

**ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI  
LAMA WAKTU MENCARI PEKERJAAN MENGGUNAKAN  
KURVA KAPLAN-MEIER, UJI *LOG-RANK*, DAN REGRESI  
*COX PROPORTIONAL HAZARD***

Studi Kasus: Data Sakernas Provinsi D.I. Yogyakarta Agustus 2022

**TUGAS AKHIR**



Bimasuci Basiludin

19611170

**PROGRAM STUDI STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS ISLAM INDONESIA  
YOGYAKARTA**

**2023**

**HALAMAN PERSETUJUAN PEMBIMBING  
TUGAS AKHIR**

Judul : Analisis Faktor-faktor yang Mempengaruhi Lama Waktu Mencari Pekerjaan Menggunakan Kurva Kaplan-Meier, Uji *Log-Rank*, dan Regresi *Cox Proportional Hazard* (Studi Kasus: Data Sakernas Provinsi D.I. Yogyakarta Agustus 2022)

Nama Mahasiswa : Bimasuci Basiludin

NIM : 19611170

**TUGAS AKHIR INI TELAH DIPERIKSA DAN DISETUJUI UNTUK  
DIUJIKAN**

Yogyakarta, 12 September 2023

Pembimbing 1



(Abdullah Ahmad Dzikrullah, S.Si., M.Sc.)

Pembimbing 2



(Muhammad Hasan Sidiq Kurniawan, S.Si., M.Sc.)

Ketua Prodi



(Dr. Atina Andika, S.Si., M.Si.)

**HALAMAN PENGESAHAN**  
**TUGAS AKHIR**

**ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI LAMA  
WAKTU MENCARI PEKERJAAN MENGGUNAKAN KURVA KAPLAN-  
MEIER, UJI LOG-RANK, DAN REGRESI COX *PROPORTIONAL***

***HAZARD***

**(Studi Kasus: Data Sakernas Provinsi D.I. Yogyakarta Agustus 2022)**

**Nama Mahasiswa : Bimasuci Basiludin**

**NIM : 19611170**

**TUGAS AKHIR INI TELAH DIUJIKAN  
PADA TANGGAL: 4 Oktober 2023**

**Nama Penguji:**

**Tanda Tangan**

1. Arum Handini Primandari, S.Pd.Si., M.Si. ....

2. Ayundyah Kesumawati, S.Si., M.Si. ....

3. Abdullah Ahmad Dzikrullah, S.Si., M.Sc. ....

Mengetahui,

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



(Prof. Riyanto, S.Pd., M.Si., Ph.D.)



## KATA PENGANTAR

*Assalamu'alaikum Wr.Wb*

Puji dan syukur penulis panjatkan kepada Allah SWT yang atas berkat rahmat dan karunianya penulis dapat menyelesaikan penyusunan tugas akhir ini yang berjudul: “Analisis Faktor-faktor yang Mempengaruhi Lama Waktu Mencari Pekerjaan Menggunakan Kurva Kaplan-Meier, Uji *Log-Rank*, dan Regresi Cox *Proportional Hazard* (Studi Kasus: Data Sakernas Provinsi D.I. Yogyakarta Agustus 2022)”.

Disusunnya tugas akhir ini bertujuan untuk memenuhi salah satu syarat untuk mendapatkan gelar Sarjana Statistika di Program Studi Statistika Universitas Islam Indonesia Yogyakarta, sekaligus mewujudkan rasa cinta penulis kepada Daerah Istimewa Yogyakarta untuk setidaknya berkontribusi dalam memberikan wawasan terkait ketenagakerjaan.

Penulis ingin menyampaikan ucapan terima kasih dan apresiasi yang setinggi-tingginya kepada pihak-pihak yang telah berkontribusi, memberikan dukungan, dan bantuan kepada penulis dengan tulus dan tanpa pamrih sehingga dapat menyelesaikan serangkaian perkuliahan sejak awal hingga penyelesaian tugas akhir ini. Pihak-pihak tersebut di antaranya adalah:

1. Bapak Abdullah Ahmad Dzirkullah, S.Si., M.Sc. sebagai dosen pembimbing utama yang senantiasa tulus dan ikhlas dalam membimbing penulis untuk menyelesaikan tugas akhir ini.
2. Bapak Muhammad Hasan Sidiq Kurniawan, S.Si., M.Sc. sebagai dosen pembimbing kedua yang juga telah berkontribusi besar dalam penulisan tugas akhir ini karena selalu sabar dan tulus dalam membimbing dan menghadapi pertanyaan dari penulis yang seringkali kebingungan.
3. Bapak Dayatmoko selaku ayah kandung dari penulis yang telah memfasilitasi, membiayai, dan mendukung penulis sejak awal perkuliahan hingga terselesaikannya penulisan tugas akhir ini.



