namun tidak diketahui kapan pastinya terjadi *event* tersebut. Keadaan tersebut

membuat nilai waktu *survival* kurang dari atau sama dengan waktu pengamatan.

3. Data tersensor interval (*Interval censored*)

Data tersensor interval terjadi apabila *event* yang ingin diamati dari individu terjadi diantara dua waktu observasi tertentu. Misalnya saja pada pengamatan pertama belum terjadi *event* pada individu dan pada pengamatan berikutnya dengan selang waktu tertentu telah terjadi *event* pada individu tersebut sehingga tidak tau kapan persisnya *event* terjadi.

2.2.1 Fungsi *Survival* dan Fungsi *Hazard*

Terdapat dua kuantitas dasar yang sering digunakan dalam analisis *survival* yaitu fungsi *survival* yang dilambangkan dengan $S(t)$ dan fungsi *hazard* dilambangkan dengan $h(t)$. Fungsi *survival* didefinisikan sebagai probabilitas individu dapat bertahan lebih dari waktu ke- t , sedangkan fungsi *hazard* didefinisikan sebagai laju terjadinya *event* setelah individu bertahan hingga waktu ke- t . Jika T merupakan variabel random kontinu yang memiliki distribusi peluang $f(t)$, maka fungsi kepadatan peluang dapat dinyatakan seperti pada persamaan (2.3) berikut (Kleinbaum & Klein, 2012).

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t} \quad (2.3)$$

Sedangkan fungsi distribusi kumulatif dapat dirumuskan persamaan (2.4) berikut.

$$F(t) = P(T \leq t) = \int_0^t f(t) dt \quad (2.4)$$

Maka, fungsi *survival* dinyatakan dengan persamaan (2.5) berikut.

$$S(t) = P(T > t) = 1 - P(T \leq t) = 1 - F(t) \quad (2.5)$$

Fungsi *hazard* dapat diartikan sebagai kebalikan dari fungsi *survival* yang ditunjukkan pada persamaan (2.8). Fungsi *hazard* dinyatakan dengan persamaan (2.6) berikut.

$$h(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \right\} \quad (2.6)$$

Dengan menggunakan teori probabilitas bersyarat maka diperoleh persamaan (2.7) berikut.

$$h(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{P((t \leq T < t + \Delta t) \cap (T \geq t))}{\Delta t \times P(T \geq t)} \right\} \quad (2.7)$$

$$h(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t \leq T < t + \Delta t)}{\Delta t \times S(t)} \right\}$$

$$h(t) = \frac{1}{S(t)} \times \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t \leq T < t + \Delta t)}{\Delta t} \right\}$$

$$h(t) = \frac{f(t)}{S(t)} \quad (2.8)$$

Dengan $F(t) = 1 - S(t)$ maka $f(t) = \frac{d(F(t))}{dt} = \frac{d(1 - S(t))}{dt}$ sehingga diperoleh fungsi *hazard* $h(t)$ pada persamaan (2.9) berikut.

$$h(t) = \frac{\left(\frac{d(1 - S(t))}{dt} \right)}{S(t)} \quad (2.9)$$

$$h(t) = \frac{\left(\frac{-d(S(t))}{dt} \right)}{S(t)}$$

$$-h(t)dt = \frac{1}{S(t)} d(S(t)) \quad (2.10)$$

Dengan mengintegrasikan kedua ruas fungsi tersebut seperti pada persamaan (2.11), maka diperoleh persamaan (2.12) berikut.

$$-\int_0^t h(t)dt = \int_0^t \frac{1}{S(t)} d(S(t)) \quad (2.11)$$

$$-\int_0^t h(t)dt = \ln S(t) - \ln S(0)$$

$$-\int_0^t h(t)dt = \ln S(t) \quad (2.12)$$

Fungsi distribusi kumulatif dari fungsi *hazard* dan hubungan antara fungsi *hazard* dan fungsi *survival* dapat dirumuskan persamaan (2.14) berikut.

$$H(t) = \int_0^t h(t)dt \quad (2.13)$$

$$H(t) = -\ln S(t) \quad (2.14)$$

2.2.2 Kurva Kaplan Meier

Kurva *Kaplan Meier* merupakan suatu metode statistika pada analisis data *survival* yang digunakan untuk menghitung peluang *survival* dengan menggambarkan hubungan antara estimasi fungsi *survival* pada waktu t dengan waktu *survival*. Metode *Kaplan Meier* didasarkan pada waktu kelangsungan hidup subjek dan mengasumsikan bahwa data sensor adalah prediktor berdasarkan waktu kelangsungan hidup (yaitu, alasan observasi yang disensor tidak berhubungan dengan penyebab *failure time*) (Kleinbaum & Klein, 2012). Metode *Kaplan Meier* berdasarkan dua konsep sederhana, yaitu jika probabilitas dari *Kaplan Meier* dinotasikan dengan $\hat{S}(t_{(f)})$ maka persamaan umum *Kaplan Meier* adalah sebagai berikut.

$$\hat{S}(t_{(f)}) = \hat{S}(t_{(f-1)}) \times \hat{P}_r(T > t_{(f)} | T \geq t_{(f)}) \quad (2.15)$$

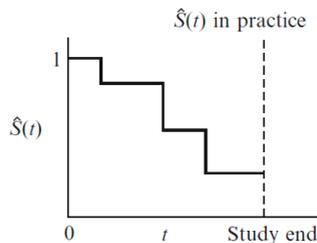
Dimana $\hat{S}(t_{(f-1)})$ merupakan probabilitas suatu objek bertahan sampai waktu ke $t_{(f-1)}$ dan $\hat{P}_r(T > t_{(f)} | T \geq t_{(f)})$ probabilitas objek bertahan sampai setelah waktu ke $t_{(f)}$ dengan syarat objek mampu bertahan hidup tidak kurang dari $t_{(f)}$.

$$\hat{S}(t_{(f-1)}) = \prod_{i=1}^{f-1} \hat{P}_r(T > t_{(i)} | T \geq t_{(i)}) \quad (2.16)$$

Persamaan umum kurva *Kaplan Meier* atau estimator *Kaplan Meier* dirumuskan seperti persamaan (2.17) berikut.

$$\hat{S}(t_{(f)}) = \prod_{i=1}^f \hat{P}_r(T > t_{(i)} | T \geq t_{(i)}) \quad (2.17)$$

Secara visual kurva *Kaplan Meier* memiliki sumbu vertikal yang menggambarkan estimasi fungsi *survival* dan sumbu horizontal yang menggambarkan waktu *survival*. Berikut ini merupakan ilustrasi dari kurva *Kaplan Meier*.



Sumber: Buku *Survival Analysis* (Kleinbaum & Klein, 2012)

Gambar 2.1 Ilustrasi Kurva *Kaplan Meier*.

2.2.3 Uji *Log-Rank*

Uji *Log-Rank* adalah uji statistik nonparametrik yang digunakan ketika data tidak simetris. Uji *Log-Rank* digunakan untuk membandingkan kurva *Kaplan Meier* dalam kelompok yang

berbeda. Berikut merupakan hipotesis dan statistik uji dengan pendekatan rumus pada persamaan (2.18) yang digunakan pada pengujian *Log-Rank* (Kleinbaum & Klein, 2012).

H_0 : tidak terdapat perbedaan pada kurva *survival* antara kelompok yang berbeda

H_1 : minimal terdapat satu perbedaan pada kurva *survival* antara kelompok yang berbeda

Statistik uji:

$$X^2 \approx \sum_{g=1}^G \frac{(O_g - E_g)^2}{E_g} \quad (2.18)$$

dengan,

$$O_g - E_g = \sum_{i=1}^n (m_{gi} - \hat{e}_{gi}) \quad (2.19)$$

$$\hat{e}_{gi} = \left(\frac{n_{gi}}{\sum_{g=1}^G n_{gi}} \right) \left(\sum_{g=1}^G m_{gi} \right) \quad (2.20)$$

Daerah kritis: Tolak H_0 apabila nilai $X^2 > \chi_{\alpha, (G-1)}^2$.

Keterangan:

O_g : nilai observasi individu kelompok ke- g

E_g : nilai ekspektasi individu kelompok ke- g

m_{gi} : jumlah individu yang mengalami *event* dalam kelompok ke- g pada waktu $t_{(i)}$

n_{gi} : jumlah individu yang beresiko gagal seketika dalam kelompok ke- i sebelum waktu $t_{(i)}$

\hat{e}_{gi} : nilai ekspektasi dalam kelompok ke- g pada waktu $t_{(i)}$

G : banyak kelompok

2.2.4 Regresi *Cox Proportional Hazard*

Regresi *Cox* merupakan salah satu analisis *survival* yang pertama kali diperkenalkan oleh Cox dan respon yang digunakan merupakan data yang diperoleh dari perhitungan waktu suatu

peristiwa tertentu (Cox, 1972). Model *Cox Proportional Hazard* (*Cox PH*) disebut model semiparametrik yang dapat digunakan ketika data yang ada tidak diketahui jelas distribusinya. Hal ini dikarenakan meskipun *baseline hazard* tidak diketahui bentuk fungsionalnya, *Cox PH* tetap dapat memberikan informasi yang berguna berupa *hazard ratio* (Aini, 2011). Model regresi *Cox Proportional Hazard* ditunjukkan pada persamaan (2.21) berikut.

$$h(t_i, \mathbf{x}_i) = h_0(t_i) \exp\left(\sum_{j=1}^p \beta_j x_{ij}\right); \mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip}) \quad (2.21)$$

Keterangan:

$h_0(t_i)$: *baseline hazard* pada waktu ke- t_i

β_j : parameter regresi ke- j

x_{ij} : nilai dari variabel prediktor ke- j dari individu ke- i

Model terbaik diperoleh dengan melakukan estimasi koefisien variabel prediktor X_1, X_2, \dots, X_p yaitu $\beta_1, \beta_2, \dots, \beta_p$. Parameter atau koefisien regresi β pada model *Cox PH* dapat diestimasi menggunakan metode *maximum partial likelihood estimation* (MPLE). Waktu *survival* terurut r individu yang mengalami *event* dari n individu dilambangkan $t_{(1)} < t_{(2)} < \dots < t_{(i)} < \dots < t_{(r)}$ dan $t_{(i)}$ merupakan urutan waktu *event* ke- i . $R(t_{(i)})$ merupakan himpunan dari individu yang berisiko mengalami *event* pada waktu $t_{(i)}$ yang terdiri dari individu-individu yang bertahan hidup hingga $t_{(i)}$. $\mathbf{x}_{(i)}$ merupakan vektor variabel dari individu yang mengalami *event* pada waktu $t_{(i)}$. Fungsi parsial *likelihood* model *Cox PH* dapat dituliskan persamaan (2.22) berikut.

$$L(\boldsymbol{\beta}) = \prod_{i=1}^r \frac{\exp\left(\sum_{j=1}^p \beta_j x_{ij}\right)}{\sum_{l \in R(i)} \exp\left(\sum_{j=1}^p \beta_j x_{ij}\right)} \quad (2.22)$$

Fungsi *ln-likelihood* dari persamaan (2.22) dapat diuraikan sebagai berikut.

$$\ln L(\boldsymbol{\beta}) = \sum_{i=1}^r \sum_{j=1}^p \beta_j x_{ij} - \sum_{i=1}^r \ln \left[\sum_{l \in R(i)} \exp\left(\sum_{j=1}^p \beta_j x_{il}\right) \right] \quad (2.23)$$

Langkah selanjutnya adalah memaksimumkan turunan pertama dari persamaan *ln-likelihood* dengan menggunakan metode Newton-Rhapon (Collett, 2003). Berikut disajikan langkah-langkah memaksimuman menggunakan metode Newton-Rhapon.

1. Menentukan *initial value* parameter.

$$\hat{\boldsymbol{\beta}}_0 = [\beta_{10} \beta_{20} \dots \beta_{p0}]; \text{ iterasi pada } k=0.$$

2. Membentuk vektor *gradient*.

$$\nabla L'(\hat{\boldsymbol{\beta}}_{(k)})_p = \left[\frac{\partial \ln L(\cdot)}{\partial \beta_1} \quad \frac{\partial \ln L(\cdot)}{\partial \beta_2} \quad \frac{\partial \ln L(\cdot)}{\partial \beta_3} \quad \dots \quad \frac{\partial \ln L(\cdot)}{\partial \beta_p} \right]$$

3. Membentuk matriks *Hessian*.

$$\mathbf{H}(\hat{\boldsymbol{\beta}}_{(k)})_{p \times p} = \begin{bmatrix} \frac{\partial^2 L(\cdot)}{\partial \beta_1^2} & \frac{\partial^2 L(\cdot)}{\partial \beta_1 \partial \beta_2} & \dots & \frac{\partial^2 L(\cdot)}{\partial \beta_1 \partial \beta_p} \\ \frac{\partial^2 L(\cdot)}{\partial \beta_2^2} & \dots & \frac{\partial^2 L(\cdot)}{\partial \beta_2 \partial \beta_p} & \\ \vdots & & & \\ \dots & \frac{\partial^2 L(\cdot)}{\partial \beta_p^2} & & \end{bmatrix}$$

4. Melakukan iterasi mulai dari iterasi ke-0 ($k = 0$).

$$\hat{\boldsymbol{\beta}}_{(k+1)} = \hat{\boldsymbol{\beta}}_{(k)} - \mathbf{H}^{-1}(\hat{\boldsymbol{\beta}}_{(k)}) \nabla L(\hat{\boldsymbol{\beta}}_{(k)})$$

5. Iterasi akan berhenti jika $\|\hat{\beta}_{(k+1)} - \hat{\beta}_{(k)}\| \leq \varepsilon$ dimana ε adalah bilangan yang sangat kecil.

2.2.5 Asumsi *Proportional Hazard*

Asumsi *proportional hazard* didefinisikan sebagai suatu keadaan dimana *hazard ratio* dari model *Cox Proportional Hazard* bersifat konstan terhadap waktu (Kleinbaum & Klein, 2012). Pengujian asumsi *proportional hazard* dapat dilakukan dengan tiga pendekatan, yaitu grafik $-\ln(-\ln S(t))$, uji *Goodness of Fit*, dan variabel *time dependent*. Tetapi penelitian ini hanya menggunakan uji *Goodness of Fit*.

Pengujian asumsi *proportional hazard* dengan menggunakan statistik yaitu uji *goodness of fit* dan pada uji *goodness of fit* dihasilkan *p-value*. Berikut disajikan langkah-langkah pengujian asumsi *proportional hazard* dengan uji *goodness of fit*.

1. Mendapatkan residual *schoenfeld* untuk setiap variabel prediktor dari model *Cox PH*. Residual *schoenfeld* ada pada setiap variabel prediktor pada model dan pada setiap observasi yang mengalami *event*.
2. Membuat variabel *rank* waktu *survival* yang diurutkan dari kecil ke besar.
3. Menghitung korelasi antara residual *schoenfeld* dan *rank* waktu *survival* (r_{RT,PR_i}) menggunakan formula persamaan (2.24) berikut.

$$r_{RT,PR_i} = \frac{\sum_{j=1}^n (PR_{ij} - \overline{PR}_i)(RT_j - \overline{RT})}{\sqrt{\sum_{j=1}^n (PR_{ij} - \overline{PR}_i)^2} \sqrt{\sum_{j=1}^n (RT_j - \overline{RT})^2}} \quad (2.24)$$

Keterangan :

PR_{ij} : residual *schoenfeld* untuk variabel prediktor ke-*i* individu yang mengalami *event* pada $t_{(j)}$.

\overline{PR}_i : rata-rata nilai residual *schoenfeld* untuk variabel prediktor ke-*i*

RT_j : rank waktu *survival* yang diurutkan dari kecil ke besar pada $t_{(j)}$

\overline{RT} : rata-rata rank waktu *survival*

4. Melakukan pengujian korelasi antara residual *schoenfeld* dan rank waktu *survival*. Hipotesis yang digunakan dalam pengujian korelasi adalah sebagai berikut.

H_0 : $\rho = 0$

H_1 : $\rho \neq 0$

Dengan statistik uji,

$$t_{hit} = \frac{r_{RT, PR_i} \sqrt{n-2}}{\sqrt{1-r_{RT, PR_i}^2}} \quad (2.25)$$

Tolak H_0 jika $|t_{hit}| > t_{\alpha/2, n-2}$ atau nilai *p-value* kurang dari taraf signifikansi (α) yang digunakan. Asumsi *propotional hazard* terpenuhi jika uji korelasi antara residual *schoenfeld* dan rank waktu *survival* tidak signifikan atau gagal tolak H_0 (Kleinbaum & Klein, 2012).

2.2.6 Model *Cox Extended*

Model *Cox extended* merupakan salah satu metode yang digunakan jika terdapat variabel prediktor bergantung pada waktu (*time dependent*) sehingga menyebabkan asumsi *proportional hazard* tidak terpenuhi. Variabel prediktor yang nilainya dapat berubah setiap saat bergantung kepada waktu disebut variabel yang bergantung terhadap waktu (Kleinbaum & Klein, 2012). Pada model *Cox extended*, variabel prediktor yang bergantung terhadap waktu diinteraksikan dengan fungsi waktu $g(t)$. Fungsi waktu yang digunakan berdasarkan variabel prediktor yang tidak memenuhi asumsi *proportional hazard* dengan *p-value* terkecil.

Apabila terdapat q variabel prediktor yang tidak memenuhi asumsi *proportional hazard*, maka sebanyak q variabel yang harus diinteraksikan dengan fungsi waktu, dan apabila fungsi waktu yang digunakan adalah fungsi waktu T atau $\ln(T)$ maka model *Cox extended* dapat dituliskan seperti persamaan (2.26) berikut.

$$h(t_i, \mathbf{x}_i) = h_0(t_i) \exp \left[\sum_{j=1}^s \beta_j x_{ij} + \sum_{k=1}^q \beta_k^* x_{ik}^* + \sum_{k=1}^q \delta_k x_{ik}^* g_k(t_i) \right] \quad (2.26)$$

$$\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})$$

Keterangan:

$h_0(t_i)$: *baseline hazard* pada waktu ke- t

β_j : parameter regresi dari variabel prediktor *time independent* ke- j

β_k^* : parameter regresi dari variabel prediktor *time dependent* ke- k

δ_k : parameter regresi dari variabel prediktor *time dependent* ke- k yang telah diinteraksikan dengan fungsi waktu $Ln(t)$.

x_{ij} : nilai dari variabel prediktor *time independent* ke- j dari individu ke- i

x_{ik}^* : nilai dari variabel prediktor *time dependent* ke- k dari individu ke- i

$g_k(t_i)$: fungsi waktu $g(t)$ pada indivisu ke- i yang diinteraksikan dengan variabel *time dependent* ke- k

s : banyaknya variabel prediktor *time independent*

q : banyaknya variabel prediktor *time dependent*

p : banyaknya variabel prediktor

Apabila model *Cox extended* diasumsikan terdapat satu variabel *time dependent* dengan f kategori yang menggunakan *heavyside function* sebagai fungsi waktu maka dapat dituliskan persamaan (2.27) berikut ini.

$$h(t_i, \mathbf{x}_i) = h_0(t_i) \exp \left[\sum_{j=1}^s \beta_j x_{ij} + \sum_{m=1}^{f-1} \delta_{1m}^{**} x_{im}^{**} g_{11}(t_i) + \sum_{m=1}^{f-1} \delta_{2m}^{**} x_{im}^{**} g_{21}(t_i) \right] \quad (2.27)$$

$$\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})$$

Keterangan:

δ_{1m}^{**} : parameter regresi dari variabel prediktor *time dependent* kategori ke- m yang diinteraksikan dengan fungsi waktu $g_{11}(t_i)$

δ_{2m}^{**} : parameter regresi dari variabel prediktor *time dependent* kategori ke- m yang diinteraksikan dengan fungsi waktu $g_{21}(t_i)$

x_{im}^{**} : nilai dari variabel prediktor *time dependent* kategori ke- m dari individu ke- i

f : banyaknya kategori dari variabel prediktor *time dependent*

Apabila model *Cox extended* menggunakan kombinasi fungsi waktu $g(t)$ yang berbeda, misalnya kombinasi satu variabel prediktor yang diinteraksikan dengan fungsi waktu $\ln(T)$ dan satu variabel lainnya diinteraksikan dengan fungsi waktu *heavyside*, maka model *Cox extended* dapat dituliskan seperti pada persamaan (2.28) sebagai berikut.

$$h(t, \mathbf{x}_i) = h_0(t_i) \exp \left[\sum_{j=1}^s \beta_j x_{ij} + \sum_{k=1}^q \beta_k^* x_{ik}^* + \sum_{m=1}^{f-1} \delta_{1m}^{**} x_{im}^{**} g_{12}(t_i) + \sum_{k=1}^q \delta_k x_{ik}^* g_k(t_i) + \sum_{m=1}^{f-1} \delta_{2m}^{**} x_{im}^{**} g_{22}(t_i) \right] \quad (2.28)$$

$$\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})$$

Menurut Collett (2003) pendugaan parameter pada model *Cox extended* menggunakan metode *maximum partial likelihood estimation*. Persamaan (2.29) merupakan fungsi parsial *likelihood* model *Cox extended* dari persamaan (2.28) seperti berikut ini.

$$L(\boldsymbol{\beta}) = \prod_{i=1}^r \frac{\exp \left(\sum_{j=1}^s \beta_j x_{ij} + \sum_{k=1}^q \beta_k^* x_{ik}^* + \sum_{m=1}^{f-1} \delta_{1m}^{**} x_{im}^{**} g_{12}(t) + \sum_{k=1}^q \delta_k x_{ik}^* g_k(t) + \sum_{m=1}^{f-1} \delta_{2m}^{**} x_{im}^{**} g_{22}(t) \right)}{\sum_{i \in R(t_i)} \exp \left(\sum_{j=1}^s \beta_j x_{ij} + \sum_{k=1}^q \beta_k^* x_{ik}^* + \sum_{m=1}^{f-1} \delta_{1m}^{**} x_{im}^{**} g_{12}(t) + \sum_{k=1}^q \delta_k x_{ik}^* g_k(t) + \sum_{m=1}^{f-1} \delta_{2m}^{**} x_{im}^{**} g_{22}(t) \right)} \quad (2.29)$$

Sedangkan persamaan (2.30) merupakan persamaan *log-likelihood* ($\ln L(\boldsymbol{\beta})$) yang digunakan untuk mengestimasi parameter model *Cox extended*.

$$\ln L(\boldsymbol{\beta}) = \sum_{i=1}^r \left(\sum_{j=1}^s \beta_j x_{ij} + \sum_{k=1}^q \beta_k^* x_{ik}^* + \sum_{m=1}^{f-1} \delta_{1m}^{**} x_{im}^{**} g_{12}(t) + \sum_{k=1}^q \delta_k x_{ik}^* g_k(t) + \sum_{m=1}^{f-1} \delta_{2m}^{**} x_{im}^{**} g_{22}(t) \right) - \sum_{i=1}^r \ln \left[\sum_{i \in R(t_i)} \exp \left(\sum_{j=1}^s \beta_j x_{ij} + \sum_{k=1}^q \beta_k^* x_{ik}^* + \sum_{m=1}^{f-1} \delta_{1m}^{**} x_{im}^{**} g_{12}(t) + \sum_{k=1}^q \delta_k x_{ik}^* g_k(t) + \sum_{m=1}^{f-1} \delta_{2m}^{**} x_{im}^{**} g_{22}(t) \right) \right] \quad (2.30)$$

Selanjutnya persamaan 2.30 dimaksimumkan dengan menggunakan metode Newton-Rhapon untuk mendapatkan

estimasi parameter model *Cox extended* dengan langkah-langkah sebagai berikut.

1. Menentukan *initial value* parameter.

$$\hat{\theta}_0 = \left[\beta_{10} \dots \beta_{s0} \beta_{10}^* \dots \beta_{q0}^* \delta_{110}^{**} \dots \delta_{1(f-1)0}^{**} \delta_{10} \dots \delta_{q0} \dots \delta_{210}^{**} \dots \delta_{2(f-1)0}^{**} \right];$$

iterasi pada $c=0$.

2. Membentuk vektor *gradient*.

$$\nabla L(\hat{\theta}_{(c)})_p = \left[\frac{\partial \ln L(.)}{\partial \beta_1} \dots \frac{\partial \ln L(.)}{\partial \beta_s} \frac{\partial \ln L(.)}{\partial \beta_1^*} \dots \frac{\partial \ln L(.)}{\partial \beta_q^*} \frac{\partial \ln L(.)}{\partial \delta_{11}^{**}} \dots \frac{\partial \ln L(.)}{\partial \delta_{1(f-1)}^{**}} \frac{\partial \ln L(.)}{\partial \delta_1} \dots \frac{\partial \ln L(.)}{\partial \delta_q} \frac{\partial \ln L(.)}{\partial \delta_{21}^{**}} \dots \frac{\partial \ln L(.)}{\partial \delta_{2(f-1)}^{**}} \right]$$

3. Membentuk matriks *Hessian*.

$$\mathbf{H}(\hat{\theta}_{(c)})_{p \times p} = \begin{bmatrix} \frac{\partial^2 L(.)}{\partial \beta_1^2} & \frac{\partial^2 L(.)}{\partial \beta_1 \partial \beta_2} & \dots & \frac{\partial^2 L(.)}{\partial \beta_1 \partial \delta_{2(f-1)}^{**}} \\ \frac{\partial^2 L(.)}{\partial \beta_2^2} & \dots & \frac{\partial^2 L(.)}{\partial \beta_2 \partial \delta_{2(f-1)}^{**}} & \\ & & \ddots & \vdots \\ & & & \frac{\partial^2 L(.)}{\partial \delta_{2(f-1)}^{**2}} \end{bmatrix}$$

4. Melakukan iterasi mulai dari iterasi ke-0 ($c = 0$).

$$\hat{\theta}_{(c+1)} = \hat{\theta}_{(c)} - \mathbf{H}^{-1}(\hat{\theta}_{(c)}) \nabla L(\hat{\theta}_{(c)})$$

5. Iterasi akan berhenti jika $\|\hat{\theta}_{(c+1)} - \hat{\theta}_{(c)}\| \leq \varepsilon$ dimana ε adalah bilangan yang sangat kecil.

Nilai fungsi *hazard* untuk model *Cox extended* dapat diestimasi dengan menggunakan persamaan (2.31) berikut.

$$\hat{h}_0(t) = -\ln \hat{S}_0(t) = \sum_{i=1}^s \frac{1}{\sum_{j \in R(t)} \exp \left(\sum_{j=1}^s \hat{\beta}_j x_{ij} + \sum_{k=1}^q \hat{\beta}_k^* x_{ik}^* + \sum_{m=1}^{f-1} \hat{\delta}_{1m}^{**} g_{12}(t) + \sum_{k=1}^q \hat{\delta}_k x_{ik} g_1(t) + \sum_{m=1}^{f-1} \hat{\delta}_{2m}^{**} g_{22}(t) \right)} \quad (2.31)$$

Penentuan bentuk fungsi waktu $g(t)$ harus tepat dan sesuai dengan kondisi kasus, sehingga sangat penting menentukan bentuk $g(t)$ yang tepat. Menurut Kleinbaum dan Klein (2012) bahwa terdapat beberapa kemungkinan bentuk fungsi waktu $g(t)$ yang dijelaskan sebagai berikut.

- $g(t) = 0$ menghasilkan model *Cox Proportional Hazard*
- $g(t) = t$ jika hasil pengujian signifikan maka model *Cox extended* lebih baik daripada model *Cox Proportional Hazard* sehingga *hazard ratio* merupakan fungsi terhadap waktu.
- $g(t) = \ln(t)$
- $g(t)$ *heavyside function*, *hazard ratio* akan konstan untuk interval waktu yang berbeda. Menurut Aini (2011) fungsi *heavyside* digunakan jika *hazard ratio* berubah pada waktu tertentu atau bernilai konstan pada interval waktu tertentu namun berbeda antar selang waktu.

Sesuai dengan kondisi kasus pada persamaan (2.28), misalkan *hazard ratio* berubah setelah interval waktu t_0 , maka fungsi *heavyside* dapat dituliskan seperti persamaan (2.32) berikut.

$$g_{12}(t) = \begin{cases} 1 & ; t < t_0 \\ 0 & ; t \geq t_0 \end{cases} \quad \text{dan} \quad g_{22}(t) = \begin{cases} 1 & ; t \geq t_0 \\ 0 & ; t < t_0 \end{cases} \quad (2.32)$$

Secara umum *Hazard ratio* dari beberapa kemungkinan fungsi waktu $g_k(t)$ dapat disajikan sebagai berikut.

Tabel 2. 1 Hazard Ratio Masing-masing $g_m(t)$

Fungsi Waktu	Interval	Hazard Ratio
$g_k(t) = 0$	$\forall t$	$\hat{HR} = \exp(\hat{\beta}_j)$
$g_k(t) = t$	$\forall t$	$\hat{HR} = \exp(\hat{\beta}_k^* + \hat{\delta}_k t)$
$g_k(t) = \ln(t)$	$\forall t$	$\hat{HR} = \exp(\hat{\beta}_k^* + \hat{\delta}_k \ln t)$
$g_{12}(t) = \begin{cases} 1 \\ 0 \end{cases}$ dan $g_{22}(t) = \begin{cases} 0 \\ 1 \end{cases}$	$t < t_0$	$\hat{HR} = \exp(\hat{\delta}_{1m}^{**})$
	$t \geq t_0$	$\hat{HR} = \exp(\hat{\delta}_{2m}^{**})$

2.2.7 Pengujian Signifikansi Parameter

Pengujian signifikansi parameter digunakan untuk mengetahui variabel prediktor yang signifikan berpengaruh terhadap model atau fungsi *survival*. Pengujian signifikansi parameter dilakukan secara serentak dan parsial. Berikut ini adalah pengujian signifikansi parameter secara serentak maupun parsial.

a) Uji Serentak

Uji serentak digunakan untuk mengetahui apakah variabel prediktor yang digunakan pada model berpengaruh signifikan secara serentak atau bersama-sama. Hipotesis pengujian signifikansi parameter secara serentak adalah sebagai berikut.

$$H_0 : \beta_1 = \dots = \beta_s = \beta_1^* = \dots = \beta_q^* = \delta_{11}^{**} = \dots = \delta_{1(f-1)}^{**} = \delta_1 = \dots = \delta_q = \delta_{21}^{**} = \dots = \delta_{2(f-1)}^{**} = 0$$

(tidak terdapat variabel yang berpengaruh signifikan terhadap usia pertama melahirkan)

$$H_1 : \text{Minimal ada satu } \theta_v \neq 0 ; v = 1, 2, \dots, p;$$

$$\theta_v = (\beta_1, \dots, \beta_s, \beta_1^*, \dots, \beta_q^*, \delta_{11}^{**}, \dots, \delta_{1(f-1)}^{**}, \delta_1, \dots, \delta_q, \delta_{21}^{**}, \dots, \delta_{2(f-1)}^{**})$$

(minimal ada satu variabel yang berpengaruh signifikan terhadap usia pertama melahirkan)

Statistik uji:

$$G_{hit}^2 = -2 \ln \left(\frac{\max_{\omega} \ln L(\omega)}{\max_{\Omega} \ln L(\Omega)} \right) = -2 \ln \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) \quad (2.33)$$

dengan

$$L(\hat{\Omega}) = \prod_{i=1}^f \frac{\exp \left(\sum_{j=1}^s \hat{\beta}_j x_{ij} + \sum_{k=1}^q \hat{\beta}_k^* x_{ik}^* + \sum_{m=1}^{f-1} \hat{\delta}_{1m}^{**} x_{im}^{**} g_{12}(t) + \sum_{k=1}^q \hat{\delta}_k x_{ik}^* g_k(t) + \sum_{m=1}^{f-1} \hat{\delta}_{2m}^{**} x_{im}^{**} g_{22}(t) \right)}{\sum_{l \in R(i)} \exp \left(\sum_{j=1}^s \hat{\beta}_j x_{ij} + \sum_{k=1}^q \hat{\beta}_k^* x_{ik}^* + \sum_{m=1}^{f-1} \hat{\delta}_{1m}^{**} x_{im}^{**} g_{12}(t) + \sum_{k=1}^q \hat{\delta}_k x_{ik}^* g_k(t) + \sum_{m=1}^{f-1} \hat{\delta}_{2m}^{**} x_{im}^{**} g_{22}(t) \right)} \quad (2.34)$$

$$L(\hat{\omega}) = \prod_{i=1}^f \frac{n_{(i)} - m_{(i)}}{n_{(i)}} \quad (2.35)$$

Daerah Kritis: Tolak H_0 , jika $G_{hit}^2 > \chi_{\alpha, p}^2$

Keterangan:

$L(\hat{\omega})$: nilai *likelihood* untuk model tanpa menyertakan variabel prediktor

$L(\hat{\Omega})$: nilai *likelihood* untuk model dengan menyertakan semua variabel prediktor

p : banyak parameter dalam model

$n_{(i)}$: jumlah individu yang beresiko mengalami *event* sebelum waktu $t_{(i)}$

$m_{(i)}$: jumlah *event* pada waktu $t_{(i)}$

b) Uji Parsial

Secara parsial, hipotesis uji signifikansi parameter secara umum adalah sebagai berikut.

Hipotesis:

$H_0 : \theta_v = 0$ (variabel prediktor ke- v tidak berpengaruh signifikan terhadap variabel respon)

$H_1 : \theta_v \neq 0$ (variabel prediktor ke- v berpengaruh signifikan terhadap variabel respon)

Statistik Uji:

$$W^2 = \frac{(\hat{\theta}_v)^2}{(SE(\hat{\theta}_v))^2} \quad (2.36)$$

Daerah kritis: Tolak H_0 jika $W^2 > \chi^2_{(\alpha,1)}$ atau $p\text{-value} < \alpha$.

2.2.8 Pemilihan Model Terbaik

Seleksi model dilakukan untuk mendapatkan model terbaik sehingga mengetahui gambaran hubungan antar waktu *survival* dengan variabel prediktor secara tepat. *Akaike Information Criterion* (AIC) merupakan suatu kriteria pemilihan model untuk menganalisis data aktual (Konishi & Kitagawa, 2008). Kriteria AIC berguna untuk mengevaluasi pemodelan regresi yang bertujuan untuk mengidentifikasi faktor-faktor yang berpengaruh terhadap model (Akaike, Information Theory as an Extension of The Maximum Likelihood, 1973). AIC biasanya digunakan dalam berbagai bidang seperti ilmu alam, ilmu sosial, dan teknik (Konishi & Kitagawa, 2008). Persamaan (2.37) berikut ini merupakan persamaan AIC.

$$AIC = -2\log L(\hat{\theta}) + 2p \quad (2.37)$$

Dimana L merupakan nilai *Likelihood* sedangkan p adalah banyaknya parameter. Sedangkan AICc lebih general dibanding AIC dan umumnya digunakan sebagai pengganti AIC. Perhitungan nilai AICc pada persamaan (2.38) sebagai berikut (Burnham & Anderson, 2002).

$$AICc = AIC + \frac{2p(p+1)}{n-p-1} \quad (2.38)$$

dengan n merupakan jumlah unit penelitian. Model terbaik adalah model dengan kriteria nilai AIC dan AICc terkecil.

2.2.9 Hazard Ratio

Hazard Ratio adalah suatu ukuran yang digunakan untuk mengetahui tingkat risiko dari perbandingan antara individu dengan kondisi variabel prediktor pada kategori sukses dengan kategori gagal (Hosmer, Lameshow, & May, 2008). Nilai *hazard ratio* diperoleh menggunakan persamaan (2.40) dan didapatkan nilai *hazard ratio* dari model *Cox proportional hazard* pada persamaan (2.41) berikut.

$$\hat{HR} = \frac{\hat{h}(t, x_j^n)}{\hat{h}(t, x_j)} \quad (2.39)$$

$$\hat{HR} = \exp \left[\sum_{j=1}^p \hat{\beta}_j (x_j^n - x_j) \right] \quad (2.40)$$

Dimana $j = 1, 2, \dots, p$ dan x_j adalah nilai dari variabel prediktor ke- j . Rumus *hazard ratio* secara umum dari model *Cox extended* pada persamaan ditunjukkan oleh persamaan (2.42) berikut.

$$\hat{HR} = \frac{\hat{h}(t, x_v^n(t))}{\hat{h}(t, x_v(t))} \quad (2.41)$$

$$\hat{HR} = \exp \left[\sum_{j=1}^s \hat{\beta}_j (x_j^n - x_j) + \sum_{k=1}^q \hat{\beta}_k (x_k^{**} - x_k^*) + \sum_{m=1}^{f-1} \hat{\delta}_{1m}^{**} (x_m^{**}(t) - x_m^*(t)) + \sum_{k=1}^q \hat{\delta}_k^{**} (x_k^{**}(t) - x_k^*(t)) + \sum_{m=1}^{f-1} \hat{\delta}_{2m}^{**} (x_m^{**}(t) - x_m^*(t)) \right] \quad (2.42)$$

Dimana $j = 1, 2, \dots, s$; $k = 1, 2, \dots, q$; dan $m = 1, 2, \dots, (f-1)$. x_v adalah nilai dari variabel prediktor ke- v yang mana $v = 1, 2, \dots, p$.

2.3 Usia Pertama Melahirkan Beserta Faktor-faktor yang Diduga Berpengaruh

Usia pertama melahirkan memiliki konsep yang tidak jauh berbeda dengan konsep awal mula subjek untuk fertilitas. Fertilitas bagian dari perilaku sosial yang memiliki konsep yang kompleks dan rinci. Konsep fertilitas tidak terbatas pada satu fase saja, namun berpengaruh pada keseluruhan siklus kehidupan manusia. Penjelasan menurut Matthew (2012) dan Sri (2010) mendefinisikan fertilitas merupakan istilah demografi yang diartikan sebagai hasil reproduksi yang nyata dari seorang wanita berupa jumlah anak yang pernah dilahirkan sebelum mencapai usia 50 tahun. Jumlah anak yang pernah dilahirkan oleh seorang perempuan menurut Adhikari (2010) ialah pengalaman melahirkan anak dari kelompok usia nyata yang mencerminkan perilaku fertilitas saat ini dan masa lalu. Jumlah anak lahir hidup dan lebih mudah dihitung untuk wanita, karena yang menjalani proses melahirkan. Anak yang dilahirkan dengan keadaan mati, tidak dapat dikatakan sebagai fertilitas.

Menurut Rusli (1996) bahwa fertilitas dipengaruhi dan ditentukan oleh dua faktor, yaitu faktor demografi dan non demografi. Faktor demografi meliputi umur, umur perkawinan pertama, lama perkawinan, paritas atau jumlah persalinan yang pernah dialami dan proporsi perkawinan, sedangkan faktor non demografi meliputi keadaan ekonomi penduduk, tingkat pendidikan, perbaikan status wanita, urbanisasi dan industrialisasi (Rusli, 1996).

Penelitian ini menggunakan delapan faktor yang diduga mempengaruhi usia pertama melahirkan. Faktor-faktor tersebut yaitu Usia Kawin Pertama (UKP), Pendidikan, dan Region. Deskripsi dari masing-masing faktor yang selanjutnya disebut variabel prediktor adalah sebagai berikut.

2.3.1 Usia Kawin Pertama

Usia Kawin Pertama (UKP) merupakan usia responden pada pernikahan pertama atau persatuan pertama yang diperoleh dari

kode bulan-tahun pernikahan/persatuan pertama dan kode bulan-tahun kelahiran responden (ICF, 2018).

2.3.2 Pendidikan

Pendidikan digolongkan menjadi tidak sekolah, tidak tamat SD, tamat SD, tidak tamat SLTA, tamat SLTA, dan perguruan tinggi. Kategori tidak sekolah ditujukan kepada responden yang tidak pernah sekolah. Kategori tidak tamat SD ditujukan kepada responden yang pernah atau sedang bersekolah di bangku sekolah dasar sedangkan responden yang sudah tamat sekolah dasar dikategorikan tamat SD. Kategori tidak tamat SLTA ditujukan kepada responden yang pernah atau sedang bersekolah di SMP, tamat SMP, atau pernah/sedang bersekolah di SLTA sedangkan responden yang sudah tamat SLTA dikategorikan tamat SLTA. Responden dengan tingkat pendidikan di atas SLTA dikategorikan dalam pendidikan perguruan tinggi (BKKBN, 2018).

2.3.3 Region

Variabel Region menunjukkan jenis tempat tinggal di mana rumah tangga berada yaitu perkotaan maupun pedesaan (ICF, 2018). Penelitian ini membatasi data yang digunakan merupakan data wanita yang sedang bertempat tinggal di tempat tinggal atau daerah yang sama dengan saat pertama kali melahirkan. Jika wanita tersebut belum pernah melahirkan maka data region menunjukkan daerah tempat tinggal saat ini.

(Halaman ini sengaja dikosongkan)

BAB III METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari hasil Survei Demografi dan Kesehatan Indonesia (SDKI) 2017, diunduh pada situs resmi *Demographic and Health Surveys Program* yaitu <https://dhsprogram.com>, dan telah mendapatkan izin resmi dengan bukti yang terlampir pada Lampiran 20. Data yang digunakan merupakan data dari 2.756 responden wanita usia subur (WUS) di Provinsi Jawa Timur yang berusia 15-49 tahun, pernah menikah, bertempat tinggal di region yang sama dengan ketika saat pertama kali melahirkan, dan pertama kali melahirkan dalam rentang tahun 1968 hingga 30 September 2017.

3.2 Metodologi SDKI 2017

Survei Demografi dan Kesehatan Indonesia (SDKI) 2017 dilaksanakan bersama oleh Badan Pusat Statistik (BPS), Badan Kependudukan dan Keluarga Berencana Nasional (BKKBN), dan Kementerian Kesehatan (Kemenkes). Pengumpulan data berlangsung dari 24 Juli hingga 30 September 2017. Dalam teknis pelaksanaannya, Pemerintah Indonesia dibantu oleh *ICF* melalui proyek *Demographic and Health Surveys (DHS) Program*, yaitu program *United States Agency for International Development (USAID)* yang menyediakan dana dan bantuan teknis dalam pelaksanaan survei kependudukan dan kesehatan di banyak negara (BKKBN, 2018). Tujuan utama SDKI 2017 adalah menyediakan estimasi terbaru indikator dasar demografi dan kesehatan. SDKI 2017 menyediakan gambaran menyeluruh tentang kependudukan serta kesehatan ibu dan anak di Indonesia.

3.2.1 Rancangan Sampel dan Hasil Kunjungan

Sampel SDKI 2017 mencakup 1.970 blok sensus yang meliputi daerah perkotaan dan perdesaan. Jumlah sampel rumah

tangga yang diharapkan sebanyak 49.250 rumah tangga. Dari seluruh sampel rumah tangga tersebut diharapkan akan dapat diperoleh sekitar 59.100 responden wanita usia subur umur 15-49, 24.625 responden remaja pria belum kawin umur 15-24, dan 14.193 responden pria kawin umur 15-54. Desain *sampling* yang digunakan dalam SDKI 2017 adalah *sampling* dua tahap berstrata, yaitu memilih sejumlah blok sensus secara *probability proportional to size* (PPS) sistematis dengan *size* jumlah rumah tangga hasil *listing* SP2010. Dalam hal ini, sistematis dilakukan dengan proses implisit stratifikasi menurut perkotaan dan perdesaan serta dengan mengurutkan blok sensus berdasarkan kategori *Wealth Index* dari hasil SP2010. Kemudian memilih 25 rumah tangga biasa di setiap blok sensus terpilih secara sistematis dari hasil pemutakhiran rumah tangga di setiap blok sensus tersebut. Sampel pria kawin (PK) akan dipilih 8 rumah tangga secara sistematis dari 25 rumah tangga tersebut.

Secara umum, hasil kunjungan rumah tangga dan perseorangan relatif tinggi. Dari 49.261 rumah tangga terpilih SDKI 2017, sebanyak 48.216 rumah tangga ditemukan, dan dari jumlah tersebut sebanyak 47.963 atau 99,5 persen rumah tangga berhasil diwawancarai. Dari 47.963 rumah tangga yang diwawancarai, terdapat 50.730 wanita yang memenuhi syarat untuk diwawancarai, dan sebanyak 49.627 wanita yang berhasil diwawancarai menghasilkan tingkat respon sebesar 97,8 persen. Dari rumah tangga terpilih sampel pria kawin, terdapat 10.440 pria yang memenuhi syarat untuk diwawancarai, dan yang berhasil diwawancarai ada 10.009 pria, menghasilkan tingkat respon sebesar 95,9 persen.

Dari 49.627 responden wanita yang berhasil diwawancarai, terdapat 3.729 responden yang berasal dari Provinsi Jawa Timur dan 45.898 lainnya berasal dari provinsi lain. 3.729 responden wanita usia subur Jawa Timur terdiri dari 2.974 wanita yang pernah menikah dan 755 wanita yang belum pernah menikah. Pada tahun 2017, dari 2.974 responden wanita terdapat sebanyak 2.756 responden wanita usia subur di Jawa Timur yang pernah menikah

bertempat tinggal di region yang sama dengan region saat pertama kali melahirkan atau region saat ini jika belum pernah melahirkan. Dari jumlah tersebut, terdapat sebanyak 2519 wanita yang pernah melahirkan dalam rentang tahun 1968 hingga 30 September 2017. Oleh karena itu, sesuai dengan batasan unit penelitian, data sekunder yang digunakan sebanyak 2.756 data responden wanita usia subur di Jawa Timur yang pernah menikah dan tinggal di region yang sama dengan region saat pertama kali melahirkan untuk wanita yang pernah melahirkan.