4. Mbak Nurul Indra yang telah membantu penyusunan tugas akhir ini sejak awal, memberikan ide penulisan, membantu penyusunan kata, mencarikan sumber dan referensi, memberikan motivasi, memberikan dukungan, bahkan memberikan makan minum dan tempat menginap di rumah.
5. Teman-teman Statistika kelas D angkatan 2019, Salsa, Dhan, dkk. yang telah senantiasa memberikan dukungan, bantuan dan motivasi kepada penulis.
6. Teruntuk yang istimewa, Rif'atur Ayunda Fitri yang telah menjadi sumber semangat dan motivasi utama bagi penulis dan Insya Allah akan menjadi calon istri bagi penulis di masa yang akan datang.
7. Pihak-pihak lain yang tidak dapat disebutkan secara satu persatu, yang telah memberikan segala bantuan dan dukungan kepada penulis baik secara langsung maupun tidak langsung.

Akhir kata, penulis memohonkan kepada Allah agar semoga segala bantuan dan perhatian sekecil apapun dari berbagai pihak kepada penulis agar diberikan imbalan dan balasan yang setinggi-tingginya, baik di dunia maupun di akhirat. Penulis menyadari bahwa penulisan tugas akhir ini memiliki banyak kekurangan dan kelemahan, oleh karena itu penulis terbuka terhadap segala kritik dan saran untuk dapat menjadi lebih baik terutama dalam membuat tulisan-tulisan yang akan datang. Penulis memiliki harapan besar untuk bergunanya tulisan ini baik kepada penulis sendiri maupun kepada para pembaca.

*Wassalamualaikum Wr.Wb*

Yogyakarta, September 2023

Bimasuci Basiludin

## DAFTAR ISI

HALAMAN SAMPUL .....	i
HALAMAN PERSETUJUAN PEMBIMBING .....	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR .....	iv
DAFTAR ISI.....	vi
DAFTAR TABEL.....	ix
DAFTAR GAMBAR .....	x
DAFTAR LAMPIRAN .....	xi
PERNYATAAN.....	xii
ABSTRAK .....	xiii
ABSTRACT.....	xiv
BAB 1 PENDAHULUAN .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	4
1.3 Batasan Masalah .....	4
1.4 Tujuan Penelitian .....	5
1.5 Manfaat Penelitian.....	5
BAB 2 TINJAUAN PUSTAKA .....	6
2.1 Penelitian Terdahulu.....	6
BAB 3 LANDASAN TEORI .....	9
3.1 Teori Ketenagakerjaan.....	9
3.1.1 Angkatan Kerja.....	9
3.2 Faktor-faktor yang Dapat Mempengaruhi Lama Waktu Mencari Kerja	
10	
3.3 Statistika Deskriptif .....	14
3.3.1 Diagram Batang .....	14
3.4 Analisis Survival .....	15

3.4.1	Data Survival .....	16
3.4.2	Teori Data Tersensor .....	17
3.4.3	Fungsi Survival .....	19
3.4.4	Fungsi Hazard dan Hazard Kumulatif .....	20
3.5	Estimasi dan Kurva Kaplan-Meier .....	21
3.6	Uji <i>Log-Rank</i> .....	21
3.7	Regresi .....	23
3.7.1	Regresi Cox Proportional Hazard .....	23
3.7.2	Estimasi Parameter .....	25
3.7.3	Ties dalam Partial Likelihood .....	26
3.7.4	Uji Signifikansi Parameter .....	26
3.7.5	Asumsi Proportional Hazard .....	28
BAB 4	METODOLOGI PENELITIAN .....	31
4.1	Populasi dan Sampel .....	31
4.2	Data dan Sumber Data .....	31
4.3	Variabel dan Definisi Operasional .....	31
4.4	Metode Analisis Data .....	34
4.5	Alur Penelitian .....	34
BAB 5	HASIL DAN PEMBAHASAN .....	36
5.1	Pengumpulan Data .....	36
5.2	Identifikasi Data .....	36
5.3	Kurva Kaplan-Meier dan Uji <i>Log-Rank</i> .....	42
5.3.1	Membuat Kurva Kaplan-Meier dan Melakukan Uji <i>Log-Rank</i> .....	42
5.3.2	Interpretasi Pengaruh Variabel Bebas Berdasarkan Kurva Kaplan-Meier dan Uji <i>Log-Rank</i> .....	60
5.4	Regresi Cox <i>Proportional Hazard</i> .....	61
5.4.1	Uji Asumsi Proportional Hazard .....	61
5.4.2	Estimasi Parameter .....	62
5.4.3	Uji Signifikansi Parameter .....	63
5.4.4	Pembentukan Model Alternatif .....	65
5.4.5	Interpretasi Parameter Regresi dan Pengaruh Variabel Bebas .....	66

5.5	Signifikansi Pengaruh Variabel Bebas Berdasarkan Hasil Uji <i>Log-Rank</i> dan Regresi Cox <i>Proportional Hazard</i> .....	68
BAB 6	PENUTUP .....	69
6.1	Kesimpulan .....	69
6.2	Saran .....	70
DAFTAR PUSTAKA .....		71
LAMPIRAN .....		75