3.2.2 Kuesioner

Pelaksanaan SDKI 2017 menggunakan 4 (empat) jenis kuesioner yaitu kuesioner rumah tangga, wanita usia subur (WUS), pria kawin (PK), dan remaja pria (RP). Kuesioner rumah tangga dan wanita umur 15-49 mengacu pada kuesioner DHS (*Demographic Health Surveys*) Phase 7 tahun 2015 yang sudah mengakomodasi beberapa isu terbaru sesuai keterbandingan internasional dan telah disesuaikan dengan kondisi di Indonesia.

3.3 Kerangka Konsep

Berdasarkan penelitian Fagbamigbe & Idemudia (2016) bahwa karakteristik latar belakang serorang wanita berpengaruh terhadap keputusan melahirkan pertama kali dan akan mempengaruhi perkembangan dirinya setelah kelahiran pertama. Faktor latar belakang tersebut dibagi menjadi beberapa indikator yang diduga mempengaruhi usia pertama melahirkan dan dikelompokkan menjadi faktor demografi, faktor biologis, faktor sosial, faktor ekonomi, dan faktor wilayah. Berdasarkan faktor-faktor yang berpengaruh terhadap fertilitas menurut Rusli (1996) bahwa terdapat 2 faktor yaitu faktor demografi dan faktor nondemografi. Pada penelitian ini, tidak semua indikator dalam faktor-faktor tersebut diteliti dan juga terdapat beberapa indikator tambahan yang disesuaikan dengan variabel pada Data Survei Demografi dan Kesehatan Indonesia (SDKI) 2017. Rincian indikator dari setiap faktor yaitu sebagai berikut.

a) Faktor Demografi

1. Usia responden

Usia responden termasuk dalam konsep demografi yang digunakan untuk mengetahui karakteristik usia wanita saat ini terhadap usia pertama melahirkan, dan ingin diketahui apakah ada perbedaan karakteristik usia pertama melahirkan dalam rentang usia produktif wanita saat ini.

2. Usia Kawin Pertama (UKP)

Penelitian mengenai usia pada pernikahan pertama menjadi penting karena adanya hubungan yang erat antara pernikahan dengan timbulnya persalinan. Pernikahan telah dicatat berkorelasi positif terhadap persalinan pertama kali (Adebowale, Fagbamigbe, Okareh, & Lawal, 2012). Hasil dari penelitian Fagbamigbe & Idemudia (2016) juga mengatakan bahwa usia wanita saat pernikahan pertama secara signifikan berpengaruh terhadap usia pertama melahirkan.

3. Status perkawinan

Menurut (Mantra, 2003) bahwa status perkawinan merupakan salah satu faktor demografi yang mempengaruhi tingkat fertilitas. Tetapi status perkawinan ini tidak menentukan atau mempengaruhi keputusan wanita untuk melahirkan karena wanita pada saat melahirkan telah memiliki status perkawinan yang jelas.

4. Agama

Indikator agama ini merupakan salah satu pertanyaan yang diajukan pada kuesioner DHS (*Demographic Health Surveys*) Phase 7 dan diteliti dalam penelitian Fagbamigbe & Idemudia (2016), tetapi indikator ini tidak sesuai dengan kondisi Indonesia yang mana isu agama merupakan isu sensitif sehingga data agama responden tidak tercatat dalam *dataset* hasil SDKI 2017.

5. Pendidikan orang tua

Indikator pendidikan orang tua diteliti dalam penelitian Fagbamigbe & Idemudia (2016). Variabel pendidikan orang tua ini sudah diwakili oleh variabel pendidikan yang termasuk ke dalam faktor sosial.

b) Faktor Biologis

6. Status penggunaan kontrasepsi

Status penggunaan kontrasepsi yang dimaksudkan adalah status responden apakah sedang menggunakan alat kontrasepsi atau tidak sedang menggunakan alat kontrasepsi (ICF, 2018). Variabel ini tidak menunjukkan adanya keterkaitan terhadap usia pertama kali responden melahirkan sehingga variabel ini tidak digunakan pada penelitian ini.

7. Pengetahuan metode kontrasepsi

Berdasarkan penelitian Fagbamigbe & Idemudia (2016) bahwa wanita yang pernah menggunakan suatu alat/cara untuk mencegah kehamilan telah memiliki pengetahuan mengenai alat kontrasepsi sehingga kemungkinan kelahiran anak pertama dini lebih rendah. Pengetahuan ini diukur pada saat survei sehingga menjadi tidak relevan dengan kondisi pada saat melahirkan pertama kali.

8. Pernah memakai alat/cara KB

Informasi mengenai pernah tidaknya responden memakai alat/cara KB tidak menunjukkan status responden dalam penggunaan alat/cara KB untuk mencegah kehamilan yang pertama sehingga tidak diketahui pengaruhnya terhadap usia pertama melahirkan, adanya kekurangan informasi tersebut menyebabkan variabel pernah tidaknya memakai alat/cara KB tidak digunakan dalam penelitian ini.

9. Pernah tidaknya merokok

Variabel pernah tidaknya merokok termasuk ke dalam salah satu pertanyaan pada kuesioner DHS (*Demographic Health Surveys*) Phase 7 dan diteliti oleh Fagbamigbe & Idemudia (2016). Hasil penelitian tersebut menunjukkan bahwa variabel pernah tidaknya merokok tidak signifikan berpengaruh terhadap usia pertama melahirkan. Oleh karena itu, variabel ini tidak digunakan ke dalam penelitian ini.

10. Pernah tidaknya mengakhiri kehamilan

Variabel pernah tidaknya wanita mengakhiri kehamilan digunakan dalam penelitian Fagbamigbe & Idemudia (2016). Hasil

penelitian tersebut menunjukkan bahwa variabel pernah tidaknya mengakhiri kehamilan tidak signifikan berpengaruh terhadap usia pertama melahirkan.

c) Faktor Ekonomi

11. Kuintil Kekayaan

Menurut Noorkasiani, Heryati, & Rita (2009) bahwa salah satu faktor yang mendorong terjadinya perkawinan di usia muda yaitu keadaan ekonomi. Tidak jarang ditemukan perkawinan yang berlangsung dalam usia sangat muda, diantaranya disebabkan karena keinginan status ekonomi yang lebih tinggi dan sebagai jalan keluar untuk berbagai kesulitan ekonomi (Noorkasiani, Heryati, & Ismail, 2009). Kondisi kekayaan responden pada saat disurvei bisa jadi berbeda dengan pada saat pertama kali melahirkan, sehingga indikator tidak dapat mendeskripsikan karakteristik responden pada saat melahirkan pertama kali.

12. Status ekonomi

Sama halnya dengan kuintil kekayaan, indikator status ekonomi juga menunjukkan kondisi atau status ekonomi responden.

d) Faktor Wilayah

13. Region

Wilayah tempat tinggal yang dikategorikan perkotaan dan pedesaan merupakan salah satu faktor yang signifikan mempengaruhi usia pertama melahirkan menurut hasil penelitian Fagbamigbe & Idemudia (2016). Menurut Noorkasiani, Heryati, & Rita (2009), faktor yang mendorong terjadinya perkawinan di usia muda salah satunya adalah faktor lingkungan, termasuk faktor tempat tinggal dan adat istiadat yang mana karakteristik perkotaan dan pedesaan berbeda.

14. Wilayah bagian

Indikator wilayah bagian digunakan dalam penelitian Fagbamigbe & Idemudia (2016) yang bertujuan untuk mengetahui pola usia pertama melahirkan berdasarkan setiap wilayah bagian

negara di negara Nigeria. Sedangkan objek dalam penelitian ini termasuk dalam cakupan provinsi yaitu Jawa Timur sehingga indikator ini tidak perlu digunakan dan telah diwakili oleh variabel region.

e) Faktor Sosial

15. Pendidikan

Salah satu faktor yang mendorong terjadinya perkawinan di usia muda yaitu tingkat pendidikan seorang wanita (Noorkasiani, Heryati, & Ismail, 2009). Menurut Noorkasiani, Heryati, & Rita (2009) semakin rendah tingkat pendidikan semakin mendorong berlangsungnya perkawinan usia muda. Hal tersebut disebabkan peran tingkat pendidikan berhubungan erat dengan pemahaman tentang kehidupan berkeluarga.

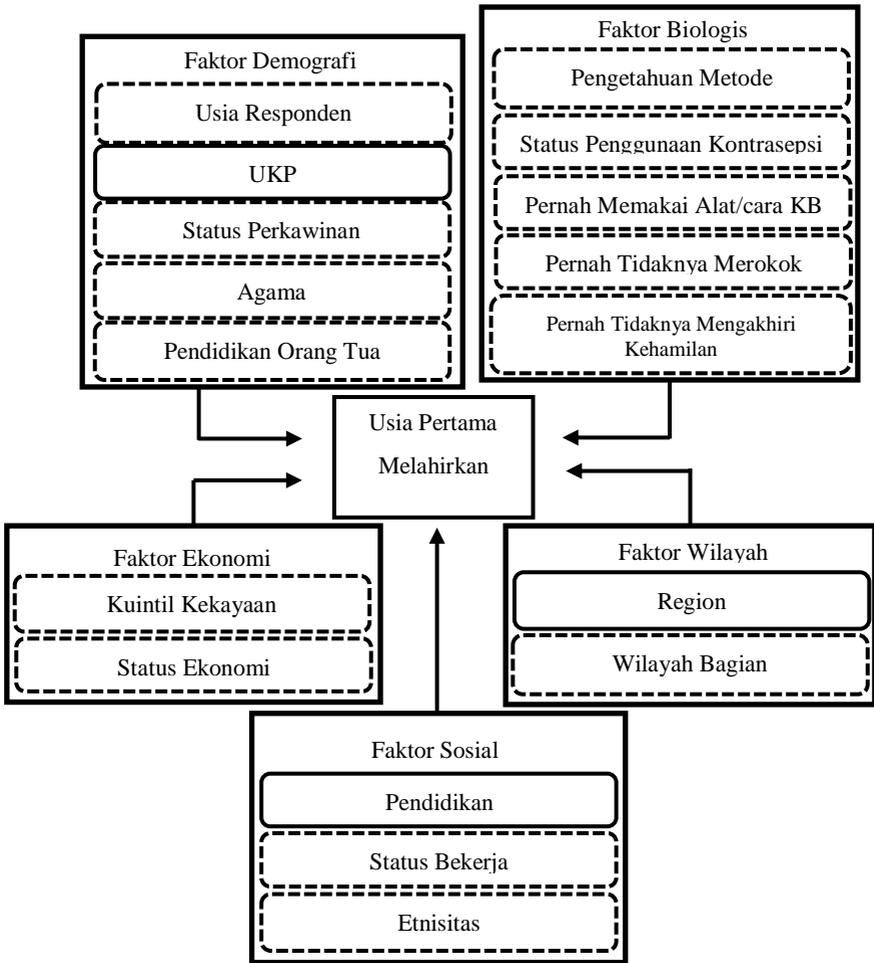
16. Status bekerja

Pekerjaan memiliki hubungan dua arah dengan usia pertama melahirkan. Suatu keputusan menjadi seorang ibu baik disengaja ataupun tidak, akan memiliki efek konsekuensial pada kemampuan kerja (Fagbamigbe & Idemudia, 2016). Kelemahan jika menggunakan data hasil survei ini adalah data yang didapatkan diukur pada waktu yang berbeda dengan pertama kali melahirkan sehingga faktor status bekerja menjadi tidak relevan dengan usia pertama melahirkan.

17. Etnisitas

Indikator etnitas juga merupakan variabel yang diteliti oleh Fagbamigbe & Idemudia (2016) dan pertanyaan pada kuesioner DHS (*Demographic Health Surveys*) Phase 7. Karena beragamnya etnik dan isu mengenai etnik di Indonesia merupakan isu sensitif, data identitas etnik responden juga tidak tercatat dalam *dataset* hasil SDKI 2017.

Pada penelitian ini digunakan 3 indikator yang diketahui dari 17 indikator yang diduga mempengaruhi usia pertama melahirkan. Berikut ini kerangka konsep yang digunakan dalam penelitian ini.



Gambar 3.1 Kerangka Konsep Penelitian.

Keterangan: Variabel yang diteliti Variabel yang tidak diteliti

Sumber: Rusli, 1996; Fagbamigbe & Idemudia, 2016; dan Pendataan Survei Demografi dan Kesehatan Indonesia (SDKI) 2017 yang telah disesuaikan dan dimodifikasi.

3.4 Variabel Penelitian

Variabel penelitian yang digunakan dalam penelitian ini terdiri atas variabel dependen dan variabel prediktor. Variabel dependen yang digunakan dalam penelitian ini terdiri atas *survival time* (T) dan status tersensor (d) yang menunjukkan responden mengalami *event* atau tidak. Usia pertama melahirkan adalah usia pertama kali melahirkan bagi seorang wanita usia subur yang berusia 15-49 tahun dan pernah menikah. Skema *survival* pada penelitian ini adalah sebagai berikut.

- a. *Event* yang diteliti pada penelitian ini adalah kejadian wanita melahirkan pertama kali dalam rentang tahun 1968 hingga 30 September 2017.
- b. Data dikatakan tersensor apabila responden yang terdata belum pernah melahirkan hingga berakhirnya waktu survei pada 30 September 2017.

Detail deskripsi telah dijelaskan pada Tinjauan Pustaka yang sesuai dengan deskripsi masing-masing kode variabel pada *Demographic and Health Surveys Standar Recode Manual for DHS-7*, kode V212 menunjukkan variabel Usia Pertama Melahirkan, kode V511 menunjukkan variabel Usia Kawin Pertama, kode V149 menunjukkan variabel Pendidikan, dan kode V025 menunjukkan variabel Region. Adapun variabel-variabel penelitian dijelaskan melalui tabel sebagai berikut.

Tabel 3.1 Variabel Penelitian

Simbol Variabel	Nama Variabel	Skala	Deskripsi
T	Usia Pertama Melahirkan	Rasio	Periode waktu responden dari usia nol sampai melahirkan pertama atau sampai waktu survei berakhir tanpa terjadinya <i>event</i> .
d	Status Tersensor	Nominal	Responden mengalami <i>event</i> atau tidak 0 = tersensor (tidak terjadi <i>event</i>) 1 = tidak tersensor (terjadi <i>event</i>)

Tabel 3.1 Variabel Penelitian (*Lanjutan*)

Simbol Variabel	Nama Variabel	Scala	Deskripsi
X_1	Usia Kawin Pertama (UKP)	Rasio	Usia responden saat pertama kali menikah/mulai hidup bersama dengan suami/pasangan.
X_2	Pendidikan	Ordinal	Jenjang pendidikan yang ditempuh responden. 0 = Tidak sekolah 1 = Tidak tamat SD 2 = Tamat SD 3 = Tidak tamat SLTA 4 = Tamat SLTA 5 = Perguruan Tinggi
X_3	Region	Nominal	Klasifikasi daerah tempat tinggal responden. 1 = Perkotaan 2 = Pedesaan

3.5 Struktur Data

Struktur data dari variabel-variabel yang digunakan dalam penelitian ini adalah sebagai berikut.

Tabel 3. 2 Struktur Data Penelitian

id	T	d	X_1	X_2	X_3
1	16	1	16	2	2
2	22	1	20	3	2
3	24	1	23	3	2
⋮	⋮	⋮	⋮	⋮	⋮
2756	21	0	21	4	1

3.6 Penulisan Model *Cox Extended*

Pemodelan *Cox Extended* pada penelitian ini menggunakan dua skenario kombinasi fungsi waktu yang diinteraksikan dengan variabel yang tidak memenuhi asumsi *proportional hazard*. Skenario pertama yaitu pemodelan *Cox Extended* menggunakan kombinasi fungsi waktu *heavyside* dengan mengelompokkan variabel usia kawin pertama menjadi empat kategori. Sedangkan skenario kedua yaitu pemodelan *Cox Extended* menggunakan kombinasi fungsi waktu $\ln(T)$ dan *heavyside*. Penulisan model lengkap *Cox Extended* skenario pertama adalah sebagai berikut.

$$h(t_i, \mathbf{x}_i) = h_0(t_i) \exp \left[\beta_3 x_{i3} + \sum_{k=1}^3 \delta_{1k}^* x_{ik}^* g_{11}(t_i) + \sum_{m=1}^5 \delta_{1m}^{**} x_{im}^{**} g_{12}(t_i) + \sum_{k=1}^3 \delta_{2k}^* x_{ik}^* g_{21}(t_i) + \sum_{m=1}^5 \delta_{2m}^{**} x_{im}^{**} g_{22}(t_i) \right]$$

$$\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})$$

Keterangan:

$h_0(t_i)$: *baseline hazard* pada waktu ke- t_i .

β_3 : parameter regresi dari variabel Region.

δ_{1k}^* : parameter regresi dari variabel Kelompok Usia Kawin Pertama kategori ke- k yang telah diinteraksikan dengan fungsi waktu $g_{11}(t_i)$.

δ_{2k}^* : parameter regresi dari variabel Kelompok Usia Kawin Pertama kategori ke- k yang telah diinteraksikan dengan fungsi waktu $g_{21}(t_i)$.

δ_{1m}^{**} : parameter regresi dari variabel Pendidikan kategori ke- m yang telah diinteraksikan dengan fungsi waktu $g_{12}(t_i)$.

δ_{2m}^{**} : parameter regresi dari variabel Pendidikan kategori ke- m yang telah diinteraksikan dengan fungsi waktu $g_{22}(t_i)$.

x_{i3} : nilai dari variabel Region.

x_{ik}^* : nilai dari variabel Kelompok Usia Kawin Pertama kategori ke- k .

x_{im}^{**} : nilai dari variabel Pendidikan kategori ke- m .

Penulisan model lengkap *Cox Extended* skenario kedua adalah sebagai berikut.

$$h(t_i, \mathbf{x}_i) = h_0(t_i) \exp \left[\beta_3 x_{i3} + \beta_1 x_{i1}^* + \sum_{m=1}^5 \delta_{1m}^{**} x_{im}^{**} g_{12}(t_i) + \delta_1 x_{i1}^* \ln(t_i) + \sum_{m=1}^5 \delta_{2m}^{**} x_{im}^{**} g_{22}(t_i) \right]$$

$$\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})$$

Keterangan:

$h_0(t_i)$: *baseline hazard* pada waktu ke- t_i .

β_1 : parameter regresi dari variabel Usia Kawin Pertama.

β_3 : parameter regresi dari variabel Region.

δ_1 : parameter regresi dari variabel Usia Kawin Pertama yang telah diinteraksikan dengan fungsi waktu $\ln(t_i)$.

δ_{1m}^{**} : parameter regresi dari variabel Pendidikan kategori ke- m yang telah diinteraksikan dengan fungsi waktu $g_{12}(t_i)$.

δ_{2m}^{**} : parameter regresi dari variabel Pendidikan kategori ke- m yang telah diinteraksikan dengan fungsi waktu $g_{22}(t_i)$.

x_{i1}^* : nilai dari variabel Usia Kawin Pertama.

x_{i3} : nilai dari variabel Region.

x_{im}^{**} : nilai dari variabel Pendidikan kategori ke- m .

3.7 Langkah Analisis

Langkah analisis dalam penelitian ini adalah sebagai berikut.

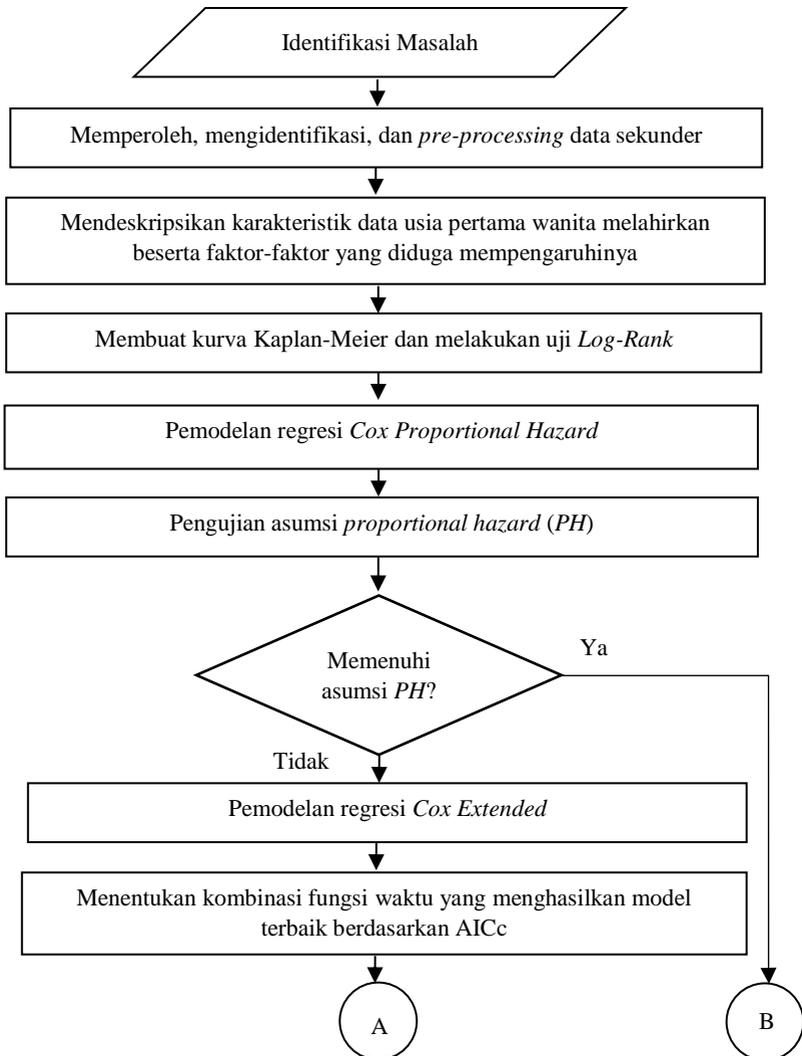
1. Memperoleh data Survei Demografi dan Kesehatan Indonesia (SDKI) 2017 tentang wanita usia subur (WUS) di Provinsi Jawa Timur yang pernah menikah dan bertempat tinggal di region yang sama dengan region saat pertama kali melahirkan.
2. Mengidentifikasi data usia pertama melahirkan dan variabel-variabel prediktor yang diduga berpengaruh terhadap usia pertama melahirkan yang akan digunakan dalam penelitian ini. Serta melakukan *pre-processing* data seperti memilah data sesuai dengan batasan unit penelitian yang digunakan.