## DAFTAR TABEL

<b>Tabel 3.1</b> Contoh Kasus Data <i>Ties</i> .....	26
<b>Tabel 4.1</b> Definisi operasional variabel.....	31
<b>Tabel 5.1</b> Jumlah data berdasarkan status penyensoran .....	37
<b>Tabel 5.2</b> Interpretasi Pengaruh Variabel Bebas Berdasarkan Uji <i>Log-Rank</i> .....	60
<b>Tabel 5.3</b> Hasil uji asumsi <i>proportional hazard</i> .....	62
<b>Tabel 5.4</b> Estimasi parameter model awal.....	62
<b>Tabel 5.5</b> Hasil uji parsial.....	65
<b>Tabel 5.6</b> Hasil estimasi model alternatif .....	66
<b>Tabel 5.7</b> Interpretasi Parameter Variabel X5 .....	67
<b>Tabel 5.8</b> Hasil identifikasi pengaruh variabel bebas.....	68



## DAFTAR GAMBAR

<b>Gambar 3.1</b> Ilustrasi data survival.....	17
<b>Gambar 3.2</b> Ilustrasi penyensoran data .....	18
<b>Gambar 3.3</b> Contoh bentuk fungsi survival.....	20
<b>Gambar 4.1</b> Diagram Alir Penelitian.....	35
<b>Gambar 5.1</b> Jumlah data berdasarkan jenis kelamin .....	38
<b>Gambar 5.2</b> Jumlah data berdasarkan usia .....	38
<b>Gambar 5.3</b> Jumlah data berdasarkan hubungan dengan KRT .....	39
<b>Gambar 5.4</b> Jumlah data berdasarkan status perkawinan.....	39
<b>Gambar 5.5</b> Jumlah data berdasarkan tingkat pendidikan.....	40
<b>Gambar 5.6</b> Jumlah data berdasarkan pengalaman kerja .....	40
<b>Gambar 5.7</b> Jumlah data berdasarkan pelatihan kerja .....	41
<b>Gambar 5.8</b> Jumlah data berdasarkan tempat tinggal.....	41
<b>Gambar 5.9</b> Kurva Kaplan-Meier variabel X1 .....	43
<b>Gambar 5.10</b> Kurva Kaplan-Meier variabel X2 .....	45
<b>Gambar 5.11</b> Kurva Kaplan-Meier variabel X3 .....	47
<b>Gambar 5.12</b> Kurva Kaplan-Meier variabel X4 .....	49
<b>Gambar 5.13</b> Kurva Kaplan-Meier variabel X5 .....	51
<b>Gambar 5.14</b> Kurva Kaplan-Meier variabel X6 .....	53
<b>Gambar 5.15</b> Kurva Kaplan-Meier variabel X7 .....	55
<b>Gambar 5.16</b> Kurva Kaplan-Meier variabel X8 .....	57
<b>Gambar 5.17</b> Kurva kaplan-meier variabel gabungan X1 dan X4.....	59

## DAFTAR LAMPIRAN

<b>Lampiran 1</b> Surat Perjanjian Penggunaan Data.....	75
<b>Lampiran 2</b> Sintaks Regresi Cox PH.....	76
<b>Lampiran 3</b> Sintaks Kurva Kaplan-Meier & Uji <i>Log-Rank</i> .....	77
<b>Lampiran 4</b> Output Regresi Cox PH .....	80
<b>Lampiran 5</b> Output Uji <i>Log-Rank</i> .....	81
<b>Lampiran 6</b> Dalil Tentang Pernikahan .....	85

## PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang sebelumnya pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan orang lain, kecuali yang diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 12 September 2023

A 1000 Rupiah postage stamp is shown, featuring the Garuda Pancasila emblem and the text 'SEPULUH RIBU RUPIAH' and '1000'. Below the stamp, the text '20 METERA TEMPEL' and the serial number 'AF704AKX549281215' are visible. A handwritten signature in black ink is written over the stamp and extends to the right.

**Bimasuci Basiludin**

**ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI  
LAMA WAKTU MENCARI PEKERJAAN MENGGUNAKAN  
KURVA KAPLAN-MEIER, UJI *LOG-RANK*, DAN REGRESI  
*COX PROPORTIONAL HAZARD***

Bimasuci Basiludin

Program Studi Statistika, Fakultas MIPA

Universitas Islam Indonesia

**ABSTRAK**

Ketatnya persaingan dalam mencari pekerjaan akan mengakibatkan banyaknya faktor yang menjadi penentu terhadap lama waktu yang dibutuhkan bagi seseorang untuk mendapatkan pekerjaan. Hal ini dapat dianalisis dengan pendekatan menggunakan analisis survival, dengan beberapa di antara metode yang bisa digunakan adalah regresi *cox proportional hazard*, estimasi Kaplan-Meier, dan uji *log-rank*. Melalui pendekatan regresi *cox PH*, dilakukan analisis secara simultan untuk mengetahui faktor-faktor apa saja yang mempengaruhi lama waktu mencari kerja. Dengan metode tersebut dihasilkan bahwa jenis kelamin, tingkat pendidikan, pengalaman kerja, dan klasifikasi tempat tinggal menjadi faktor-faktor yang mempengaruhi lama waktu mencari kerja secara signifikan. Penerapan metode ini menghasilkan nilai *Concordance* sebesar 0.644 yang artinya model memiliki kemampuan yang cukup baik dalam membandingkan urutan waktu kejadian di antara pasangan observasi. Sementara itu untuk melakukan analisis secara parsial, dilakukan estimasi Kaplan-Meier yang dibentuk menjadi kurva dan uji *log-rank*, kemudian dihasilkan bahwa jenis kelamin, usia, hubungan dengan kepala rumah tangga, status perkawinan, tingkat pendidikan, pengalaman kerja, pelatihan kerja, dan klasifikasi tempat tinggal menjadi faktor-faktor yang mempengaruhi lama waktu mencari kerja secara signifikan.

**Kata Kunci:** Analisis Survival, Lama Waktu Mencari Kerja, Regresi Cox *Proportional Hazard*, Kaplan-Meier, Uji *Log-Rank*

**ANALYSIS OF FACTORS AFFECTING THE DURATION OF JOB  
SEARCHING SIMULTANEOUSLY AND PARTIALLY USING COX  
PROPORTIONAL HAZARD REGRESSION METHOD AND LOG-RANK  
TEST**

Bimasuci Basiludin  
Program Studi Statistika, Fakultas MIPA  
Universitas Islam Indonesia

**ABSTRACT**

*The tight competition in job seeking leads to numerous determining factors influencing the duration it takes for an individual to secure employment. This can be analyzed using a survival analysis approach, with several methods such as Cox proportional hazard regression, Kaplan-Meier estimation, and log-rank test being applicable. Through the Cox PH regression approach, a simultaneous analysis is conducted to identify the significant factors that affect the duration of job search. The application of this method reveals that gender, educational level, work experience, and residential classification are substantial factors significantly influencing the duration of job search. The implementation of this method yields a Concordance value of 0.644, signifying a reasonably strong ability of the model to compare the order of event times among pairs of observations. Meanwhile, for partial analysis, Kaplan-Meier estimation is performed, which is then transformed into curves, and a log-rank test is conducted. Consequently, gender, age, household relationship, marital status, educational level, work experience, job training, and residential classification are identified as significant factors affecting the duration of job search.*

**Keywords:** *Survival Analysis, Duration of Job Search, Cox Proportional Hazard Regression, Kaplan-Meier, Log-Rank Test.*



# **BAB 1 PENDAHULUAN**

## **1.1 Latar Belakang**

Daerah Istimewa Yogyakarta (DIY) termasuk dalam provinsi yang memiliki permasalahan terkait kependudukan, salah satunya berupa tingginya angka pengangguran. Hal tersebut menjadi tanda bahwa peluang untuk dapat bekerja di D.I. Yogyakarta tidak sebanding dengan banyaknya angkatan kerja yang tersedia. Menurut Badan Pusat Statistik (BPS) DIY (2022), banyaknya penduduk yang berusia kerja di DIY pada Agustus 2022 adalah 3,127 juta orang. Sebanyak 2,336 juta orang dari jumlah tersebut merupakan angkatan kerja, dan 881,80 ribu orang bukan angkatan kerja. Dari banyaknya jumlah angkatan kerja tersebut, DIY memiliki nilai Tingkat Pengangguran Terbuka (TPT) sebesar 4,06% secara keseluruhan (BPS Provinsi D.I. Yogyakarta, 2022).

Berdasarkan data dari BPS, tren angka TPT secara proporsi di DIY dari tahun 2020 hingga 2022 terus mengalami penurunan. Angka TPT tahun 2020 yaitu 4,57%, tahun 2021 sebesar 4,56% dan tahun 2022 sebesar 4,06% (Badan Pusat Statistik, 2022). Jumlah angkatan kerja di DIY pada tahun 2020 yaitu 2,228 juta orang. Angka ini naik di tahun 2021 menjadi 2,334 juta orang, kemudian meningkat kembali di tahun 2022 menjadi 2,336 juta orang. Sementara itu pada tahun 2020 jumlah pengangguran di DIY adalah sebanyak 101,85 ribu orang, meningkat pada tahun 2021 menjadi 106,43 ribu orang, lalu menurun pada tahun 2022 menjadi 94,95 ribu orang (BPS Provinsi D.I. Yogyakarta, 2022).

Besarnya perkembangan jumlah penduduk DIY berpengaruh pada jumlah angkatan kerja, dimana besarnya angkatan kerja tentu akan mencerminkan besarnya ketersediaan tenaga kerja. Namun besarnya ketersediaan tenaga kerja tersebut kurang disertai dengan peningkatan kebutuhan tenaga kerja. Sehingga hal ini berdampak pada banyaknya angkatan kerja yang tidak mendapatkan pekerjaan dan menjadi pengangguran (BPS Provinsi D.I. Yogyakarta, 2020).