3. Mendeskripsikan karakteristik data usia pertama wanita melahirkan berdasarkan faktor-faktor yang diduga mempengaruhinya dengan melakukan analisis statistika deskriptif menggunakan ukuran pemusatan data yaitu *mean* (rata-rata) dan ukuran penyebaran data yaitu deviasi standar.
4. Membuat kurva *survival* waktu usia pertama wanita melahirkan menggunakan pendekatan *Kaplan Meier* dan melakukan uji *Log-Rank* untuk mengetahui ada tidaknya perbedaan kurva *survival* untuk beberapa kategori pada beberapa variabel prediktor yang bersifat kategorik.
5. Melakukan pengujian asumsi *proportional hazard* dengan langkah sebagai berikut.
 - a. Memodelkan usia pertama wanita melahirkan menggunakan regresi *Cox Proportional Hazard*.
 - b. Melakukan pengujian asumsi *proportional hazard* menggunakan uji *Goodness of Fit*.
6. Melakukan pemodelan Regresi *Cox Extended* data usia pertama wanita melahirkan dengan langkah sebagai berikut.
 - a. Identifikasi variabel prediktor yang tidak memenuhi asumsi *proportional hazard* hasil uji *Goodness of Fit* menggunakan Variabel *Time Dependent* yaitu dengan memasukkan variabel tersebut ke dalam model *Cox Extended*
 - b. Menentukan fungsi waktu yang akan digunakan sebagai interaksi waktu dari variabel-variabel prediktor yang tidak memenuhi asumsi *proportional hazard*.
 - c. Melakukan pemodelan Regresi *Cox Extended* usia pertama wanita melahirkan dengan kombinasi fungsi waktu yang berbeda
 - d. Melakukan pengujian signifikansi parameter faktor-faktor yang diduga mempengaruhi usia pertama wanita melahirkan yaitu dengan melakukan uji serentak dan parsial

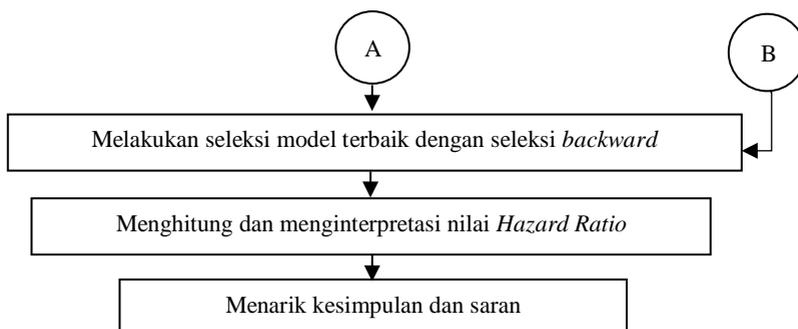
- e. Menentukan variabel-variabel yang berpengaruh signifikan terhadap usia pertama wanita melahirkan berdasarkan hasil pengujian signifikansi parameter.
 - f. Menentukan kombinasi fungsi waktu yang menghasilkan model terbaik berdasarkan kriteria kebaikan model AICc.
 - g. Melakukan seleksi model terbaik dengan menggunakan *backward elimination* pada pemodelan yang menggunakan kombinasi fungsi waktu yang terpilih
 - h. Menghitung dan menginterpretasi nilai *Hazard Ratio* dari setiap variabel prediktor yang berpengaruh signifikan terhadap model.
7. Menarik kesimpulan dan saran.

3.8 Diagram Alir Langkah-langkah Analisis

Adapun diagram alir untuk langkah-langkah analisis adalah sebagai berikut.



Gambar 3. 2 Diagram Alir Penelitian.



Gambar 3. 2 Diagram Alir Penelitian (*Lanjutan*).

BAB IV ANALISIS DAN PEMBAHASAN

Karakteristik responden wanita usia subur di Jawa Timur yang pernah menikah berikut ini akan dideskripsikan oleh usia pertama melahirkan berdasarkan faktor-faktor yang diduga berpengaruh seperti Usia Kawin Pertama (X_1), Pendidikan (X_2), dan Region (X_3). Setelahnya, akan dilakukan analisis *survival* untuk mendapatkan faktor-faktor yang signifikan berpengaruh terhadap usia pertama melahirkan dengan beberapa tahap, mulai dari eksplorasi data menggunakan kurva *survival Kaplan Meier* dan uji *Log-Rank*, pemodelan menggunakan Regresi *Cox Proportional Hazard* dan pengujian asumsi *proportional hazard*, hingga pemodelan menggunakan Regresi *Cox Extended*.

4.1 Karakteristik Usia Pertama Melahirkan Pada Wanita Usia Subur di Jawa Timur

Karakteristik usia pertama melahirkan pada wanita usia subur di Jawa Timur yang pernah menikah berdasarkan masing-masing faktor yang diduga berpengaruh akan dideskripsikan melalui analisis statistika deskriptif dengan menggunakan ukuran pemusatan data yaitu *mean* (rata-rata) dan ukuran penyebaran data yaitu deviasi standar. Dalam analisis statistika deskriptif kali ini menggunakan bobot untuk perhitungan *mean* dan deviasi standar. Bobot yang digunakan telah tersedia dalam dataset seperti pada Lampiran 1. Berikut ini karakteristik usia pertama melahirkan berdasarkan pernah tidaknya responden melahirkan (status tersensor).

Tabel 4. 1 Statistika Deskriptif Usia Pertama Melahirkan

Variabel	Status Tersensor (d)	Mean	Deviasi Standar
T	Tersensor	29,6293	9,624
	Tidak tersensor	21,6180	4,484

Berdasarkan Tabel 4.1 bahwa rata-rata wanita usia subur di Jawa Timur yang pernah menikah melahirkan pertama kali di usia 21 tahun. Sedangkan wanita yang belum pernah melahirkan hingga

30 September 2017 rata-rata berusia 29 tahun, hal ini berarti rata-rata dari mayoritas wanita usia subur di Jawa Timur yang pernah menikah melahirkan di usia muda. Walaupun demikian, usia dari wanita yang belum pernah melahirkan lebih bervariasi dibanding usia wanita pertama kali melahirkan, hal tersebut ditunjukkan oleh nilai deviasi standar yang lebih besar.

4.1.1 Karakteristik Usia Pertama Melahirkan Berdasarkan Faktor Usia Kawin Pertama

Karakteristik usia pertama melahirkan berdasarkan faktor usia kawin pertama akan dideskripsikan melalui *mean* (rata-rata), dan deviasi standar. Berikut ini karakteristik usia pertama kali menikah dari wanita usia subur di Jawa Timur beserta keterkaitannya dengan usia pertama melahirkan.

Tabel 4. 2 Statistika Deskriptif Usia Kawin Pertama

Variabel	Status Tersensor (d)	Mean	Deviasi Standar
Usia Kawin Pertama (X_1)	Tersensor	22,654	6,2759
	Tidak tersensor	19,844	4,5737

Rata-rata wanita usia subur di Jawa Timur yang pernah melahirkan menikah di usia 19 tahun, sedangkan wanita usia subur di Jawa Timur yang belum pernah melahirkan hingga 30 September 2017 menikah di usia 22 tahun. Rata-rata wanita usia subur di Jawa Timur melahirkan untuk pertama kalinya setelah 2 tahun menikah. Sedangkan wanita usia subur yang belum pernah melahirkan hingga 30 September 2017 rata-rata belum melahirkan untuk pertama kalinya setelah 7 tahun menikah, hal tersebut dapat diindikasikan oleh faktor lain yang perlu untuk diidentifikasi.

4.1.2 Karakteristik Usia Pertama Melahirkan Berdasarkan Faktor Pendidikan

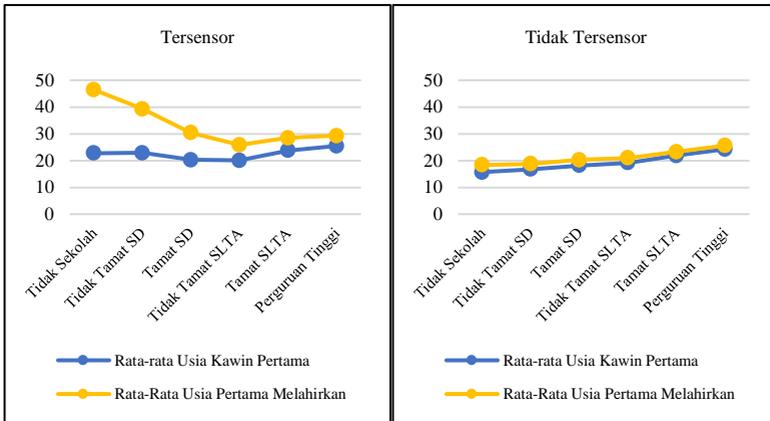
Tingkat pendidikan terakhir yang ditempuh responden dikategorikan menjadi 6 kategori yaitu tidak sekolah, tidak tamat SD, tamat SD, tidak tamat SLTA, tamat SLTA, dan perguruan tinggi. Rata-rata dan deviasi standar usia pertama melahirkan dibedakan berdasarkan kategori pendidikan seperti berikut ini.

Tabel 4. 3 Statistika Deskriptif Usia Pertama Melahirkan Berdasarkan Pendidikan

Variabel Pendidikan	Tersensor		Tidak Tersensor	
	<i>Mean</i>	Deviasi Standar	<i>Mean</i>	Deviasi Standar
Tidak Sekolah	46,427	2,114	18,354	5,149
Tidak Tamat SD	39,302	7,480	18,757	4,327
Tamat SD	30,366	10,422	20,298	4,191
Tidak Tamat SLTA	25,864	9,129	20,986	3,919
Tamat SLTA	28,474	8,619	23,258	3,807
Perguruan Tinggi	29,267	7,037	25,638	3,882

Tabel 4.3 menunjukkan rata-rata usia pertama melahirkan dari wanita usia subur yang pernah menikah cenderung naik searah dengan semakin tingginya tingkat pendidikan terakhir yang ditempuh. Tetapi terdapat beberapa rata-rata usia yang sama untuk beberapa kategori tingkat pendidikan yaitu rata-rata usia pertama melahirkan dari wanita yang tidak sekolah sama dengan wanita yang tidak tamat SD, begitu juga wanita lulusan SD memiliki rata-rata usia pertama melahirkan yang sama dengan wanita yang tidak tamat SLTA. Sedangkan wanita yang belum pernah melahirkan memiliki usia yang bervariasi pada setiap kategori pendidikan dan rata-rata usia paling muda dimiliki oleh wanita yang tidak tamat SLTA.

Rata-rata Usia Pertama Melahirkan dari kategori Tersensor merupakan usia wanita yang belum pernah melahirkan hingga berakhirnya waktu survei pada 30 September 2017. Sedangkan rata-rata Usia Pertama Melahirkan dari kategori Tidak Tersensor merupakan usia wanita pada saat melahirkan pertama kali. Berikut ini merupakan grafik perbandingan rata-rata Usia Pertama Melahirkan dan rata-rata Usia Kawin Pertama pada setiap kategori tingkat pendidikan wanita untuk mengetahui interval jarak usia wanita pada saat menikah pertama kali hingga melahirkan pertama kali.



Gambar 4. 1 Rata-rata Usia Pertama Melahirkan dan Rata-rata Usia Kawin Pertama Berdasarkan Pendidikan.

Jarak usia wanita pada saat menikah hingga 30 September 2017 yang terlalu lama untuk tidak melahirkan yang pertama kali hingga waktu berakhirnya survei yaitu pada wanita yang tidak sekolah hingga berpendidikan tamat SD. Sedangkan wanita yang pernah melahirkan dari setiap kategori pendidikan rata-rata memiliki jarak 1 hingga 3 tahun untuk melahirkan yang pertama kalinya dari waktu setelah menikah

4.1.3 Karakteristik Usia Pertama Melahirkan Berdasarkan Faktor Region

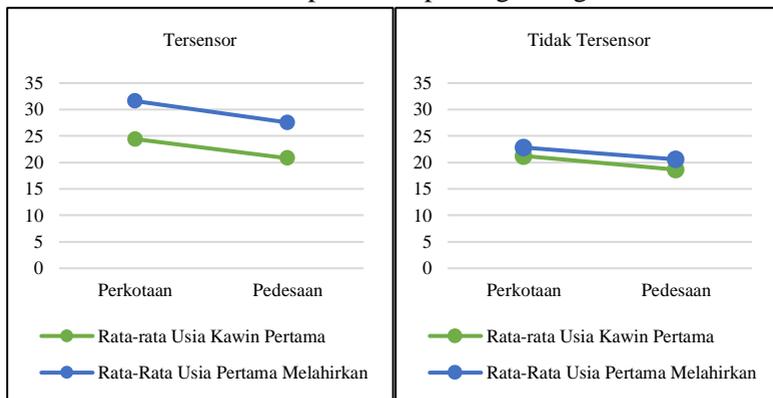
Lingkungan perkotaan dan pedesaan memiliki karakteristik yang berbeda seperti adat istiadat, pola pikir, ataupun gaya hidup, sehingga memungkinkan region memiliki pengaruh yang berbeda terhadap usia pertama melahirkan.

Tabel 4. 4 Statistika Deskriptif Usia Pertama Melahirkan Berdasarkan Region

Variabel Region	Tersensor		Tidak Tersensor	
	Mean	Deviasi Standar	Mean	Deviasi Standar
Perkotaan	31,599	9,448	22,819	4,587
Pedesaan	27,553	9,367	20,578	4,119

Terlihat pada Tabel 4.4 bahwa rata-rata usia pertama melahirkan dari wanita yang tinggal di perkotaan yaitu 22 tahun, sedangkan wanita yang tinggal di pedesaan rata-rata melahirkan di usia yang lebih muda daripada wanita yang tinggal di perkotaan yaitu 20 tahun. Disamping itu, usia wanita yang belum pernah melahirkan dari wanita yang tinggal di perkotaan juga lebih dewasa dibanding wanita yang tinggal di pedesaan. Hal tersebut menimbulkan dugaan region tempat tinggal wanita memberi pengaruh yang berbeda terhadap usia pertama melahirkan.

Kategori Tersensor diperuntukkan untuk wanita yang belum pernah melahirkan hingga 30 September 2017, sehingga variabel Usia Pertama Melahirkan merupakan usia wanita per 30 September 2017. Interval jarak usia wanita pada saat menikah pertama kali hingga melahirkan pertama kali dapat diketahui dari grafik berikut ini yang menunjukkan rata-rata Usia Pertama Melahirkan dan rata-rata Usia Kawin Pertama pada setiap kategori region.



Gambar 4. 2 Rata-rata Usia Pertama Melahirkan dan Rata-rata Usia Kawin Pertama Berdasarkan Region.

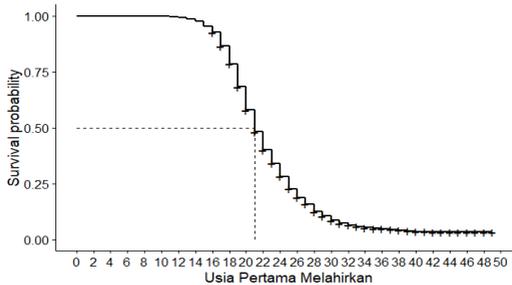
Jarak waktu wanita melahirkan untuk pertama kali atau jarak usia wanita hingga 30 September 2017 dari usia menikah pertama untuk wanita yang tinggal di perkotaan maupun pedesaan tidak berbeda. Hal tersebut berarti region tempat tinggal wanita tidak mempengaruhi jarak usia wanita untuk melahirkan pertama kali dari waktu pernikahan pertama.

4.2 Pemodelan Regresi *Cox Extended* Pada Usia Pertama Melahirkan

Berbagai tahap dilakukan dalam analisis *survival* sebelum dilakukannya pemodelan Regresi *Cox Extended* diantaranya yaitu analisis kurva *survival Kaplan Meier* dan uji *Log-Rank*, pengujian asumsi *proportional hazard*, identifikasi variabel prediktor yang tidak memenuhi asumsi *proportional hazard*, hingga penentuan fungsi waktu yang akan digunakan dalam pemodelan regresi *Cox Extended*.

4.2.1 Kurva *Survival Kaplan Meier* dan Uji *Log-Rank*

Kurva *survival Kaplan Meier* digunakan untuk menggambarkan probabilitas wanita usia subur di Jawa Timur yang pernah menikah untuk tidak melahirkan yang pertama kalinya hingga usia 49 tahun berdasarkan faktor faktor yang diduga mempengaruhi usia pertama melahirkan. Uji *Log-Rank* digunakan untuk membandingkan kurva *survival* dalam kategori yang berbeda. Berikut ini merupakan kurva *survival* untuk mengetahui karakteristik secara umum dari usia pertama melahirkan.

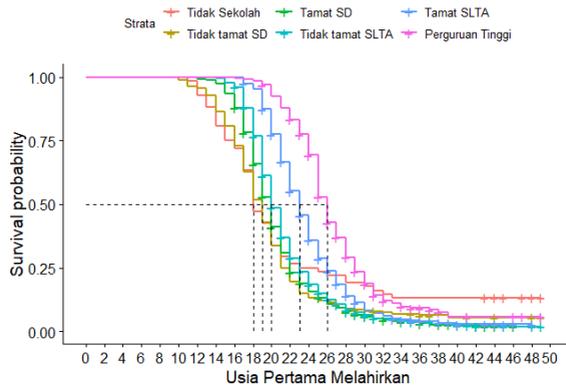


Gambar 4. 3 Kurva *Survival* Usia Pertama Melahirkan.

Probabilitas wanita usia subur di Jawa Timur yang pernah menikah untuk tidak melahirkan yang pertama kalinya sampai usia lebih dari 17 tahun menurun cepat hingga usia 25 tahun seperti pada Gambar 4.3 yang menunjukkan kurva turun secara cepat pada usia wanita ke-17 hingga usia ke-25. Probabilitas wanita untuk tidak melahirkan yang pertama kalinya sampai usia lebih dari 21 tahun yaitu sebesar 0,5.

A. Faktor Pendidikan

Faktor pendidikan dibagi berdasarkan jenjang pendidikan terakhir yang ditempuh responden. Jenjang pendidikan tersebut dibagi menjadi enam kategori yaitu tidak sekolah, tidak tamat SD, tamat SD, tidak tamat SLTA, tamat SLTA, dan perguruan tinggi. Berikut ini disajikan kurva *survival Kaplan Meier* berdasarkan faktor pendidikan.



Gambar 4. 4 Kurva *Kaplan Meier* Berdasarkan Faktor Pendidikan.

Gambar 4.4 menunjukkan semakin tinggi jenjang pendidikan terakhir yang di tempuh maka semakin tinggi pula probabilitas wanita untuk melahirkan yang pertama kalinya di usia yang semakin dewasa. Hal tersebut juga dapat diketahui dari kurva *Kaplan Meier* semakin ke kanan kecuali pada kurva kategori “tidak sekolah” dan “tidak tamat SD” yang saling berimpit, kurva kategori “tamat SD” dan “tidak tamat SLTA” sedikit menjauh dari kedua kurva tersebut apabila usia pertama melahirkan kurang dari usia 20 tahun. Probabilitas wanita yang tidak tamat SLTA, tamat SLTA, atau mengenyam perguruan tinggi untuk tidak melahirkan yang pertama kalinya sampai usia lebih dari 20 tahun lebih tinggi dari wanita yang tidak sekolah, tidak tamat SD, atau tamat SD. Dengan propobalitas sekitar 0,5, wanita yang tidak sekolah atau tidak tamat SD tidak melahirkan untuk pertama kalinya sampai usia lebih dari 18 tahun, wanita yang tamat SD tidak melahirkan untuk pertama

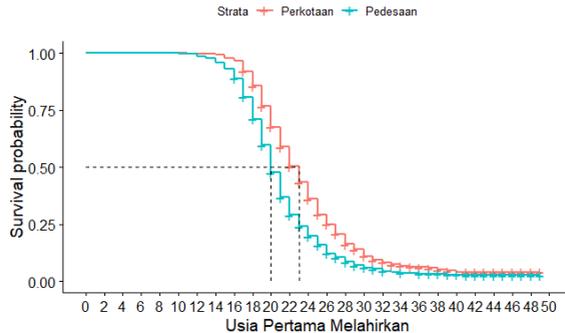
kalinya sampai usia lebih dari 19 tahun, wanita yang tidak tamat SLTA tidak melahirkan untuk pertama kalinya sampai usia lebih dari 20 tahun, wanita yang menyelesaikan pendidikan SLTA tidak melahirkan untuk pertama kalinya sampai usia lebih dari 23 tahun, dan wanita yang pernah mengenyam pendidikan tinggi tidak melahirkan untuk pertama kalinya sampai usia lebih dari 26 tahun. Setelah usia ke-21 probabilitas wanita yang tidak sekolah untuk tidak melahirkan yang pertama kalinya lebih tinggi dari wanita yang tidak tamat SD. Setelah usia ke-22 probabilitas wanita yang tidak sekolah untuk tidak melahirkan yang pertama kalinya lebih tinggi dari wanita yang tamat SD. Setelah usia ke-24 probabilitas wanita yang tidak sekolah untuk tidak melahirkan yang pertama kalinya lebih tinggi dari wanita yang tidak tamat SLTA. Setelah usia ke-27 probabilitas wanita yang tidak sekolah untuk tidak melahirkan yang pertama kalinya lebih tinggi dari wanita yang tamat SLTA. Dan setelah usia ke-31 probabilitas wanita yang tidak sekolah untuk tidak melahirkan yang pertama kalinya lebih tinggi dari wanita berpendidikan tinggi. Untuk mengetahui adanya perbedaan yang signifikan diantara keenam kurva *survival* maka dilakukan pengujian *Log-Rank*.

Didapatkan nilai statistik uji *Log-Rank* sebesar 324 dengan derajat bebas 5 dan *p-value* kurang dari taraf signifikansi 0,05 sesuai dengan *output* pada Lampiran 12. Nilai $\chi^2_{(5;0,05)}$ didapatkan sebesar 11,07. Karena nilai statistik uji *Log-Rank* lebih besar dari nilai $\chi^2_{(5;0,05)}$ dan *p-value* kurang dari 0,05 maka dapat diputuskan tolak H_0 yang berarti bahwa minimal terdapat satu perbedaan yang signifikan pada kurva *Kaplan Meier* antara kategori pendidikan yang berbeda.

B. Faktor Region

Region dibagi menjadi dua kategori yaitu perkotaan dan pedesaan. Uji *Log-Rank* digunakan untuk mengetahui kebenaran adanya perbedaan yang signifikan antara kedua kurva. Berikut ini

merupakan gambar kurva *Kaplan Meier* berdasarkan faktor region dan hasil pengujian *Log-Rank* .



Gambar 4. 5 Kurva *Kaplan Meier* Berdasarkan Faktor Region.