Terbatasnya penawaran tenaga kerja di Yogyakarta akan membuat persaingan dalam mencari kerja menjadi ketat, sehingga akan banyak faktor-faktor yang menjadi penentu terhadap lama waktu yang dibutuhkan bagi seseorang untuk

mendapatkan pekerjaan di DIY. Menurut BPS DIY, banyaknya lapangan pekerjaan yang tersedia di DIY di tahun 2022 adalah sebanyak 2,241 juta orang yang mana didominasi oleh bidang Perdagangan Besar dan Eceran (19,84%), Pertanian, Kehutanan, Perikanan (17,94%), Industri Pengolahan (17,28%), serta Akomodasi dan Makan Minum (10,07%) (BPS Provinsi D.I. Yogyakarta, 2022).

Beberapa faktor yang dapat mempengaruhi lama waktu mencari pekerjaan diantaranya yaitu jenis kelamin, usia, dan tingkat pendidikan. Hal ini sesuai dengan penelitian yang dilakukan oleh Yosef Hartoko (2019), Fahrudin Arrozi dan Ketut Sutrisna (2018) yang menyatakan bahwa ketiga variabel bebas tersebut mempengaruhi lama waktu mencari kerja secara signifikan. Selain faktor-faktor tersebut, status perkawinan juga dapat mempengaruhi lama waktu mencari kerja. Hal ini berdasarkan hasil penelitian oleh Yosef Hartoko (2019) dan I Gede Ari Sudana, dkk (2013). Faktor lain yang juga dapat berpengaruh terhadap lama waktu mencari kerja adalah pengalaman kerja dan hubungan dengan kepala rumah tangga, yang mana sesuai dengan hasil penelitian yang dilakukan oleh Muji Nurjanah Susenati (2015). Selain itu lama waktu mencari kerja juga dapat dipengaruhi oleh pelatihan kerja, dimana hal ini sesuai dengan hasil penelitian oleh Yosef Hartoko (2019). Klasifikasi tempat tinggal juga dapat mempengaruhi lama waktu mencari pekerjaan, sesuai dengan hasil penelitian oleh I Gede Ari Sudana, dkk (2013).

Faktor-faktor yang mempengaruhi lama waktu mencari pekerjaan dapat dianalisis dengan menggunakan analisis survival. Analisis survival adalah prosedur statistik yang dilakukan untuk melakukan analisis terhadap data kelangsungan hidup (survival), yaitu data yang berhubungan dengan waktu yang memiliki waktu awal (*origin*) hingga terjadinya suatu kejadian khusus (*event*) (Kleinbaum & Klein, 2005).

Salah satu metode yang dapat digunakan dalam analisis survival adalah analisis regresi cox *proportional hazard*. Metode tersebut merupakan regresi semiparametrik yang tidak mengasumsikan persebaran data mengikut suatu distribusi tertentu. Model regresi cox *proportional hazard*, tidak membutuhkan fungsi *baseline hazard*, yaitu fungsi *hazard* yang melibatkan waktu tanpa memperhatikan kovariat (Riangkaryaman, 2017). Metode ini dapat digunakan

sebagai alternatif ketika data yang tersedia tidak memenuhi asumsi sebaran distribusi tertentu baik itu eksponensial, weibull, maupun log-logistik. Sehingga data yang ada tidak dapat dianalisis menggunakan jenis metode regresi parametrik. Metode ini dapat digunakan untuk melakukan analisis pengaruh variabel-variabel bebas secara bersama-sama (simultan).

Selain regresi *cox proportional hazard*, data survival juga dapat diteliti dengan menggunakan uji *log-rank* terhadap variabel bebas yang bersifat kategorik. Sehingga untuk variabel bebas yang bersifat numerik harus dikategorikan terlebih dahulu untuk dilakukan analisis menggunakan uji *log-rank*. Uji *log-rank* adalah sebuah metode statistik yang digunakan untuk membandingkan dua kelompok atau lebih dan menguji apakah ada perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok tersebut (Kleinbaum & Klein, 2005). Salah satu tujuan digunakannya uji *log-rank* dalam suatu penelitian yaitu untuk mengetahui apakah terdapat pengaruh dari variabel bebas terhadap variabel terikat secara parsial. Berbeda dengan regresi *cox proportional hazard*, uji *log-rank* hanya dapat dilakukan untuk menguji pengaruh setiap variabel bebas secara satu per satu (parsial). Selain itu, uji *log-rank* juga tidak dapat digunakan untuk membuat pemodelan fungsi survival maupun fungsi *hazard* seperti halnya regresi *cox ph*. Penggunaan uji *log-rank* biasanya juga dibarengi dengan penggunaan estimasi kaplan-meier untuk mengestimasi fungsi survival yang kemudian dibentuk menjadi kurva.

Dalam melakukan penelitian yang menggunakan analisis survival, seringkali terdapat sebuah masalah dalam data. Yaitu adanya dua individu atau lebih yang mengalami *event* pada waktu yang bersamaan, atau disebut sebagai kejadian bersama (*ties*). Kejadian ini dapat menimbulkan masalah pada saat dilakukan penghitungan *partial likelihood* dalam proses melakukan analisis regresi *cox proportional hazard*. Disebut sebagai *partial* karena hanya sebagian dari fungsi *likelihood* yang digunakan, yaitu pada data tersensor (Iskandar, Kusumawati, & Subekti, 2015). Dalam mengatasi masalah ini, ada 3 metode pendekatan yaitu *Efron*, *Breslow*, dan *Exact*. Di antara ketiga metode tersebut, metode *efron*

merupakan metode yang cepat, sederhana, dan juga akurat ketika data yang dianalisis mengandung banyak *ties* dalam jumlah besar (Rahmadeni & Ranti, 2016).

Berdasarkan penjelasan di atas, penulis mendapatkan ketertarikan untuk melakukan analisis survival dalam meneliti faktor-faktor yang mempengaruhi lama waktu mencari pekerjaan di Daerah Istimewa Yogyakarta menggunakan metode regresi *cox proportional hazard* dengan metode pendekatan *efron* untuk menganalisis pengaruh seluruh variabel bebas secara simultan, serta menggunakan metode uji *log-rank* untuk menganalisis pengaruh variabel bebas secara parsial. Tujuan digunakannya kedua metode tersebut adalah untuk membandingkan pengaruh variabel-variabel bebas yang dianalisis secara simultan dan secara parsial. Dalam melakukan analisis ataupun pemodelan secara simultan, seringkali didapatkan hasil yang kurang memuaskan karena adanya variabel bebas yang seharusnya dapat berpengaruh terhadap variabel terikat secara signifikan namun tertutupi pengaruhnya dikarenakan adanya variabel-variabel bebas lainnya yang jauh lebih signifikan. Oleh karena itu, penelitian ini dilakukan dengan membandingkan hasil analisis secara simultan dan secara parsial.

## **1.2 Rumusan Masalah**

Berdasarkan permasalahan yang telah dipaparkan dalam latar belakang, rumusan masalah yang dapat disusun adalah sebagai berikut:

1. Bagaimana perbandingan jumlah individu yang belum dan telah mendapatkan pekerjaan berdasarkan masing-masing faktor yang diamati?
2. Bagaimana *hazard ratio* seseorang mendapatkan pekerjaan berdasarkan faktor-faktor yang diamati?
3. Apa saja faktor-faktor yang mempengaruhi waktu tunggu kerja di DIY berdasarkan hasil regresi *cox proportional hazard* dan kurva kaplan meier yang disertai uji *log-rank*?

## **1.3 Batasan Masalah**

Penelitian ini merupakan penelitian kuantitatif yang menggunakan metode analisis survival melalui pendekatan regresi *cox proportional hazard* dan uji *log-*

*rank*. Data yang digunakan adalah data Sakernas Agustus 2022 yang bersumber dari Badan Pusat Statistik Daerah Istimewa Yogyakarta. Variabel bebas yang digunakan adalah jenis kelamin, usia, hubungan dengan kepala rumah tangga, status perkawinan, tingkat pendidikan, pengalaman kerja, pelatihan, dan klasifikasi tempat tinggal. Sedangkan variabel terikat yang digunakan adalah lama waktu mencari pekerjaan. Data yang digunakan pada penelitian ini dibatasi dengan hanya melibatkan responden yang belum mendapatkan pekerjaan sampai bulan Agustus 2022 dan responden yang mulai bekerja dalam waktu kurang dari satu tahun pada bulan Agustus 2022.

#### **1.4 Tujuan Penelitian**

Penelitian ini memiliki tujuan diantaranya:

1. Mengetahui perbandingan jumlah individu yang belum dan telah mendapatkan pekerjaan berdasarkan faktor-faktor yang diamati.
2. Mendapatkan *hazard ratio* seseorang mendapatkan pekerjaan pada suatu periode tertentu berdasarkan faktor yang mempengaruhi.
3. Mengetahui faktor-faktor yang mempengaruhi waktu tunggu kerja seseorang di DIY berdasarkan hasil regresi cox *proportional hazard* dan uji *log-rank*.

#### **1.5 Manfaat Penelitian**

Secara terapan, hasil penelitian ini dapat digunakan untuk mengetahui dan memahami faktor apa saja yang berpengaruh terhadap lama waktu mencari kerja hingga memperoleh pekerjaan. Selain itu, dapat juga mengetahui bagaimana implementasi analisis survival untuk mengetahui signifikansinya.