Gambar 4.5 menunjukkan kedua kurva *survival* kategori perkotaan dan pedesaan berjauhan atau tidak berimpit, hal tersebut menimbulkan dugaan bahwa terdapat perbedaan kurva *survival* antara wanita yang tinggal di perkotaan dengan wanita yang tinggal di pedesaan. Probabilitas wanita yang tinggal di perkotaan untuk tidak melahirkan yang pertama kalinya selalu lebih tinggi dari wanita yang tinggal di pedesaan. Dengan probabilitas yang sama yaitu 0,5, wanita yang tinggal di pedesaan tidak melahirkan untuk pertama kalinya sampai usia lebih dari 21 tahun sedangkan wanita yang tinggal di perkotaan tidak melahirkan untuk pertama kalinya sampai usia lebih dari 23 tahun. Hal tersebut berarti bahwa wanita yang tinggal di perkotaan melahirkan pertama kali di usia yang lebih dewasa dibandingkan wanita yang tinggal di pedesaan.

Uji *Log-Rank* seperti pada Lampiran 12 menghasilkan nilai statistik uji sebesar 120 dengan derajat bebas 1 dan *p-value* kurang dari 2×10^{-16} . Jika dibandingkan dengan nilai taraf signifikansi α yaitu sebesar 0,05 dan nilai $\chi^2_{(1;0,05)}$ sebesar 3,841, maka *p-value* kurang dari α dan nilai statistik uji *Log-Rank* lebih besar dari nilai $\chi^2_{(1;0,05)}$, sehingga uji ini menghasilkan keputusan tolak H_0 . Kesimpulan yang diperoleh yaitu terdapat perbedaan kurva

survival yang signifikan antara wanita yang tinggal di perkotaan dan wanita yang tinggal di pedesaan.

4.2.2 Pengujian Asumsi *Proportional Hazard*

Pemeriksaan asumsi *proportional hazard* dengan pendekatan statistik digunakan untuk memperoleh keputusan yang lebih objektif, pemeriksaan asumsi *proportional hazard* dapat menggunakan uji *goodness of fit*. H_0 memiliki arti bahwa faktor yang diduga berpengaruh terhadap usia pertama melahirkan telah memenuhi asumsi *proportional hazard*. Sedangkan H_1 memiliki arti sebaliknya yaitu faktor yang diduga berpengaruh terhadap usia pertama melahirkan tidak memenuhi asumsi *proportional hazard*. Berikut ini disajikan hasil pengujian *Goodness of Fit* pada Tabel 4.5.

Tabel 4. 5 Hasil Uji *Goodness of Fit*

Variabel	ρ	P-Value	Keputusan
Usia Kawin Pertama	0,5187	0,0000	Tolak H_0
Pendidikan (1)	0,0387	0,0422	Tolak H_0
Pendidikan (2)	0,0774	0,0000	Tolak H_0
Pendidikan (3)	0,0911	0,0000	Tolak H_0
Pendidikan (4)	0,1049	0,0000	Tolak H_0
Pendidikan (5)	0,0913	0,0000	Tolak H_0
Region (2)	0,0347	0,0685	Gagal Tolak H_0

Hasil uji *Goodness of Fit* pada Tabel 4.5 menunjukkan *p-value* yang dihasilkan bernilai kurang dari taraf signifikansi $\alpha = 0,05$ adalah variabel usia kawin pertama dan pendidikan. Sehingga faktor usia pertama melahirkan (X_1) dan pendidikan (X_2) memberikan keputusan tolak H_0 , dapat disimpulkan bahwa kedua faktor tersebut tidak memenuhi asumsi *proportional hazard*. Hal tersebut berarti bahwa variabel prediktor usia kawin pertama dan pendidikan bergantung pada waktu atau nilainya dapat berubah bergantung pada waktu. Oleh karena itu, terdapat dua variabel yang harus diinteraksikan dengan fungsi waktu $g(t)$.

Fungsi waktu $g(t)$ yang digunakan dalam penelitian ini adalah fungsi $\ln(T)$ dan fungsi *heavyside* sehingga terdapat dua jenis pemodelan. Pemodelan pertama merupakan pemodelan

dengan memasukkan kedua variabel yang diinteraksikan dengan fungsi waktu *heavyside* (kombinasi fungsi *heavyside*). Pemodelan kedua merupakan pemodelan dengan memasukkan satu variabel yang diinteraksikan dengan fungsi $\ln(T)$ dan variabel lainnya diinteraksikan dengan fungsi *heavyside* (kombinasi fungsi $\ln(T)$ dan *heavyside*).

4.2.3 Pemodelan Regresi *Cox Extended* Menggunakan Kombinasi Fungsi Waktu *Heavyside*

Pemodelan regresi *cox extended* dengan menggunakan kombinasi fungsi waktu *heavyside* pada penelitian ini merupakan pemodelan regresi *cox extended* dengan memasukkan variabel yang memenuhi asumsi *proportional hazard* yaitu variabel region, variabel usia kawin pertama yang diinteraksikan dengan fungsi *heavyside*, dan variabel pendidikan yang diinteraksikan dengan fungsi *heavyside*.

Variabel usia kawin pertama yang digunakan pada pemodelan ini merupakan variabel kategorik yang berasal dari pembagian usia kawin pertama menjadi beberapa kelompok usia. Pembagian kelompok usia kawin pertama ini berdasarkan hasil grafik *Martingale Residual* yang dapat dilihat pada Lampiran 13 dan didapatkan 4 kelompok usia kawin pertama yaitu kurang dari 15 tahun, 15-19 tahun, 20-24 tahun, dan di atas 24 tahun. Waktu berubahnya *hazard* ratio yang digunakan untuk variabel usia kawin pertama dan pendidikan secara berturut-turut yaitu usia ke-18 dan usia ke-23.

$$g_{11}(t) = \begin{cases} 1 & ; t < 18 \\ 0 & ; t \geq 18 \end{cases} \quad g_{21}(t) = \begin{cases} 1 & ; t \geq 18 \\ 0 & ; t < 18 \end{cases}$$

$$g_{12}(t) = \begin{cases} 1 & ; t < 23 \\ 0 & ; t \geq 23 \end{cases} \quad g_{22}(t) = \begin{cases} 1 & ; t \geq 23 \\ 0 & ; t < 23 \end{cases}$$

Pemodelan regresi *cox extended* usia pertama melahirkan dengan menggunakan kombinasi fungsi waktu *heavyside* menghasilkan estimasi parameter model *cox extended* seperti sebagai berikut.

Tabel 4. 6 Hasil estimasi Parameter Model *Cox Extended* Kombinasi Fungsi Waktu *Heavyside*

Variabel	Estimasi Parameter	W^2	<i>P-value</i>
Region (2)	0,0951	4,4359	0,03519
Kelompok Usia Kawin Pertama(1)* $g_{11}(t)$	-1,6360	268,4257	0,00000
Kelompok Usia Kawin Pertama(2)* $g_{11}(t)$	-5,8984	134,8087	0,00000
Kelompok Usia Kawin Pertama(3)* $g_{11}(t)$	-6,4639	41,1146	0,00000
Pendidikan(1)* $g_{12}(t)$	0,3748	3,9454	0,04700
Pendidikan(2)* $g_{12}(t)$	0,1637	0,8775	0,34888
Pendidikan(3)* $g_{12}(t)$	0,1425	0,6582	0,41721
Pendidikan(4)* $g_{12}(t)$	-0,1073	0,3647	0,54590
Pendidikan(5)* $g_{12}(t)$	-0,8371	16,9783	0,00004
Kelompok Usia Kawin Pertama(1)* $g_{21}(t)$	0,8893	12,9636	0,00032
Kelompok Usia Kawin Pertama(2)* $g_{21}(t)$	0,0010	0,0000	0,99678
Kelompok Usia Kawin Pertama(3)* $g_{21}(t)$	-1,1656	21,9191	0,00000
Pendidikan(1)* $g_{22}(t)$	0,7739	2,5210	0,11234
Pendidikan(2)* $g_{22}(t)$	1,2408	8,7732	0,00306
Pendidikan(3)* $g_{22}(t)$	1,4968	12,7357	0,00036
Pendidikan(4)* $g_{22}(t)$	2,1067	26,9710	0,00000
Pendidikan(5)* $g_{22}(t)$	2,0638	26,1180	0,00000

Keterangan: *P-value* yang dicetak tebal menunjukkan keputusan Tolak H_0 atau signifikan.

Pengujian signifikansi parameter dilakukan secara serentak dan parsial. Pengujian signifikansi parameter secara serentak digunakan untuk mengetahui apakah variabel prediktor yang

digunakan pada model berpengaruh signifikan secara serentak atau bersama-sama.

Berdasarkan hasil pengujian signifikansi parameter secara serentak terhadap model dengan menggunakan *likelihood ratio* didapatkan G_{hit}^2 dengan derajat bebas 17 yaitu sebesar 2253 dan diperoleh nilai $\chi_{(17;0,05)}^2$ yaitu sebesar 27,59. Karena nilai G_{hit}^2 lebih besar dari $\chi_{(17;0,05)}^2$ dan *p-value* kurang dari taraf signifikansi 0,05 maka diperoleh keputusan tolak H_0 yang berarti minimal terdapat satu variabel dalam model yang signifikan memengaruhi usia pertama melahirkan.

Pengujian signifikansi parameter secara parsial bertujuan untuk mengetahui pengaruh signifikansi variabel secara parsial atau individu. Hasil pengujian signifikansi parameter secara parsial/individu dapat dilihat pada Tabel 4.6. Diketahui nilai $\chi_{(1;0,05)}^2$ adalah sebesar 3,841 dan taraf signifikansi α sebesar 0,05.

Jika nilai statistik uji W^2 lebih besar dari $\chi_{(1;0,05)}^2$ dan *p-value* kurang dari α maka dapat diputuskan tolak H_0 yang berarti bahwa variabel prediktor berpengaruh secara signifikan terhadap model. Berdasarkan hasil tersebut, secara umum variabel Region, Kelompok Usia Kawin Pertama, dan Pendidikan berpengaruh secara signifikan terhadap model usia pertama melahirkan dari wanita usia subur di Jawa Timur yang pernah menikah. Status pendidikan tamat SD, tidak tamat SLTA, dan tamat SLTA menunjukkan gagal tolak H_0 , hal tersebut berarti bahwa status pendidikan tamat SD, tidak tamat SLTA, dan tamat SLTA dari wanita yang melahirkan di usia kurang dari 23 tahun tidak berpengaruh signifikan terhadap usia pertama melahirkan, terdapat kemungkinan lain yang menyebabkan status pendidikan tersebut tidak signifikan yaitu waktu berubahnya *hazard ratio* untuk kategori pendidikan ini bukan di usia 23 tahun. Dan juga kelompok usia kawin pertama yaitu 20-24 tahun dari wanita yang melahirkan di usia lebih dari 18 tahun tidak berpengaruh signifikan terhadap

usia pertama melahirkan atau waktu berubahnya *hazard ratio* untuk kelompok usia kawin pertama 20-24 tahun bukan di usia 18 tahun.

4.2.4 Pemodelan Regresi *Cox Extended* Menggunakan Kombinasi Fungsi Waktu $\ln(T)$ dan *Heavyside*

Perbedaan pemodelan regresi *cox extended* yang menggunakan kombinasi fungsi waktu *heavyside* dengan kombinasi fungsi waktu $\ln(T)$ dan *heavyside* terletak pada fungsi waktu yang diinteraksikan dengan variabel usia pertama melahirkan. Pemodelan regresi *cox extended* dengan menggunakan kombinasi fungsi waktu $\ln(T)$ dan *heavyside* pada penelitian ini merupakan pemodelan regresi *cox extended* dengan memasukkan variabel yang memenuhi asumsi *proportional hazard* yaitu variabel region, variabel usia kawin pertama yang diinteraksikan dengan fungsi $\ln(T)$, dan variabel pendidikan yang diinteraksikan dengan fungsi *heavyside*. Waktu berubahnya *hazard ratio* yang digunakan untuk variabel pendidikan yaitu usia ke-23.

$$g_{12}(t) = \begin{cases} 1 & ; t < 23 \\ 0 & ; t \geq 23 \end{cases} \quad g_{22}(t) = \begin{cases} 1 & ; t \geq 23 \\ 0 & ; t < 23 \end{cases}$$

Berikut ini merupakan hasil pemodelan regresi *cox extended* usia pertama melahirkan dengan menggunakan kombinasi fungsi waktu $\ln(T)$ dan *heavyside* berupa estimasi parameter model *cox extended*.

Tabel 4. 7 Hasil estimasi Parameter Model *Cox Extended* Kombinasi Fungsi Waktu $\ln(T)$ dan *Heavyside*

Variabel	Estimasi Parameter	W^2	<i>P-value</i>
Region (2)	-0,02910	0,2964	0,5861
Usia Kawin Pertama	-2,52225	332,5112	0,0000
Pendidikan(1)* $g_{12}(t)$	0,49015	2,4732	0,1158
Pendidikan(2)* $g_{12}(t)$	0,72246	6,9533	0,0084
Pendidikan(3)* $g_{12}(t)$	0,76270	7,1572	0,0075
Pendidikan(4)* $g_{12}(t)$	0,76185	7,3111	0,0069

Tabel 4. 7 Hasil estimasi Parameter Model *Cox Extended* Kombinasi Fungsi Waktu $\ln(T)$ dan *Heavyside* (Lanjutan)

Variabel	Estimasi Parameter	W^2	<i>P-value</i>
Pendidikan(5)* $g_{12}(t)$	0,15615	0,2675	0,6050
Usia Kawin Pertama* $\ln(T)$	0,73273	300,7700	0,0000
Pendidikan(1)* $g_{22}(t)$	0,72941	2,5443	0,1107
Pendidikan(2)* $g_{22}(t)$	1,45250	15,0758	0,0001
Pendidikan(3)* $g_{22}(t)$	1,64359	19,3024	0,0000
Pendidikan(4)* $g_{22}(t)$	1,97747	28,7282	0,0000
Pendidikan(5)* $g_{22}(t)$	1,92002	26,8295	0,0000

Keterangan: *P-value* yang dicetak tebal menunjukkan keputusan Tolak H_0 atau signifikan.

Pengujian signifikansi parameter pada pemodelan ini juga dilakukan secara serentak dan parsial. Berikut ini merupakan hasil pengujian signifikansi parameter secara serentak pada pemodelan regresi *cox extended* yang menggunakan kombinasi fungsi waktu $\ln(T)$ dan *heavyside*.

Didapatkan nilai uji *Likelihood Ratio* G_{hit}^2 yaitu sebesar 2810 dengan derajat bebas 13 dan diperoleh nilai $\chi_{(13;0,05)}^2$ yaitu sebesar 22,36. Karena nilai G_{hit}^2 lebih besar dari $\chi_{(13;0,05)}^2$ dan *p-value* kurang dari taraf signifikansi 0,05, sehingga dapat diputuskan tolak H_0 yang berarti minimal terdapat satu variabel dalam model yang signifikan memengaruhi usia pertama melahirkan.

Diketahui nilai $\chi_{(1;0,05)}^2$ sebesar 3,841 dan taraf signifikansi α sebesar 0,05. Berdasarkan hasil pengujian signifikansi parameter secara parsial/individu pada Tabel 4.7, jika nilai statistik uji W^2 lebih besar dari $\chi_{(1;0,05)}^2$ dan *p-value* kurang dari α maka dapat diputuskan tolak H_0 yang berarti bahwa variabel prediktor berpengaruh secara signifikan terhadap model. Pada pemodelan ini variabel Region tidak berpengaruh secara signifikan terhadap model sedangkan variabel Usia Kawin Pertama dan Pendidikan

berpengaruh secara signifikan terhadap model usia pertama melahirkan dari wanita usia subur di Jawa Timur yang pernah menikah. Lebih dari itu, status pendidikan tidak tamat SD dari wanita yang melahirkan pertama kali di usia kurang dari maupun lebih dari 23 tahun tidak berpengaruh signifikan terhadap usia pertama melahirkan, hal ini juga dapat diindikasikan oleh pemilihan waktu berubahnya *hazard ratio* belum tepat untuk kategori pendidikan tidak tamat SD. Selain itu, status pendidikan perguruan tinggi dari wanita yang pertama kali melahirkan di usia kurang dari 23 tahun tidak berpengaruh signifikan terhadap usia pertama melahirkan.

4.2.5 Pemilihan Kombinasi Fungsi Waktu Terbaik Berdasarkan Kriteria AICc

Model terbaik dari pemodelan Regresi *Cox Extended* merupakan model terbaik yang dihasilkan dari pemodelan yang memasukkan kombinasi fungsi waktu yang terpilih. Pemilihan kombinasi fungsi waktu ini berdasarkan kriteria kebaikan model AICc terkecil. Berikut ini disajikan pada Tabel 4.8 hasil pemilihan kombinasi fungsi waktu terbaik yang akan digunakan dalam pemodelan selanjutnya.

Tabel 4. 8 Perbandingan Nilai AICc Dari Kombinasi Fungsi Waktu yang Berbeda

Kombinasi Fungsi Waktu	Variabel	AICc
Fungsi <i>Heavyside</i>	Region, Kelompok Usia Kawin Pertama* $g_{11}(t)$, Pendidikan* $g_{12}(t)$, Kelompok Usia Kawin Pertama* $g_{21}(t)$, Pendidikan* $g_{22}(t)$.	33660,16
Fungsi $Ln(T)$ dan <i>Heavyside</i>	Region, Usia Kawin Pertama, Pendidikan* $g_{12}(t)$, Usia Kawin Pertama* $Ln(T)$, Pendidikan* $g_{22}(t)$.	33093,59

Kombinasi fungsi waktu terbaik yang akan digunakan dalam pemodelan selanjutnya untuk mendapatkan model terbaik adalah kombinasi fungsi waktu $Ln(T)$ dan *Heavyside*.

4.2.6 Pemilihan Model Terbaik Menggunakan *Backward Elimination*

Variabel yang mempengaruhi usia pertama melahirkan secara signifikan yaitu usia kawin pertama dan pendidikan, hal ini sesuai pembahasan pada sub bab 4.2.4 di atas. Untuk mendapatkan model terbaik dari pemodelan yang menggunakan kombinasi fungsi waktu $\ln(T)$ dan *Heavyside* maka dilakukan *Backward Elimination* dengan beberapa tahap hingga pemodelan terbentuk dari semua variabel yang signifikan berpengaruh. *Backward Elimination* dilakukan dengan cara mengeliminasi secara bertahap variabel yang paling tidak berpengaruh signifikan. Sehingga didapatkan model terbaik yang terdiri dari variabel yang berpengaruh signifikan terhadap model yaitu variabel Usia Pertama Melahirkan dan Pendidikan.

Tabel 4. 9 Hasil estimasi Parameter Model *Cox Extended* Terbaik

Variabel	Estimasi Parameter	W^2	<i>P-value</i>
Usia Kawin Pertama	-2,518280	325,3681	0,000000
Pendidikan(1)* $g_{12}(t)$	0,489980	2,4720	0,115890
Pendidikan(2)* $g_{12}(t)$	0,723030	6,9826	0,008231
Pendidikan(3)* $g_{12}(t)$	0,766270	7,2599	0,007051
Pendidikan(4)* $g_{12}(t)$	0,769960	7,5628	0,005959
Pendidikan(5)* $g_{12}(t)$	0,165450	0,3042	0,581253
Usia Kawin Pertama* $\ln(T)$	0,731600	294,6595	0,000000
Pendidikan(1)* $g_{22}(t)$	0,727680	2,5326	0,111514
Pendidikan(2)* $g_{22}(t)$	1,452590	15,1537	0,000099
Pendidikan(3)* $g_{22}(t)$	1,651340	19,6801	0,000009
Pendidikan(4)* $g_{22}(t)$	1,989060	29,5251	0,000000
Pendidikan(5)* $g_{22}(t)$	1,933760	27,7015	0,000000

Keterangan: *P-value* yang dicetak tebal menunjukkan keputusan Tolak H_0 atau signifikan.

Model *Cox Extended* terbaik yang terdiri dari variabel yang signifikan mempengaruhi usia pertama melahirkan dapat dituliskan seperti sebagai berikut.

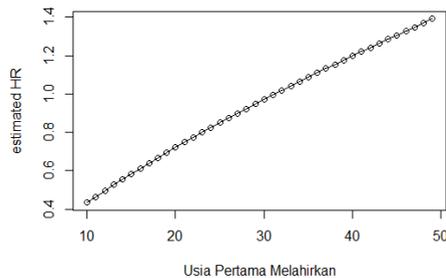
$$\begin{aligned} \hat{h}(t, \mathbf{x}) = & \hat{h}_0(t) \exp[-2,51828 \text{UsiaKawinPertama} + 0,489980 \text{Pendidikan}(1) * g_{12}(t) \\ & + 0,72303 \text{Pendidikan}(2) * g_{12}(t) + 0,766270 \text{Pendidikan}(3) * g_{12}(t) \\ & + 0,76996 \text{Pendidikan}(4) * g_{12}(t) + 0,16545 \text{Pendidikan}(5) * g_{12}(t) \\ & + 0,7316 \text{UsiaKawinPertama} * \ln(t) + 0,72768 \text{Pendidikan}(1) * g_{22}(t) \\ & + 1,45259 \text{Pendidikan}(2) * g_{22}(t) + 1,65134 \text{Pendidikan}(3) * g_{22}(t) \\ & + 1,98906 \text{Pendidikan}(4) * g_{22}(t) + 1,93376 \text{Pendidikan}(5) * g_{22}(t)] \end{aligned}$$

Nilai uji *Likelihood Ratio* G_{hit}^2 didapatkan sebesar 2810 dan derajat bebas sebesar 12. Didapatkan nilai $\chi_{(12;0,05)}^2$ yaitu sebesar 21,03. Karena nilai G_{hit}^2 2810 yaitu lebih besar dari $\chi_{(12;0,05)}^2$ 21,03 dan *p-value* kurang dari taraf signifikansi 0,05, sehingga dapat diputuskan tolak H_0 . Hal tersebut berarti minimal terdapat satu variabel dalam model yang signifikan memengaruhi usia pertama melahirkan.

Nilai $\chi_{(1;0,05)}^2$ didapatkan sebesar 3,841 dan taraf signifikansi α sebesar 0,05. Jika nilai statistik uji W^2 yang didapatkan sesuai dengan hasil Tabel 4.9 lebih besar dari $\chi_{(1;0,05)}^2$ dan *p-value* kurang dari α maka dapat diputuskan tolak H_0 yang berarti bahwa variabel prediktor berpengaruh secara signifikan terhadap model sehingga didapatkan variabel usia kawin pertama dan pendidikan berpengaruh signifikan terhadap usia pertama melahirkan. Jika status pendidikan terakhir dari wanita yang pertama kali melahirkan di usia kurang dari 23 tahun adalah perguruan tinggi maka status pendidikan tersebut tidak berpengaruh signifikan terhadap usia pertama melahirkan. Status pendidikan tidak tamat SD dari wanita yang melahirkan pertama kali di usia kurang dari maupun lebih dari 23 tahun tidak berpengaruh signifikan terhadap usia pertama melahirkan. Sehingga status pendidikan wanita tidak tamat SD maupun perguruan tinggi apabila usia wanita kurang dari

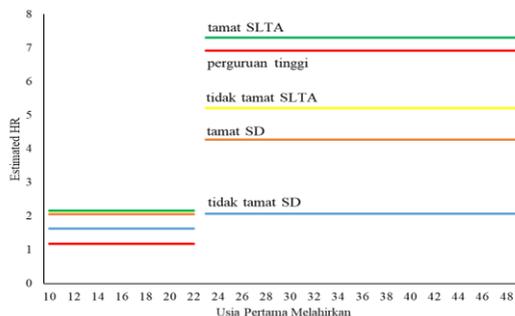
23 tahun maka kedua status pendidikan tersebut tidak berpengaruh signifikan pada usia pertama melahirkan.

Besarnya pengaruh variabel-variabel terhadap probabilitas wanita untuk tidak melahirkan yang pertama kalinya sampai lebih dari usia tertentu ditunjukkan dengan nilai *hazard ratio*. Nilai *hazard ratio* dari wanita dengan usia kawin pertama yang berubah-ubah bergantung waktu. Gambar 4.6 berikut ini menunjukkan nilai *hazard ratio* pada waktu-waktu tertentu.



Gambar 4. 6 Grafik *Hazard Ratio* Faktor Usia Kawin Pertama.

Gambar 4.6 menjelaskan bahwa semakin dewasa usia kawin pertama maka semakin dewasa usia pertama kali melahirkan yang diikuti dengan potensi yang semakin besar untuk melahirkan pertama kali. Tabel 4.10 berikut ini merupakan nilai *hazard ratio* untuk variabel pendidikan. Nilai *hazard ratio* juga dapat diilustrasikan seperti pada Gambar 4.7 untuk mempermudah dalam interpretasi *hazard ratio*.



Gambar 4. 7 Grafik *Hazard Ratio* Faktor Pendidikan.

Tabel 4. 10 *Hazard Ratio* Model Terbaik

Variabel	<i>Hazard Ratio</i> \hat{HR}
Pendidikan(1)* $g_{12}(t)$	1,6323
Pendidikan(2)* $g_{12}(t)$	2,0607
Pendidikan(3)* $g_{12}(t)$	2,1517
Pendidikan(4)* $g_{12}(t)$	2,1597
Pendidikan(5)* $g_{12}(t)$	1,1799
Pendidikan(1)* $g_{22}(t)$	2,0703
Pendidikan(2)* $g_{22}(t)$	4,2742
Pendidikan(3)* $g_{22}(t)$	5,2139
Pendidikan(4)* $g_{22}(t)$	7,3087
Pendidikan(5)* $g_{22}(t)$	6,9155

Nilai *hazard ratio* untuk variabel pendidikan tidak tamat SD maupun perguruan tinggi dari wanita yang melahirkan pertama kalinya di usia kurang dari 23 tahun yaitu mendekati 1 dan tidak signifikan, hal ini ditunjukkan pada tabel 4.9, berarti bahwa laju wanita yang tidak tamat SD maupun perguruan tinggi untuk melahirkan pertama kali di usia kurang dari 23 tahun tidak signifikan berbeda dengan wanita yang tidak sekolah.

Nilai *hazard ratio* untuk variabel pendidikan tamat SD, tidak tamat SLTA, maupun tamat SLTA dari wanita yang berusia kurang dari 23 tahun yaitu bernilai sekitar 2 sehingga dapat diartikan bahwa wanita usia subur di Jawa Timur yang tamat SD, tidak tamat SLTA, maupun tamat SLTA memiliki potensi untuk melahirkan anak pertamanya di usia kurang dari 23 tahun sebesar 2 kali lebih besar dibandingkan wanita yang tidak sekolah.

Potensi untuk melahirkan yang pertama kali dari setiap kategori status pendidikan wanita meningkat di usia lebih dari 23 tahun dibanding usia kurang dari 23 tahun. Peningkatan potensi ini bervariasi berdasarkan kategori status pendidikan wanita dan peningkatan tajam terjadi pada status pendidikan tamat SLTA dan perguruan tinggi. Wanita yang berusia lebih dari 23 tahun dan telah lulus pendidikan SD memiliki potensi untuk melahirkan anak

pertama di usia lebih dari 23 tahun 4 kali lebih besar dibandingkan wanita yang tidak sekolah. Wanita yang tidak tamat pendidikan SLTA memiliki potensi untuk melahirkan anak pertama di usia lebih dari 23 tahun 5 kali lebih besar dibandingkan wanita yang tidak sekolah. Wanita yang tamat pendidikan SLTA memiliki potensi melahirkan pertama kali di usia lebih dari 23 tahun yaitu 7 kali lebih besar dibandingkan wanita yang tidak pernah sekolah. Selain itu, wanita yang menempuh pendidikan perguruan tinggi memiliki potensi melahirkan anak pertama di usia lebih dari 23 tahun yaitu 6 kali lebih besar dibandingkan wanita yang tidak pernah sekolah.

Secara keseluruhan, semakin tinggi status pendidikan maka potensi untuk melahirkan anak pertama di usia lebih dari 23 tahun cenderung akan semakin lebih besar, terutama status pendidikan wanita tamat SLTA dan perguruan tinggi. Semakin tinggi tingkat pendidikan wanita maka semakin dewasa usia pertama kali melahirkan atau wanita semakin menunda untuk melahirkan anak pertama.

(Halaman ini sengaja dikosongkan)

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis pada data usia pertama melahirkan dari wanita usia subur di Jawa Timur yang pernah menikah didapatkan beberapa kesimpulan sebagai berikut.

1. Rata-rata wanita usia subur di Jawa Timur yang pernah menikah melahirkan pertama kali di usia 21 tahun dan setelah 2 tahun menikah. Rata-rata usia pertama melahirkan dari wanita usia subur yang pernah menikah cenderung naik searah dengan semakin tingginya tingkat pendidikan terakhir yang ditempuh. Jarak usia wanita pada saat menikah hingga 30 September 2017 yang terlalu lama untuk tidak melahirkan pertama kalinya hingga waktu berakhirnya survei terjadi pada wanita yang tidak sekolah hingga berpendidikan tamat SD. Rata-rata usia pertama melahirkan dari wanita yang tinggal di perkotaan yaitu 22 tahun, sedangkan wanita yang tinggal di pedesaan rata-rata melahirkan di usia yang lebih muda yaitu 20 tahun.
2. Kombinasi fungsi $Ln(T)$ dan *Heavyside* menjadi kombinasi fungsi waktu terbaik dan didapatkan variabel usia kawin pertama dan pendidikan sebagai faktor yang berpengaruh secara signifikan terhadap usia pertama melahirkan. Semakin dewasa usia kawin pertama maka semakin dewasa usia pertama kali melahirkan yang diikuti dengan potensi yang semakin besar untuk melahirkan pertama kali. Potensi untuk melahirkan yang pertama kali di usia lebih dari 23 tahun dari setiap kategori status pendidikan wanita mengalami peningkatan dibanding usia kurang dari 23 tahun terutama status pendidikan tamat SLTA dan perguruan tinggi. Semakin tinggi tingkat pendidikan wanita maka semakin dewasa usia pertama kali melahirkan atau wanita semakin menunda untuk melahirkan anak pertama.

5.2 Saran dan Rekomendasi

Berdasarkan hasil analisis serta kesimpulan yang diperoleh, terdapat beberapa saran untuk penelitian selanjutnya sebagai berikut.

1. Menggunakan metode lain seperti Regresi *Cox Statified* agar dapat dibandingkan dan didapatkan model terbaik dari usia pertama melahirkan.
2. Menggunakan data usia pertama melahirkan beserta variabel yang relevan atau variabel yang diukur pada saat wanita melahirkan pertama kali agar lebih banyak variabel yang dapat menjelaskan karakteristik dari usia pertama melahirkan.

Rekomendasi yang dapat diajukan sebagai bahan pertimbangan program kepada Badan Kependudukan dan Keluarga Berencana Nasional (BKKBN) Provinsi Jawa Timur yaitu Program Pendewasaan Usia Kawin masih sangat penting dilakukan khususnya untuk wilayah Provinsi Jawa Timur, sosialisasi masif terkait edukasi perencanaan sejak dini dan edukasi kesehatan reproduksi pada remaja melalui program GenRe (Generasi Berencana atau Generasi yang Punya Rencana) untuk setiap status pendidikan wanita usia subur terutama wanita yang tidak sekolah hingga wanita yang tidak tamat SLTA perlu untuk digalakkan.

DAFTAR PUSTAKA

- Adebowale, S., Fagbamigbe, F., Okareh, T., & Lawal, G. (2012). *Survival Analysis of Timing of First Marriage among Women of Reproductive age in Nigeria: Regional Differences. Afr J Reprod Health*, 95-107.
- Adhikari, R. (2010). Demographic, Socio-Economic, and Cultural Factors Affecting Fertility Differentials in Nepal. *BMC Pregnancy and Childbirth*, 10(1), 19-25.
- Aini, I. N. (2011). Extended Cox Model Untuk Time Independent Covariat yang Tidak Memenuhi Asumsi Proportional Hazard Pada Model Cox Proportional Hazard. *Skripsi Universitas Indonesia*.
- Akaike, H. (1973). Information Theory as an Extension of The Maximum Likelihood. *Petrov, B.N., Csaki, F. (Eds.), Second International Symposium* (pp. 267-281). Budapest: Akademiai Kiado.
- Baird, D., Collins, J., Egozcue, J., Evers, L., Gianaroli, L., & Leridon, H. (2005). Fertility and Ageing. *Hum Reprod Update*, 11:261-76.
- BKKBN. (2018). *Survei Demografi dan Kesehatan Indonesia 2017*. Jakarta: BKKBN.
- Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practica Information-Theoretic Approach* (2nd ed.). New York: Springer.
- Collett, D. (2003). *Modelling Survival Data in Medical Research (2nd ed)*. London: Chapman and Hall.
- Cox, D. R. (1972). Regression Model and Life Table. *J Roy Stat Soc B*, 34, 187-202.
- Fagbamigbe, A. F., & Idemudia, E. S. (2016). *Survival Analysis and Prognostic Factors of Timing of First Childbirth among Women in Nigeria. BMC Pregnancy and Childbirth*, 16, 102.
- George, K., & Kamath, M. (2010). Fertility and Age. *J Hum Reprod Sci.*, 121-123.

- Herowati, D., Kurniawan, A., Sujarwoto, & Sugiharto, M. (2018). *Survei Demografi dan Kesehatan Indonesia Provinsi Jawa Timur 2017*. Surabaya: Perwakilan BKKBN Provinsi Jawa Timur.
- Hosmer, D. W., Lameshow, S., & May, S. (2008). *Applied Survival Analysis: Regression Modelling of Time Event Data*. New Jersey: John Wiley.
- ICF. (2018). Demographic and Health Surveys Standard Recode Manual for DHS7. In *Demographic and Health Surveys Program*. Rockville, Maryland, U.S.A: ICF.
- Johnson, R. A., & Bhattacharyya, G. K. (2010). *Statistics Principles and Methods 6th Edition*. United State of America: John Wiley & Sons Inc.
- Kleinbaum, D. G., & Klein, M. (2012). *Statistics for Biology and Health Survival Analysis: A Self-Learning Text Third Edition*. New York: Springer.
- Konishi, S., & Kitagawa, G. (2008). *Information Criteria and Statistical Modeling*. Jepang: Springer.
- Mantra, I. (2003). *Demografi Umum*. Yogyakarta: Pustaka Pelajar.
- Matthew, e. a. (2012). Modelling the Determinants of Fertility among Women of Chirdbearing Age in Nigeria: Analysis Using Generalized Linier Modelling Approach. *International Journal of Humanities and Social Science*, Vol. 2(18), 167-176.
- Nevid, R., & Greene. (2005). Psikologi abnormal. In *Jilid I. AlihBahasa, TimPsikologi UI* (p. 234). Jakarta: Erlangga.
- Noorkasiani, Heryati, & Ismail, R. (2009). *Sosiologi Keperawatan*. Jakarta: EGC.
- Rao, K. V., & Balakrishnan, T. R. (1988). Age at First Birth in Canada: A Hazard Model Analysis. *Genus*, 53-72.
- Rindfuss, R., & John, C. (1983). Social Determinants of Age at First Birth. *Journal of Marriage and The Family*, 553.
- Rusli. (1996). *Pengantar Ilmu Kependudukan*. Jakarta: LP3S.
- Sri, H. (2010). *Fertilitas: Dasar-Dasar Demografi*. Jakarta: Lembaga Demografi Fakultas Ekonomi UI.

- UNICEF. (2008). *The State of The World's Children 2009*. New York: United Nations Children's Fund.
- Walpole, R. E. (1995). *Pengantar Statistika Edisi Ke-3*. Jakarta: PT Gramedia Pustaka Utama.
- Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2012). *Probability & Statistics for Engineers & Scientists* (9th ed.). United States of America: Pearson Education Inc.

(Halaman ini sengaja dikosongkan)

LAMPIRAN

Lampiran 1. Data Wanita Usia Subur di Jawa Timur yang Pernah Menikah

id	T	d	X1	X2	X3	Weight
1	16	1	16	2	2	2287801
2	22	1	20	3	2	2287801
3	24	1	23	3	2	2287801
4	17	1	15	3	2	2287801
5	22	1	21	2	2	2287801
6	20	1	19	3	2	2287801
7	18	1	17	2	2	2287801
8	17	1	15	2	2	2287801
9	20	1	19	3	2	2287801
10	39	1	22	3	2	2287801
⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮
2747	33	1	29	5	1	1386677
2748	22	1	22	5	1	1386677
2749	25	1	24	5	1	1386677
2750	28	1	26	5	1	1386677
2751	20	1	19	4	1	1386677
2752	30	1	29	5	1	1386677
2753	27	1	26	5	1	1386677
2754	27	1	26	5	1	1386677
2755	24	1	16	3	1	1386677
2756	21	0	21	4	1	1386677

Lampiran 2. *Syntax R Statistika Deskriptif*

```
dta<-read.csv("E:/2018_2019/SEMESTER 7/TA/OUTPUT
NEW/dataku.csv",sep=";")
View(dta)
dim(dta)
dta$wt<-dta$Weight/1000000
dta1<-filter(dta,d==1)
dta0<-filter(dta,d==0)
library(plyr)
install.packages("radiant.data")
library(radiant.data)
ddply(dta,~d,summarise,mean=weighted.mean(T,wt),sd=weighted.sd(T,wt))
ddply(dta,~d,summarise,mean=weighted.mean(X1,wt),sd=weighted.sd(X1,wt))
ddply(dta,~X2,summarise,mean=weighted.mean(T,wt),sd=weighted.sd(T,wt))
ddply(dta1,~X2,summarise,mean=weighted.mean(T,wt),sd=weighted.sd(T,wt))
ddply(dta0,~X2,summarise,mean=weighted.mean(T,wt),sd=weighted.sd(T,wt))
ddply(dta,~X3,summarise,mean=weighted.mean(T,wt),sd=weighted.sd(T,wt))
ddply(dta1,~X3,summarise,mean=weighted.mean(T,wt),sd=weighted.sd(T,wt))
ddply(dta0,~X3,summarise,mean=weighted.mean(T,wt),sd=weighted.sd(T,wt))
```

Lampiran 3. *Syntax R Kaplan Meier*

```
Y<-Surv(Dataku$T,Dataku$d==1)
#Kurva KM#
kmfit1<-survfit(Y~1)
summary(kmfit1)
ggsurvplot(kmfit1,data=Dataku,surv.median.line="hv",pval=F,conf.int = F,
xlab="Usia Pertama Melahirkan", legend="none", palette =
"black",break.time.by=2)

kmfitX2<-survfit(Y~Dataku$X2)
summary(kmfitX2)
ggsurvplot(kmfitX2,data=Dataku,surv.median.line="hv",pval=F,conf.int = F,
xlab="Usia Pertama Melahirkan",break.time.by=2,legend.labs=c("Tidak
Sekolah","Tidak tamat SD","Tamat SD","Tidak tamat SLTA","Tamat
SLTA","Perguruan Tinggi"))

kmfitX3<-survfit(Y~Dataku$X3)
summary(kmfitX3)
ggsurvplot(kmfitX3,data=Dataku,surv.median.line="hv",pval=F,conf.int = F,
xlab="Usia Pertama
Melahirkan",break.time.by=2,legend.labs=c("Perkotaan","Pedesaan"))
```

Lampiran 4. *Syntax* R Uji Log-Rank

```
#Uji LOG-RANK#
LRX2<-survdifff(Y~Dataku$X2)
LRX2

LRX3<-survdifff(Y~Dataku$X3)
LRX3
```

Lampiran 5. *Syntax* R Uji Asumsi PH Menggunakan Kurva $-\ln[-\ln S(t)]$

```
#kurva -log(-logS(t))
s_kmfitX2<-summary(kmfitX2)
d_kmfitX2<-data.frame(s_kmfitX2$strata,s_kmfitX2$time,s_kmfitX2$surv)
names(d_kmfitX2)<-c("Pendidikan","Usia Pertama Melahirkan","KM Survival Estimate")
Pendidikan0<-d_kmfitX2[d_kmfitX2$Pendidikan=="Dataku$X2=0", ]
Pendidikan1<-d_kmfitX2[d_kmfitX2$Pendidikan=="Dataku$X2=1", ]
Pendidikan2<-d_kmfitX2[d_kmfitX2$Pendidikan=="Dataku$X2=2", ]
Pendidikan3<-d_kmfitX2[d_kmfitX2$Pendidikan=="Dataku$X2=3", ]
Pendidikan4<-d_kmfitX2[d_kmfitX2$Pendidikan=="Dataku$X2=4", ]
Pendidikan5<-d_kmfitX2[d_kmfitX2$Pendidikan=="Dataku$X2=5", ]
X2y0<-log(-log(Pendidikan0$`KM Survival Estimate`))
X2y1<-log(-log(Pendidikan1$`KM Survival Estimate`))
X2y2<-log(-log(Pendidikan2$`KM Survival Estimate`))
X2y3<-log(-log(Pendidikan3$`KM Survival Estimate`))
X2y4<-log(-log(Pendidikan4$`KM Survival Estimate`))
X2y5<-log(-log(Pendidikan5$`KM Survival Estimate`))
plot(Pendidikan0$`Usia Pertama Melahirkan`,X2y0,xlab="Usia Pertama Melahirkan",ylab="-ln(-ln(S(t)))",ylim=range(c(X2y0,X2y1,X2y2,X2y3,X2y4,X2y5)),xlim=range(c(Pendidikan0$`Usia Pertama Melahirkan`,Pendidikan1$`Usia Pertama Melahirkan`,Pendidikan2$`Usia Pertama Melahirkan`,Pendidikan3$`Usia Pertama Melahirkan`,Pendidikan4$`Usia Pertama Melahirkan`,Pendidikan5$`Usia Pertama Melahirkan`)),col="red",type="s",lty="solid")
par(new=T)
plot(Pendidikan1$`Usia Pertama Melahirkan`,X2y1,ylim=range(c(X4y0,X4y1,X4y2,X4y3,X4y4,X4y5)),xlim=range(c(Pendidikan0$`Usia Pertama Melahirkan`,Pendidikan1$`Usia Pertama Melahirkan`,Pendidikan2$`Usia Pertama Melahirkan`,Pendidikan3$`Usia Pertama Melahirkan`,Pendidikan4$`Usia Pertama Melahirkan`,Pendidikan5$`Usia Pertama Melahirkan`)),xlab="Usia Pertama Melahirkan",ylab="-ln(-ln(S(t)))",col="chocolate4",type="s",lty="solid")
par(new=T)
```

Lampiran 5. *Syntax* R Uji Asumsi PH Menggunakan Kurva $-\ln[-\ln S(t)]$ (Lanjutan)

```

plot(Pendidikan2$`Usia Pertama
Melahirkan`,X2y2,ylim=range(c(X4y0,X4y1,X4y2,X4y3,X4y4,X4y5)),xlim=range(
c(Pendidikan0$`Usia Pertama Melahirkan`,Pendidikan1$`Usia Pertama
Melahirkan`,Pendidikan2$`Usia Pertama Melahirkan`,Pendidikan3$`Usia Pertama
Melahirkan`,Pendidikan4$`Usia Pertama Melahirkan`,Pendidikan5$`Usia Pertama
Melahirkan`)),xlab="Usia Pertama Melahirkan",ylab="-ln(-
ln(S(t)))",col="green",type="s",lty="solid")
par(new=T)
plot(Pendidikan3$`Usia Pertama
Melahirkan`,X2y3,ylim=range(c(X4y0,X4y1,X4y2,X4y3,X4y4,X4y5)),xlim=range(
c(Pendidikan0$`Usia Pertama Melahirkan`,Pendidikan1$`Usia Pertama
Melahirkan`,Pendidikan2$`Usia Pertama Melahirkan`,Pendidikan3$`Usia Pertama
Melahirkan`,Pendidikan4$`Usia Pertama Melahirkan`,Pendidikan5$`Usia Pertama
Melahirkan`)),xlab="Usia Pertama Melahirkan",ylab="-ln(-
ln(S(t)))",col="deepskyblue",type="s",lty="solid")
par(new=T)
plot(Pendidikan4$`Usia Pertama
Melahirkan`,X2y4,ylim=range(c(X4y0,X4y1,X4y2,X4y3,X4y4,X4y5)),xlim=range(
c(Pendidikan0$`Usia Pertama Melahirkan`,Pendidikan1$`Usia Pertama
Melahirkan`,Pendidikan2$`Usia Pertama Melahirkan`,Pendidikan3$`Usia Pertama
Melahirkan`,Pendidikan4$`Usia Pertama Melahirkan`,Pendidikan5$`Usia Pertama
Melahirkan`)),xlab="Usia Pertama Melahirkan",ylab="-ln(-
ln(S(t)))",col="blue",type="s",lty="solid")
par(new=T)
plot(Pendidikan5$`Usia Pertama
Melahirkan`,X2y5,ylim=range(c(X4y0,X4y1,X4y2,X4y3,X4y4,X4y5)),xlim=range(
c(Pendidikan0$`Usia Pertama Melahirkan`,Pendidikan1$`Usia Pertama
Melahirkan`,Pendidikan2$`Usia Pertama Melahirkan`,Pendidikan3$`Usia Pertama
Melahirkan`,Pendidikan4$`Usia Pertama Melahirkan`,Pendidikan5$`Usia Pertama
Melahirkan`)),xlab="Usia Pertama Melahirkan",ylab="-ln(-
ln(S(t)))",col="deeppink",type="s",lty="solid")
legend("topright",c("Tidak sekolah","Tidak tamat SD","Tamat SD","Tidak tamat
SLTA","Tamat SLTA","Perguruan
Tinggi"),lty=c("solid","solid","solid","solid","solid","solid"),col=c("red","chocolate4
","green","deepskyblue","blue","deeppink"),lwd=3,cex=0.8)
par(new=F)