Secara praktis, hasil penelitian ini diharapkan dapat dijadikan sebagai bahan referensi untuk menyusun kebijakan ketenagakerjaan, seperti kegiatan pelatihan, atau metode yang lebih efektif dalam membantu angkatan kerja untuk dapat memperoleh pekerjaan dalam waktu yang lebih efisien. Hasil penelitian ini juga dapat digunakan sebagai bahan penelitian selanjutnya.



## **BAB 2 TINJAUAN PUSTAKA**

### **2.1 Penelitian Terdahulu**

Penulis menjadikan beberapa penelitian terdahulu sebagai dasar dan acuan dalam melakukan penelitian tugas akhir ini, dimana beberapa penelitian terdahulu tersebut memiliki kesamaan baik pada tema, variabel, ataupun metode yang digunakan. Dari beberapa penelitian tersebut penulis melakukan penelitian terbaru dengan beberapa pembaruan.

Faisal Siddiq dan Mohammad Dokhi (2022) dengan penelitiannya yang berjudul *Survival Analisis Durasi Mengganggu Angkatan Kerja Disabilitas yang Mengalami Berhenti Bekerja Akibat Pandemi Covid-19*, menguraikan tentang variabel bebas yang diduga dapat mempengaruhi lama durasi angkatan kerja disabilitas mengganggu di masa pandemi sampai mendapatkan pekerjaan kembali dengan variabel yang digunakan adalah tingkat pendidikan, pelatihan bersertifikat, jenis kelamin, umur, klasifikasi wilayah tempat tinggal, status kepala rumah tangga, dan alasan utama berhenti bekerja. Metode yang digunakan adalah analisis deskriptif, estimasi Kaplan-Meier, dan uji *Log-Rank* yang menghasilkan bahwa variabel tingkat pendidikan, umur, dan status kepala rumah tangga adalah variabel-variabel bebas yang mempengaruhi secara signifikan (Siddiq & Dokhi, 2022). Perbedaan penelitian tersebut dengan penelitian ini terletak pada populasi yang menjadi sasaran penelitian. Penelitian tersebut fokus pada penyandang disabilitas, sementara penelitian ini mencakup semua pencari kerja dalam suatu daerah.

Yosef Hartoko (2019), dengan penelitiannya yang berjudul *Pengaruh Pendidikan, Pelatihan, Jenis Kelamin, Umur, Status Perkawinan, dan Daerah Tempat Tinggal Terhadap Lama Mencari Kerja Tenaga Kerja Terdidik di Indonesia*, menguraikan tentang variabel bebas yang diduga mempengaruhi lama waktu mencari kerja bagi tenaga kerja terdidik di Indonesia meliputi variabel tingkat pendidikan, pelatihan kerja, jenis kelamin, umur, status perkawinan, dan klasifikasi tempat tinggal. Dengan data yang digunakan adalah data sekunder yaitu data SAKERNAS 2015 yang didapat dari BPS. Metode analisis yang digunakan dalam penelitian ini adalah analisis regresi robust. Dari hasil penelitian ini

didapatkan bahwa variabel bebas yang mempengaruhi lama waktu mencari kerja secara signifikan adalah tingkat pendidikan, umur, status perkawinan, pelatihan, dan jenis kelamin (Hartoko, 2019). Dibandingkan dengan penelitian ini, penulis menggunakan variabel bebas tambahan yaitu hubungan dengan kepala rumah tangga dan pengalaman kerja. Selain itu metode analisis yang digunakan oleh penulis adalah regresi *cox proportional hazard*, berbeda dengan penelitian ini yang menggunakan metode regresi robust.

Fahrudin Arrozi dan Ketut Sutrisna (2018), dengan penelitiannya yang berjudul Analisis Faktor-faktor Yang Mempengaruhi Lama Mencari Kerja Bagi Tenaga Kerja Terdidik Di Kota Denpasar menguraikan faktor-faktor yang diduga mempengaruhi lama mencari kerja bagi tenaga kerja terdidik di kota Denpasar dengan variabel bebas meliputi upah, tingkat pendidikan, umur, dan jenis kelamin dengan menggunakan pendekatan analisis regresi linier. Melalui penelitian ini dihasilkan bahwa variabel bebas yang berpengaruh terhadap lama mencari kerja adalah upah, tingkat pendidikan, umur, dan jenis kelamin (Arrozi & Sutrisna, 2018). Dibandingkan dengan penelitian ini, penulis menggunakan variabel bebas tambahan yaitu klasifikasi tempat tinggal, hubungan dengan kepala rumah tangga, status perkawinan, pengalaman kerja, dan pelatihan kerja. Selain itu metode analisis yang digunakan oleh penulis adalah regresi *cox proportional hazard*, berbeda dengan penelitian ini yang menggunakan metode analisis regresi linier.