```

Lampiran 5. *Syntax R Uji Asumsi PH Menggunakan Kurva $-\ln[-\ln S(t)]$ (Lanjutan)*

```
s_kmfitX3<-summary(kmfitX3)
d_kmfitX3<-data.frame(s_kmfitX3$strata,s_kmfitX3$time,s_kmfitX3$surv)
names(d_kmfitX3)<-c("Region","Usia Pertama Melahirkan","KM Survival Estimate")
Region1<-d_kmfitX3[d_kmfitX3$Region=="Dataku$X3=1", ]
Region2<-d_kmfitX3[d_kmfitX3$Region=="Dataku$X3=2", ]
X3y1<-log(-log(Region1$`KM Survival Estimate`))
X3y2<-log(-log(Region2$`KM Survival Estimate`))
plot(Region1$`Usia Pertama Melahirkan`,X3y1,xlab="Usia Pertama Melahirkan",ylab="-ln(-ln(S(t)))",ylim=range(c(X3y1,X3y2)),xlim=range(c(Region1$`Usia Pertama Melahirkan`,Region2$`Usia Pertama Melahirkan`)),col="red",type="s",lty="solid")
par(new=T)
plot(Region2$`Usia Pertama Melahirkan`,X3y2,ylim=range(c(X3y1,X3y2)),xlim=range(c(Region1$`Usia Pertama Melahirkan`,Region2$`Usia Pertama Melahirkan`)),xlab="Usia Pertama Melahirkan",ylab="-ln(-ln(S(t)))",col="blue",type="s",lty="solid")
legend("topright",
c("Perkotaan","Pedesaan"),lty=c("solid","solid"),col=c("red","blue"),lwd=3,cex=0.8)
par(new=F)
```

Lampiran 6. *Syntax R Pemodelan Cox PH dan Uji Asumsi PH Menggunakan Uji Goodness of Fit*

```
##PEMODELAN COX PH##
mod1<-coxph(Y~X1+X2+X3,data=Dataku,method="breslow")
summary(mod1)

##Uji GOF##
gof.mod1<-cox.zph(mod1,transform="rank")
gof.mod1
```

Lampiran 7. *Syntax R Regresi Cox Extended Fungsi Waktu Heavyside*

```

##PEMODELAN COX EXTENDED HEAVYSIDE SEMUA##
#martingale residual#
martingale<-resid(mod1,type='martingale')
plot(Dataku$X1, martingale,xlab="Usia Kawin Pertama", ylab="Martingale
Residuals")
lines(lowess(Dataku$X1, martingale),col='red')
abline(v=c(15,20,25),lty="dotted", lwd=2, col="darkgrey")
#1:kurang dari 15 tahun
#2:15-19 tahun
#3:20-24 tahun
#4:di atas 24 tahun
Datakunew<-read.csv("E:/2018_2019/SEMESTER 7/TA/OUTPUT
NEW/datakunew.csv",sep=";")
View(Datakunew)
Datakunew$Id<-as.factor(Datakunew$Id)
Datakunew$d<-as.factor(Datakunew$d)
Datakunew$X2<-as.factor(Datakunew$X2)
Datakunew$X3<-as.factor(Datakunew$X3)
Datakunew$codeX1<-as.factor(Datakunew$codeX1)
#PEMODELAN COX PH#
Y<-Surv(Datakunew$T,Datakunew$d==1)
mod2<-coxph(Y~codeX1+X2+X3,data=Datakunew,method="breslow")
summary(mod2)
#Uji GOF#
gof.mod2<-cox.zph(mod2,transform="rank")
gof.mod2
#KM#
kmfitX1<-survfit(Y~Datakunew$codeX1)
summary(kmfitX1)
ggsurvplot(kmfitX1,data=Datakunew,surv.median.line="hv",pval=F,conf.int
=F, xlab="Usia Pertama
Melahirkan",break.time.by=2,legend.labs=c("Kurang dari 15 tahun","15-19
tahun","20-24 tahun","Di atas 24 tahun"))

```

Lampiran 7. *Syntax R Regresi Cox Extended Fungsi Waktu Heavyside (Lanjutan)*

```
#LOG RANK#
LRcodeX1<-survdiff(Y~Datakunew$codeX1)
LRcodeX1
#KURVA LOG-LOG#
s_kmfitX1<-summary(kmfitX1)
d_kmfitX1<-data.frame(s_kmfitX1$strata,s_kmfitX1$time,s_kmfitX1$surv)
names(d_kmfitX1)<-c("Kelompok Usia Kawin Pertama","Usia Pertama
Melahirkan","KM Survival Estimate")
KelUKP0<-d_kmfitX1[d_kmfitX1$`Kelompok Usia Kawin
Pertama`=="Datakunew$codeX1=0", ]
KelUKP1<-d_kmfitX1[d_kmfitX1$`Kelompok Usia Kawin
Pertama`=="Datakunew$codeX1=1", ]
KelUKP2<-d_kmfitX1[d_kmfitX1$`Kelompok Usia Kawin
Pertama`=="Datakunew$codeX1=2", ]
KelUKP3<-d_kmfitX1[d_kmfitX1$`Kelompok Usia Kawin
Pertama`=="Datakunew$codeX1=3", ]
X1y0<-log(-log(-log(KelUKP0$`KM Survival Estimate`)))
X1y1<-log(-log(-log(KelUKP1$`KM Survival Estimate`)))
X1y2<-log(-log(-log(KelUKP2$`KM Survival Estimate`)))
X1y3<-log(-log(-log(KelUKP3$`KM Survival Estimate`)))
plot(KelUKP0$`Usia Pertama Melahirkan`,X1y0,xlab="Usia Pertama
Melahirkan",ylab="-ln(-
ln(S(t)))",ylim=range(c(X1y0,X1y1,X1y2,X1y3)),xlim=range(c(KelUKP0$`Usia
Pertama Melahirkan`,KelUKP1$`Usia Pertama Melahirkan`,KelUKP2$`Usia Pertama
Melahirkan`,KelUKP3$`Usia Pertama Melahirkan`)),col="red",type="s",lty="solid")
par(new=T)
plot(KelUKP1$`Usia Pertama
Melahirkan`,X1y1,ylim=range(c(X1y0,X1y1,X1y2,X1y3)),xlim=range(c(KelUKP0$`
Usia Pertama Melahirkan`,KelUKP1$`Usia Pertama Melahirkan`,KelUKP2$`Usia
Pertama Melahirkan`,KelUKP3$`Usia Pertama Melahirkan`)),xlab="Usia Pertama
Melahirkan",ylab="-ln(-ln(S(t)))",col="green",type="s",lty="solid")
par(new=T)
plot(KelUKP2$`Usia Pertama
Melahirkan`,X1y2,ylim=range(c(X1y0,X1y1,X1y2,X1y3)),xlim=range(c(KelUKP0$`
Usia Pertama Melahirkan`,KelUKP1$`Usia Pertama Melahirkan`,KelUKP2$`Usia
Pertama Melahirkan`,KelUKP3$`Usia Pertama Melahirkan`)),xlab="Usia Pertama
Melahirkan",ylab="-ln(-ln(S(t)))",col="blue",type="s",lty="solid")
par(new=T)
plot(KelUKP3$`Usia Pertama
Melahirkan`,X1y3,ylim=range(c(X1y0,X1y1,X1y2,X1y3)),xlim=range(c(KelUKP0$`
Usia Pertama Melahirkan`,KelUKP1$`Usia Pertama Melahirkan`,KelUKP2$`Usia
Pertama Melahirkan`,KelUKP3$`Usia Pertama Melahirkan`)),xlab="Usia Pertama
Melahirkan",ylab="-ln(-ln(S(t)))",col="darkorchid1",type="s",lty="solid")
legend("topright",c("Kurang dari 15 tahun","15-19 tahun","20-24 tahun","Di atas 24
tahun"),lty=c("solid","solid","solid","solid"),col=c("red","green","blue","darkorchid1"
),lwd=3,cex=0.8)
par(new=F)
```

Lampiran 7. *Syntax R Regresi Cox Extended Fungsi Waktu Heavyside (Lanjutan)*

```

Datakunew<-read.csv("E:/2018_2019/SEMESTER 7/TA/OUTPUT
NEW/datakunew.csv",sep=";")
Datakunew3<-survSplit(Datakunew,cut = c(18,23),end="T",
event="d",start="start")
Datakunew3$X1hv1=Datakunew3$codeX1*(Datakunew3$start<18)
Datakunew3$X1hv2=Datakunew3$codeX1*(Datakunew3$start>=18)
Datakunew3$X2hv1=Datakunew3$X2*(Datakunew3$start<23)
Datakunew3$X2hv2=Datakunew3$X2*(Datakunew3$start>=23)

Y2=Surv(Datakunew3$start,Datakunew3$T,Datakunew3$d)
modske2<-coxph(Y2 ~
as.factor(X3)+as.factor(X1hv1)+as.factor(X2hv1)+as.factor(X1hv2)+as.fact
or(X2hv2)+cluster(id),data=Datakunew3,method = "breslow")
summary(modske2)

extractAIC(modske2)
AICc(modske2, return.K = FALSE, second.ord = TRUE, nobs = NULL)

```

Lampiran 8. *Syntax R Regresi Cox Extended Fungsi Waktu Ln(t) dan Heavyside*

```
##PEMODELAN COX EXTENDED LN(t) dan HEAVYSIDE##
Dataku<-read.csv("E:/2018_2019/SEMESTER 7/TA/OUTPUT
NEW/datakunew.csv",sep=";")
Dataku2<-survSplit(Dataku,cut = c(23),end="T", event="d",start="start")
cut.points<-unique(Dataku2$T[Dataku$d==1])
Datakunew4=survSplit(Dataku2,cut=cut.points,end="T",start="start",
event="d")
View(Dataku2)
View(Datakunew4)
Datakunew4$logtX1=Datakunew4$X1*log(Datakunew4$T)
Datakunew4$X2hv1=Datakunew4$X2*(Datakunew4$start<23)
Datakunew4$X2hv2=Datakunew4$X2*(Datakunew4$start>=23)

Y3=Surv(Datakunew4$start,Datakunew4$T,Datakunew4$d)
modske3<-coxph(Y3 ~
as.factor(X3)+X1+as.factor(X2hv1)+logtX1+as.factor(X2hv2)+cluster(id),d
ata=Datakunew4,method = "breslow")
summary(modske3) #model lengkap

modske3_1<-coxph(Y3 ~
X1+as.factor(X2hv1)+logtX1+as.factor(X2hv2)+cluster(id),data=Datakune
w4,method = "breslow")
summary(modske3_1) #tanpa X3

extractAIC(modske3) #model lengkap
extractAIC(modske3_1) #tanpa X3
AICc(modske3, return.K = FALSE, second.ord = TRUE, nobs = NULL)
AICc(modske3_1, return.K = FALSE, second.ord = TRUE, nobs = NULL)
```

Lampiran 9. *Syntax R Regresi Cox Extended Fungsi Waktu T dan Heavyside*

```
##PEMODELAN COX EXTENDED t dan HEAVYSIDE##
Dataku<-read.csv("E:/2018_2019/SEMESTER 7/TA/OUTPUT
NEW/datakunew.csv",sep=";")
Dataku3<-survSplit(Dataku,cut = c(23),end="T", event="d",start="start")
cut.points<-unique(Dataku3$T[Dataku$d==1])
Datakunew5=survSplit(Dataku3,cut=cut.points,end="T",start="start",
event="d")
View(Dataku3)
View(Datakunew5)
Datakunew5$tX1=Datakunew5$X1*Datakunew5$T
Datakunew5$X2hv1=Datakunew5$X2*(Datakunew5$start<23)
Datakunew5$X2hv2=Datakunew5$X2*(Datakunew5$start>=23)

Y4=Surv(Datakunew5$start,Datakunew5$T,Datakunew5$d)
modske4<-coxph(Y4 ~
as.factor(X3)+X1+as.factor(X2hv1)+tX1+as.factor(X2hv2)+cluster(id),data
=Datakunew5,method = "breslow")
summary(modske4) #model lengkap

modske4_1<-coxph(Y4 ~
X1+as.factor(X2hv1)+tX1+as.factor(X2hv2)+cluster(id),data=Datakunew5,
method = "breslow")
summary(modske4_1) #tanpa X3

extractAIC(modske4) #model lengkap
extractAIC(modske4_1) #tanpa X3
AICc(modske4, return.K = FALSE, second.ord = TRUE, nobs = NULL)
AICc(modske4_1, return.K = FALSE, second.ord = TRUE, nobs = NULL)
```

Lampiran 10. *Syntax R Kurva Estimated Hazard Ratio Variabel Usia Kawin Pertama*

```
##Kuva HR UKP untuk fungsi waktu Ln(t)##
beta<-modske3_1$coefficients[1]
psi<-modske3_1$coefficients[7]
T<-sort(Dataku$T)
HR_X1ln=exp(beta+psi*log(T))
plot(T,HR_X1ln,xlab="Usia Pertama Melahirkan",ylab="estimated
HR",col="black",type="o",lty="solid",lwd=1.5)

##Kuva HR UKP untuk fungsi waktu T##
beta<-modske4_1$coefficients[1]
psi<-modske4_1$coefficients[7]
T<-sort(Dataku$T)
HR_X1t=exp(beta+psi*T)
plot(T,HR_X1t,xlab="Usia Pertama Melahirkan",ylab="estimated
HR",col="black",type="o",lty="solid",lwd=1.5)
```

Lampiran 11. *Output Statistika Deskriptif*

```
> ddp1y(dta,~d, summarise, mean=weighted.mean(T,wt), sd=weighted.sd(T,wt))
d      mean      sd
1 0 29.62930 9.623592
2 1 21.61802 4.483751
> ddp1y(dta,~d, summarise, mean=weighted.mean(X1,wt), sd=weighted.sd(X1,wt))
d      mean      sd
1 0 22.65397 6.275868
2 1 19.84384 4.573694
> ddp1y(dta,~x2, summarise, mean=weighted.mean(T,wt), sd=weighted.sd(T,wt))
x2     mean      sd
1 0 22.28978 10.882096
2 1 20.01164  6.724213
3 2 20.77852  5.150114
4 3 21.39272  4.780124
5 4 23.74029  4.722277
6 5 26.25633  4.774993
> ddp1y(dta1,~x2, summarise, mean=weighted.mean(T,wt), sd=weighted.sd(T,wt))
x2     mean      sd
1 0 18.35444  5.149794
2 1 18.75650  4.327707
3 2 20.29783  4.191412
4 3 20.98582  3.919016
5 4 23.25842  3.807318
6 5 25.63803  3.882096
> ddp1y(dta0,~x2, summarise, mean=weighted.mean(T,wt), sd=weighted.sd(T,wt))
x2     mean      sd
1 0 46.42775  2.113592
2 1 39.30165  7.480293
3 2 30.36598 10.422063
4 3 25.86391  9.129736
5 4 28.47443  8.618563
6 5 29.26739  7.036804
```

Lampiran 11. Output Statistika Deskriptif (Lanjutan)

```
> ddply(dta,~X3,summarise,mean=weighted.mean(T,wt),sd=weighted.sd(T,wt))
  X3   mean      sd
1  1  23.63972  5.826557
2  2  21.12312  5.098762
> ddply(dta1,~X3,summarise,mean=weighted.mean(T,wt),sd=weighted.sd(T,wt))
  X3   mean      sd
1  1  22.81891  4.586710
2  2  20.57836  4.118995
> ddply(dta0,~X3,summarise,mean=weighted.mean(T,wt),sd=weighted.sd(T,wt))
  X3   mean      sd
1  1  31.59930  9.447715
2  2  27.55332  9.367431
```

Lampiran 12. Output R Uji Log-Rank

Call:

```
survdifff(formula = Y ~ Dataku$X2)
```

	N	Observed	Expected	(O-E) ² /E	(O-E) ² /V
Dataku\$X2=0	68	59	57.3	0.0494	0.059
Dataku\$X2=1	222	208	136.6	37.3017	45.358
Dataku\$X2=2	732	696	496.9	79.7623	115.064
Dataku\$X2=3	662	607	498.6	23.5520	34.009
Dataku\$X2=4	752	684	842.6	29.8658	52.275
Dataku\$X2=5	320	265	486.9	101.1177	149.348

Chisq= 324 on 5 degrees of freedom, p= <2e-16

Call:

```
survdifff(formula = Y ~ Dataku$X3)
```

	N	Observed	Expected	(O-E) ² /E	(O-E) ² /V
Dataku\$X3=1	1413	1282	1530	40.3	120
Dataku\$X3=2	1343	1237	989	62.4	120

Chisq= 120 on 1 degrees of freedom, p= <2e-16

Call:

```
survdifff(formula = Y ~ Datakunew$codeX1)
```

	N	Observed	Expected	(O-E) ² /E	(O-E) ² /V
Datakunew\$codeX1=0	224	212	54.2	459.1	531.8
Datakunew\$codeX1=1	1175	1106	543.0	583.6	966.8
Datakunew\$codeX1=2	905	821	985.0	27.3	55.9
Datakunew\$codeX1=3	452	380	936.7	330.9	752.8

Chisq= 1863 on 3 degrees of freedom, p= <2e-16

Lampiran 13. Output R Regresi Cox Proportional Hazard

```
> summary(mod1)
Call:
coxph(formula = Y ~ x1 + x2 + x3, data = Dataku, method = "breslow")

n= 2756, number of events= 2519

      coef exp(coef) se(coef)      z Pr(>|z|)
X1  -0.19101  0.82612  0.00579 -32.988 < 2e-16 ***
X21  0.81199  2.25239  0.14871   5.460 4.76e-08 ***
X22  0.96954  2.63672  0.13961   6.945 3.79e-12 ***
X23  0.97001  2.63796  0.14169   6.846 7.60e-12 ***
X24  0.95234  2.59176  0.14515   6.561 5.34e-11 ***
X25  0.80567  2.23819  0.15578   5.172 2.32e-07 ***
X32 -0.01535  0.98477  0.04342  -0.353  0.724
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

      exp(coef) exp(-coef) lower .95 upper .95
X1      0.8261      1.2105   0.8168   0.8356
X21     2.2524      0.4440   1.6829   3.0146
X22     2.6367      0.3793   2.0055   3.4666
X23     2.6380      0.3791   1.9983   3.4824
X24     2.5918      0.3858   1.9500   3.4447
X25     2.2382      0.4468   1.6493   3.0374
X32     0.9848      1.0155   0.9044   1.0723

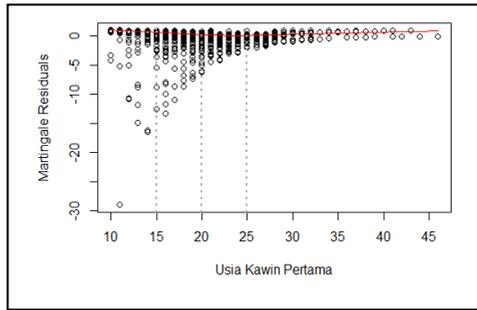
Concordance= 0.882 (se = 0.005 )
Rsquare= 0.44 (max possible= 1 )
Likelihood ratio test= 1598 on 7 df,  p=<2e-16
Wald test              = 1471 on 7 df,  p=<2e-16
Score (logrank) test = 1421 on 7 df,  p=<2e-16
```

Lampiran 14. Output R Uji Asumsi PH Menggunakan Uji Goodness of Fit

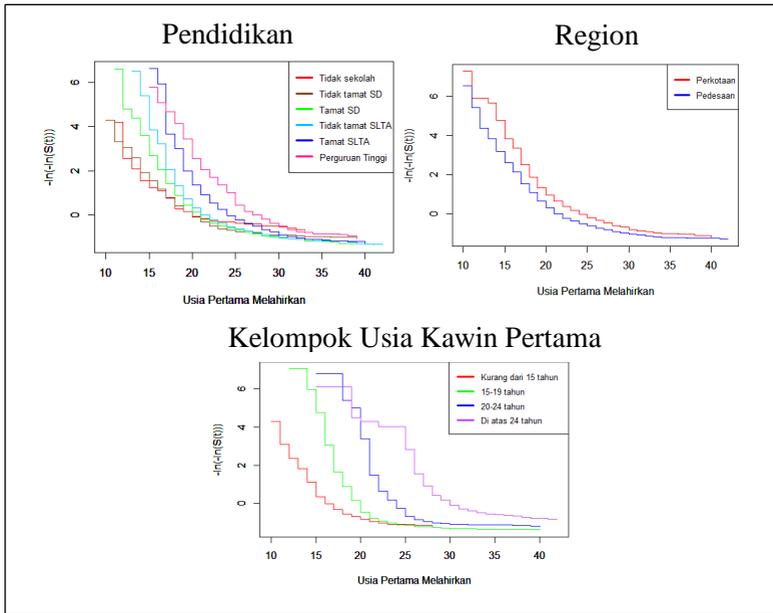
```
> gof.mod1

      rho      chisq      p
X1      0.5187  644.61 3.32e-142
X21     0.0387   3.83 5.04e-02
X22     0.0774  15.88 6.76e-05
X23     0.0911  22.00 2.73e-06
X24     0.1049  29.43 5.80e-08
X25     0.0913  22.16 2.51e-06
X32     0.0347   3.12 7.72e-02
GLOBAL      NA 1160.15 2.93e-246
```

Lampiran 15. Grafik *Martingale Residual*



Lampiran 16. Grafik $-\ln[-\ln S(t)]$



Lampiran 17. Output R Regresi Cox Extended Fungsi Waktu Heavyside

```

> summary(mod2)
Call:
coxph(formula = Y ~ codex1 + x2 + x3, data = Datakunew, method = "breslow")

n = 2756, number of events = 2519

      coef exp(coef) se(coef)      z Pr(>|z|)
codex11 -0.63146  0.53181  0.08232  -7.671 1.71e-14 ***
codex12 -1.70913  0.18102  0.08850 -19.313 < 2e-16 ***
codex13 -2.61108  0.07346  0.09842 -26.529 < 2e-16 ***
x21      0.79896  2.22323  0.14887  5.367 8.01e-08 ***
x22      0.83596  2.30703  0.14119  5.921 3.20e-09 ***
x23      0.83484  2.30444  0.14349  5.818 5.96e-09 ***
x24      0.81607  2.26161  0.14575  5.599 2.16e-08 ***
x25      0.61956  1.85811  0.15488  4.000 6.33e-05 ***
x32      0.03486  1.03547  0.04342  0.803  0.422
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

      exp(coef) exp(-coef) lower .95 upper .95
codex11  0.53181    1.8804  0.45257  0.62493
codex12  0.18102    5.5242  0.15220  0.21531
codex13  0.07346   13.6137  0.06057  0.08908
x21      2.22323    0.4498  1.66060  2.97648
x22      2.30703    0.4335  1.74934  3.04251
x23      2.30444    0.4339  1.73949  3.05286
x24      2.26161    0.4422  1.69962  3.00942
x25      1.85811    0.5382  1.37163  2.51714
x32      1.03547    0.9657  0.95099  1.12746

Concordance= 0.813 (se = 0.005 )
Rsquare= 0.41 (max possible= 1 )
Likelihood ratio test= 1456 on 9 df, p=<2e-16
Wald test = 1458 on 9 df, p=<2e-16
Score (logrank) test = 1692 on 9 df, p=<2e-16

> #Uji GOF#
> gof.mod2<-cox.zph(mod2,transform="rank")
> gof.mod2
      rho chisq      p
codex11 0.2283 138.89 4.66e-32
codex12 0.3880 356.20 1.89e-79
codex13 0.4015 359.59 3.46e-80
x21      0.0335  2.87 9.01e-02
x22      0.0700 13.37 2.56e-04
x23      0.0830 18.65 1.57e-05
x24      0.1081 31.66 1.83e-08
x25      0.1081 31.14 2.40e-08
x32      0.0289  2.16 1.41e-01
GLOBAL      NA 956.67 3.79e-200

```

Lampiran 17. Output R Regresi Cox Extended Fungsi Waktu Heavyside (Lanjutan)

```

> summary(modske2)
Call:
coxph(formula = Y2 ~ as.factor(x3) + as.factor(x1hv1) + as.factor(x2hv1) +
      as.factor(x1hv2) + as.factor(x2hv2) + cluster(id), data = DatakuneW3,
      method = "breslow")

n= 5789, number of events= 2519

              coef      exp(coef)    se(coef)    robust se         z Pr(>|z|)
as.factor(x3)2      0.0950691    1.0997349    0.0434634    0.0451384      2.106  0.035190 *
as.factor(x1hv1)1   -1.6359879    0.1947599    0.0930734    0.0998546    -16.384 < 2e-16 ***
as.factor(x1hv1)2   -5.8983917    0.0027439    0.5069357    0.5080127    -11.611 < 2e-16 ***
as.factor(x1hv1)3   -6.4639257    0.0015587    1.0036427    1.0080878    -6.412  1.44e-10 ***
as.factor(x2hv1)1    0.3748141    1.4547210    0.1583102    0.1886994     1.986  0.047000 *
as.factor(x2hv1)2    0.1637105    1.1778733    0.1476152    0.1747617     0.937  0.348880
as.factor(x2hv1)3    0.1425015    1.1531548    0.1500890    0.1756503     0.811  0.417205
as.factor(x2hv1)4   -0.1072569    0.8982949    0.1543045    0.1776007    -0.604  0.545896
as.factor(x2hv1)5   -0.8370600    0.4329816    0.1891606    0.2031465    -4.120  3.78e-05 ***
as.factor(x1hv2)1    0.8893035    2.4334343    0.1954643    0.2469941     3.601  0.000318 ***
as.factor(x1hv2)2    0.0009843    1.0009847    0.1966561    0.2435679     0.004  0.996776
as.factor(x1hv2)3   -1.1655980    0.3117362    0.2028068    0.2489645    -4.682  2.84e-06 ***
as.factor(x2hv2)1    0.7738500    2.1680975    0.4221175    0.4873818     1.588  0.112338
as.factor(x2hv2)2    1.2408051    3.4583968    0.3699826    0.4189138     2.962  0.003057 **
as.factor(x2hv2)3    1.4968437    4.4675656    0.3701278    0.4194346     3.569  0.000359 ***
as.factor(x2hv2)4    2.1067010    8.2210750    0.3656326    0.4056525     5.193  2.07e-07 ***
as.factor(x2hv2)5    2.0637647    7.8755634    0.3685928    0.4038220     5.111  3.21e-07 ***

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
as.factor(x3)2      1.099735    0.9093  1.0066214  1.201461
as.factor(x1hv1)1    0.194760    5.1345  0.1601412  0.236862
as.factor(x1hv1)2    0.002744    364.4508  0.0010138  0.007426
as.factor(x1hv1)3    0.001559    641.5747  0.0002161  0.011242
as.factor(x2hv1)1    1.454721    0.6874  1.0049825  2.105722
as.factor(x2hv1)2    1.177873    0.8490  0.8362597  1.659037
as.factor(x2hv1)3    1.153155    0.8672  0.8172855  1.627052
as.factor(x2hv1)4    0.898295    1.1132  0.6342273  1.272310
as.factor(x2hv1)5    0.432982    2.3096  0.2907708  0.644745
as.factor(x1hv2)1    2.433434    0.4109  1.4996084  3.948766
as.factor(x1hv2)2    1.000985    0.9990  0.6210150  1.613440
as.factor(x1hv2)3    0.311736    3.2078  0.1913675  0.507816
as.factor(x2hv2)1    2.168097    0.4612  0.8341011  5.635584
as.factor(x2hv2)2    3.458397    0.2892  1.5215799  7.860585
as.factor(x2hv2)3    4.467566    0.2238  1.9635753  10.164694
as.factor(x2hv2)4    8.221075    0.1216  3.7122443  18.206257
as.factor(x2hv2)5    7.875563    0.1270  3.5690095  17.378631

Concordance= 0.83 (se = 0.004 )
Rsquare= 0.322 (max possible= 0.998 )
Likelihood ratio test= 2253 on 17 df,  p<<2e-16
Wald test = 1337 on 17 df,  p<<2e-16
Score (logrank) test = 2734 on 17 df,  p<<2e-16, Robust = 1573 p<<2e-16

(Note: the likelihood ratio and score tests assume independence of
observations within a cluster, the wald and robust score tests do not).
> extractAIC(modske2)
[1] 17.00 33660.05
> AICc(modske2, return.K = FALSE, second.ord = TRUE, nobs = NULL)
[1] 33660.16

```

Lampiran 18. Output R Regresi Cox Extended Fungsi Waktu $\ln(t)$ dan *Heavyside*

```

> summary(modske3) #model lengkap
Call:
coxph(formula = Y3 ~ as.factor(X3) + X1 + as.factor(x2hv1) +
      logtX1 + as.factor(X2hv2) + cluster(id), data = Datakunew4,
      method = "breslow")

n= 36765, number of events= 2519

              coef exp(coef) se(coef) robust se          z Pr(>|z|)
as.factor(X3)2 -0.02910  0.97132  0.04380  0.05345  -0.544 0.586183
X1              -2.52225  0.08028  0.08552  0.13832 -18.235 < 2e-16 ***
as.factor(x2hv1)1 0.49015  1.63257  0.15794  0.31167   1.573 0.115796
as.factor(x2hv1)2 0.72246  2.05949  0.14787  0.27398   2.637 0.008368 **
as.factor(x2hv1)3 0.76270  2.14406  0.14999  0.28509   2.675 0.007465 **
as.factor(x2hv1)4 0.76185  2.14223  0.15642  0.28176   2.704 0.006853 **
as.factor(x2hv1)5 0.15615  1.16900  0.19152  0.30191   0.517 0.605003
logtX1          0.73273  2.08075  0.02691  0.04225  17.344 < 2e-16 ***
as.factor(x2hv2)1 0.72941  2.07386  0.42176  0.45729   1.595 0.110697
as.factor(x2hv2)2 1.45250  4.27379  0.36741  0.37409   3.883 0.000103 ***
as.factor(x2hv2)3 1.64359  5.17370  0.36740  0.37410   4.393 1.12e-05 ***
as.factor(x2hv2)4 1.97747  7.22442  0.36304  0.36894   5.360 8.33e-08 ***
as.factor(x2hv2)5 1.92002  6.82107  0.36578  0.37068   5.180 2.22e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
as.factor(X3)2  0.97132    1.0295  0.87472    1.0786
X1              0.08028   12.4566  0.06122    0.1053
as.factor(x2hv1)1 1.63257    0.6125  0.88629    3.0072
as.factor(x2hv1)2 2.05949    0.4856  1.20377    3.5235
as.factor(x2hv1)3 2.14406    0.4664  1.22623    3.7489
as.factor(x2hv1)4 2.14223    0.4668  1.23320    3.7213
as.factor(x2hv1)5 1.16900    0.8554  0.64689    2.1125
logtX1          2.08075    0.4806  1.91540    2.2604
as.factor(x2hv2)1 2.07386    0.4822  0.84632    5.0819
as.factor(x2hv2)2 4.27379    0.2340  2.05301    8.8968
as.factor(x2hv2)3 5.17370    0.1933  2.48522   10.7706
as.factor(x2hv2)4 7.22442    0.1384  3.50556   14.8884
as.factor(x2hv2)5 6.82107    0.1466  3.29860   14.1051

Concordance= 0.886 (se = 0.004 )
Rsquare= 0.074 (max possible= 0.623 )
Likelihood ratio test= 2810 on 13 df, p=<2e-16
Wald test              = 843.4 on 13 df, p=<2e-16
Score (logrank) test = 2017 on 13 df, p=<2e-16, Robust = 1214 p=<2e-16

(Note: the likelihood ratio and score tests assume independence of
      observations within a cluster, the wald and robust score tests do not).

```

Lampiran 18. Output R Regresi Cox Extended Fungsi Waktu $\text{Ln}(t)$ dan *Heavyside* (Lanjutan)

```

> summary(modske3_1) #tanpa x3
Call:
coxph(formula = Y3 ~ X1 + as.factor(x2hv1) + logtX1 + as.factor(x2hv2) +
      cluster(id), data = Dataknew4, method = "breslow")

n= 36765, number of events= 2519

              coef exp(coef) se(coef) robust se          z Pr(>|z|)
X1          -2.51828   0.08060  0.08525   0.13961 -18.038 < 2e-16 ***
as.factor(x2hv1)1  0.48998   1.63228  0.15794   0.31164   1.572  0.11590
as.factor(x2hv1)2  0.72303   2.06067  0.14787   0.27362   2.642  0.00823 **
as.factor(x2hv1)3  0.76627   2.15173  0.14989   0.28439   2.694  0.00705 **
as.factor(x2hv1)4  0.76996   2.15969  0.15592   0.27998   2.750  0.00596 **
as.factor(x2hv1)5  0.16545   1.17992  0.19098   0.29997   0.552  0.58126
logtX1         0.73160   2.07841  0.02683   0.04262  17.166 < 2e-16 ***
as.factor(x2hv2)1  0.72768   2.07027  0.42174   0.45725   1.591  0.11152
as.factor(x2hv2)2  1.45259   4.27416  0.36742   0.37315   3.893  9.91e-05 ***
as.factor(x2hv2)3  1.65134   5.21394  0.36722   0.37224   4.436  9.16e-06 ***
as.factor(x2hv2)4  1.98906   7.30869  0.36263   0.36606   5.434  5.52e-08 ***
as.factor(x2hv2)5  1.93376   6.91545  0.36520   0.36741   5.263  1.42e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
X1          0.0806    12.4072    0.0613    0.106
as.factor(x2hv1)1  1.6323    0.6126    0.8862    3.007
as.factor(x2hv1)2  2.0607    0.4853    1.2053    3.523
as.factor(x2hv1)3  2.1517    0.4647    1.2323    3.757
as.factor(x2hv1)4  2.1597    0.4630    1.2476    3.739
as.factor(x2hv1)5  1.1799    0.8475    0.6554    2.124
logtX1        2.0784    0.4811    1.9118    2.259
as.factor(x2hv2)1  2.0703    0.4830    0.8449    5.073
as.factor(x2hv2)2  4.2742    0.2340    2.0570    8.881
as.factor(x2hv2)3  5.2139    0.1918    2.5137   10.815
as.factor(x2hv2)4  7.3087    0.1368    3.5666   14.977
as.factor(x2hv2)5  6.9155    0.1446    3.3658   14.209

Concordance= 0.886 (se = 0.004 )
Rsquare= 0.074 (max possible= 0.623 )
Likelihood ratio test= 2810 on 12 df,  p=<2e-16
Wald test           = 823 on 12 df,    p=<2e-16
Score (logrank) test = 2017 on 12 df,  p=<2e-16,  Robust = 1214 p=<2e-16

(Note: the likelihood ratio and score tests assume independence of
      observations within a cluster, the Wald and robust score tests do not).

> extractAIC(modske3) #model lengkap
[1] 13.00 33095.14
> extractAIC(modske3_1) #tanpa x3
[1] 12.00 33093.58
> AICc(modske3, return.K = FALSE, second.ord = TRUE, nobs = NULL)
[1] 33095.15
> AICc(modske3_1, return.K = FALSE, second.ord = TRUE, nobs = NULL)
[1] 33093.59

```

Lampiran 19. Output R Regresi Cox Extended Fungsi Waktu *T* dan *Heavyside*

```

> summary(modske4) #model lengkap
Call:
coxph(formula = Y4 ~ as.factor(X3) + X1 + as.factor(X2hv1) +
      tx1 + as.factor(X2hv2) + cluster(id), data = Dataknew5,
      method = "breslow")

n= 36765, number of events= 2519

              coef exp(coef) se(coef) robust se      z Pr(>|z|)
as.factor(X3)2 -0.033865  0.966702  0.043777  0.053528  -0.633  0.5270
X1              -0.897631  0.407534  0.028047  0.047963 -18.715 < 2e-16 ***
as.factor(X2hv1)1  0.498342  1.645991  0.157972  0.315072   1.582  0.1137
as.factor(X2hv1)2  0.699636  2.013019  0.147895  0.278222   2.515  0.0119 *
as.factor(X2hv1)3  0.727382  2.069655  0.149952  0.287101   2.534  0.0113 *
as.factor(X2hv1)4  0.717240  2.048771  0.156290  0.284551   2.521  0.0117 *
as.factor(X2hv1)5  0.124545  1.132633  0.191484  0.304248   0.409  0.6823
tx1              0.028667  1.029082  0.001132  0.001817  15.776 < 2e-16 ***
as.factor(X2hv2)1  0.774095  2.168629  0.421751  0.475967   1.626  0.1039
as.factor(X2hv2)2  1.512740  4.539152  0.367432  0.385636   3.923  8.76e-05 ***
as.factor(X2hv2)3  1.722604  5.599089  0.367415  0.385778   4.465  8.00e-06 ***
as.factor(X2hv2)4  2.081826  8.019102  0.363186  0.380501   5.471  4.47e-08 ***
as.factor(X2hv2)5  2.045563  7.733509  0.365935  0.382118   5.353  8.64e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
as.factor(X3)2      0.9667      1.0344      0.8704      1.0736
X1                  0.4075      2.4538      0.3710      0.4477
as.factor(X2hv1)1  1.6460      0.6075      0.8876      3.0522
as.factor(X2hv1)2  2.0130      0.4968      1.1669      3.4727
as.factor(X2hv1)3  2.0697      0.4832      1.1790      3.6331
as.factor(X2hv1)4  2.0488      0.4881      1.1730      3.5785
as.factor(X2hv1)5  1.1326      0.8829      0.6239      2.0562
tx1                 1.0291      0.9717      1.0254      1.0328
as.factor(X2hv2)1  2.1686      0.4611      0.8532      5.5122
as.factor(X2hv2)2  4.5392      0.2203      2.1317      9.6656
as.factor(X2hv2)3  5.5991      0.1786      2.6287     11.9259
as.factor(X2hv2)4  8.0191      0.1247      3.8040     16.9047
as.factor(X2hv2)5  7.7335      0.1293      3.6569     16.3545

Concordance= 0.887 (se = 0.004 )
Rsquare= 0.071 (max possible= 0.623 )
Likelihood ratio test= 2727 on 13 df, p=<2e-16
Wald test              = 847.6 on 13 df, p=<2e-16
Score (logrank) test = 2014 on 13 df, p=<2e-16, Robust = 1265 p=<2e-16

(Note: the likelihood ratio and score tests assume independence of
      observations within a cluster, the wald and robust score tests do not).

```

Lampiran 19. Output R Regresi Cox Extended Fungsi Waktu T dan $Heavyside$ (Lanjutan)

```

> summary(modske4_1) #tanpa x3
Call:
coxph(formula = Y4 ~ X1 + as.factor(x2hv1) + tx1 + as.factor(x2hv2) +
      cluster(id), data = Dataknew5, method = "breslow")

n = 36765, number of events = 2519

              coef exp(coef) se(coef) robust se      z Pr(>|z|)
X1          -0.895644  0.408345  0.027904  0.048472 -18.477 < 2e-16 ***
as.factor(x2hv1)1  0.497879  1.645227  0.157967  0.315060  1.580  0.1140
as.factor(x2hv1)2  0.700180  2.014114  0.147890  0.277816  2.520  0.0117 *
as.factor(x2hv1)3  0.731412  2.078013  0.149858  0.286359  2.554  0.0106 *
as.factor(x2hv1)4  0.726483  2.067796  0.155815  0.282788  2.569  0.0102 *
as.factor(x2hv1)5  0.135105  1.144657  0.190969  0.302358  0.447  0.6550
tx1           0.028604  1.029017  0.001128  0.001834 15.598 < 2e-16 ***
as.factor(x2hv2)1  0.771310  2.162598  0.421726  0.476069  1.620  0.1052
as.factor(x2hv2)2  1.512678  4.538872  0.367434  0.384563  3.934  8.37e-05 ***
as.factor(x2hv2)3  1.731448  5.648827  0.367242  0.383742  4.512  6.42e-06 ***
as.factor(x2hv2)4  2.095252  8.127492  0.362787  0.377410  5.552  2.83e-08 ***
as.factor(x2hv2)5  2.061438  7.857259  0.365369  0.378664  5.444  5.21e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
X1           0.4083    2.4489    0.3713    0.449
as.factor(x2hv1)1  1.6452    0.6078    0.8873    3.051
as.factor(x2hv1)2  2.0141    0.4965    1.1684    3.472
as.factor(x2hv1)3  2.0780    0.4812    1.1855    3.642
as.factor(x2hv1)4  2.0678    0.4836    1.1880    3.599
as.factor(x2hv1)5  1.1447    0.8736    0.6329    2.070
tx1           1.0290    0.9718    1.0253    1.033
as.factor(x2hv2)1  2.1626    0.4624    0.8506    5.498
as.factor(x2hv2)2  4.5389    0.2203    2.1360    9.645
as.factor(x2hv2)3  5.6488    0.1770    2.6627   11.984
as.factor(x2hv2)4  8.1275    0.1230    3.8789   17.030
as.factor(x2hv2)5  7.8573    0.1273    3.7407   16.504

Concordance= 0.886 (se = 0.004 )
Rsquare= 0.071 (max possible= 0.623 )
Likelihood ratio test= 2727 on 12 df,  p=<2e-16
Wald test           = 828.2 on 12 df,    p=<2e-16
Score (logrank) test = 2014 on 12 df,    p=<2e-16,  Robust = 1265 p=<2e-16

(Note: the likelihood ratio and score tests assume independence of
      observations within a cluster, the Wald and robust score tests do not).

> extractAIC(modske4) #model lengkap
[1] 13.00 33178.21
> extractAIC(modske4_1) #tanpa x3
[1] 12.00 33176.81
> AICc(modske4, return.K = FALSE, second.ord = TRUE, nobs = NULL)
[1] 33178.22
> AICc(modske4_1, return.K = FALSE, second.ord = TRUE, nobs = NULL)
[1] 33176.82

```

Lampiran 20. Surat Pernyataan Data



Nov 07, 2019

Fitria Nur Aida
Institut Teknologi Sepuluh Nopember
Indonesia
Phone: 083830893310
Email: fitrianuraida0102@gmail.com
Request Date: 11/07/2019

Dear Fitria Nur Aida:

This is to confirm that you are approved to use the following Survey Datasets for your registered research paper titled: "Determinant of timing of first childbirth among women of childbearing age in East Java Province":

Indonesia

To access the datasets, please login at: https://www.dhsprogram.com/data/dataset_admin/login_main.cfm. The user name is the registered email address, and the password is the one selected during registration.

The IRB-approved procedures for DHS public-use datasets do not in any way allow respondents, households, or sample communities to be identified. There are no names of individuals or household addresses in the data files. The geographic identifiers only go down to the regional level (where regions are typically very large geographical areas encompassing several states/provinces). Each enumeration area (Primary Sampling Unit) has a PSU number in the data file, but the PSU numbers do not have any labels to indicate their names or locations. In surveys that collect GIS coordinates in the field, the coordinates are only for the enumeration area (EA) as a whole, and not for individual households, and the measured coordinates are randomly displaced within a large geographic area so that specific enumeration areas cannot be identified.

The DHS Data may be used only for the purpose of statistical reporting and analysis, and only for your registered research. To use the data for another purpose, a new research project must be registered. All DHS data should be treated as confidential, and no effort should be made to identify any household or individual respondent interviewed in the survey. Please reference the complete terms of use at: <https://dhsprogram.com/Data/terms-of-use.cfm>.

The data must not be passed on to other researchers without the written consent of DHS. However, if you have coresearchers registered in your account for this research paper, you are authorized to share the data with them. All data users are required to submit an electronic copy (pdf) of any reports/publications resulting from using the DHS data files to: references@dhsprogram.com.

Sincerely,

Bridgette Wellington

Bridgette Wellington
Data Archivist
The Demographic and Health Surveys (DHS) Program

(Halaman ini sengaja dikosongkan)

BIODATA PENULIS



Fitria Nur Aida merupakan anak pertama dari pasangan Mansur dan Siti Asiyah. Penulis lahir di Surabaya pada tanggal 1 Februari 1998. Pendidikan formal penulis ditempuh di SDN Rankah VII Surabaya, SMPN 1 Surabaya, dan SMAN 5 Surabaya. Selanjutnya, penulis melanjutkan pendidikan formal di Program Studi Sarjana Departemen Statistika ITS. Penulis diterima di ITS pada tahun 2016 melalui jalur SBMPTN. Selama perkuliahan, penulis aktif dalam organisasi dan kepanitiaan baik di dalam kampus maupun di luar kampus. Penulis aktif di Badan Semiotonom Himpunan Mahasiswa Departemen Statistika ITS (HIMASTA-ITS) yaitu Divisi Professional Statistics (PSt) sebagai Staff pada periode 2017-2019 dan Manajer Tim Riset pada periode 2018-2019. Penulis juga pernah menjabat sebagai Sekretaris UKM KSR PMI ITS periode 2017-2018. Selain aktif dalam organisasi dan kepanitiaan, penulis juga mengikuti berbagai perlombaan baik di bidang karya tulis ilmiah maupun analisis data. Capaian prestasi penulis antara lain Juara 2 Lomba Karya Tulis Ilmiah National Data Day 2018 dan Juara 1 Lomba Pekan Analisis Statistika JAMBORE VIII. Apabila pembaca ingin memberikan kritik dan saran serta diskusi lebih lanjut mengenai Tugas Akhir ini, dapat menghubungi penulis melalui email fitrianuraida0102@gmail.com atau nomor telepon 083830893310.