Muji Nurjanah Susenati (2015), dengan penelitiannya yang berjudul Analisis Lama Waktu Mencari Kerja dengan Pendekatan Regresi Cox Proportional Hazard menguraikan tentang faktor-faktor yang diduga mempengaruhi lama mencari kerja berdasarkan persamaan regresi *cox proportional hazard* terbaik yang terbentuk beserta hazardnya dengan menggunakan pendekatan regresi *cox*. Variabel bebas yang digunakan antara lain jenis kelamin, tingkat pendidikan, usia, hubungan dengan kepala rumah tangga, status perkawinan, klasifikasi tempat tinggal, dan pengalaman kerja, dengan data yang digunakan adalah data sekunder dari BPS Provinsi DIY. Hasil penelitian ini menunjukkan bahwa faktor yang berpengaruh terhadap lama mencari kerja bagi tenaga kerja terdidik di D.I. Yogyakarta adalah usia, hubungan dengan kepala keluarga, pendidikan, dan

pengalaman bekerja (Susenati, 2015). Dibandingkan penelitian ini, penulis menggunakan variabel bebas tambahan yaitu pelatihan kerja. Selain itu, penelitian ini hanya melibatkan responden yang merupakan pengangguran terdidik yaitu lulusan SMA, SMK, dan perguruan tinggi. Sedangkan penulis melibatkan responden dari seluruh tingkat pendidikan mulai dari tidak/belum tamat SD sampai dengan S2. Kemudian penulis juga melakukan analisis tambahan dalam penelitian, yaitu melakukan uji *log-rank* sebagai pembanding terhadap hasil metode regresi *cox proportional hazard*.

I Gede Ari Sudana, dkk. (2013) dengan penelitiannya yang berjudul Penerapan Regresi Cox Proportional Hazard untuk Menduga Faktor-faktor yang Memengaruhi Lama Mencari Kerja, menguraikan tentang variabel bebas yang diduga dapat mempengaruhi lama waktu mencari kerja dengan variabel yang digunakan diantaranya daerah tempat tinggal, hubungan dengan kepala rumah tangga, jenis kelamin, status kawin, tingkat pendidikan, pengalaman kerja, dan umur. Metode yang digunakan adalah analisis regresi *cox proportional hazard*, dimana menghasilkan bahwa daerah tempat tinggal, jenis kelamin, dan status kawin adalah variabel-variabel bebas yang mempengaruhi secara signifikan (Sudana, Suciptawati, & Harini, 2013). Dibandingkan penelitian ini, penulis menggunakan variabel bebas tambahan yaitu pelatihan kerja. Kemudian penulis juga melakukan analisis tambahan dalam penelitian, yaitu melakukan uji *log-rank* sebagai pembanding terhadap hasil metode regresi *cox proportional hazard*. Pada penelitian ini menggunakan data Sakernas 2012 wilayah Provinsi Bali, sedangkan penulis menggunakan data Sakernas 2022 wilayah Provinsi Daerah Istimewa Yogyakarta.

## **BAB 3 LANDASAN TEORI**

### **3.1 Teori Ketenagakerjaan**

Penelitian ini menggunakan konsep teori ketenagakerjaan yang digunakan oleh Badan Pusat Statistik dalam melaksanakan Survei Angkatan Kerja Nasional (SAKERNAS). Konsep yang digunakan tersebut berdasarkan pada *The Labor Force Concept* yang disarankan oleh *International Labor Organization*. Menurut konsep ini, penduduk dibagi menjadi beberapa kelompok. Berdasarkan usia penduduk terbagi menjadi dua kelompok, yaitu penduduk usia kerja dan bukan usia kerja. Penduduk usia kerja adalah penduduk yang berusia 15 tahun ke atas (Badan Pusat Statistik, 2022).

Sementara itu penduduk usia kerja terbagi menjadi dua, yaitu angkatan kerja dan bukan angkatan kerja.

#### **3.1.1 Angkatan Kerja**

Angkatan kerja merupakan sebutan bagi penduduk usia kerja yang bekerja, sementara tidak bekerja, atau pengangguran yang ingin mendapat pekerjaan. Angkatan kerja terbagi menjadi:

1. Bekerja, yaitu aktivitas yang dilakukan seseorang selama minimal 1 jam tanpa terputus dalam seminggu terakhir dengan tujuan mendapatkan atau membantu mendapatkan pendapatan atau keuntungan, baik dibayar maupun tidak dibayar.
2. Sementara tidak bekerja, yaitu kondisi seseorang yang memiliki pekerjaan namun sementara tidak bekerja dalam waktu seminggu terakhir karena berbagai sebab, yaitu: sakit, cuti, menunggu waktu panen, mogok kerja, dan sebagainya.
3. Pengangguran, yang terbagi menjadi:
  - a. Tidak punya pekerjaan dan sedang mencari pekerjaan.
  - b. Tidak punya pekerjaan dan sedang dalam proses mempersiapkan usaha.

- c. Tidak punya pekerjaan dan tidak mencari pekerjaan, karena merasa tidak mungkin untuk mendapatkan pekerjaan.
- d. Sudah mendapatkan pekerjaan, namun belum memulai bekerja (Badan Pusat Statistik, 2022).

Adapun penduduk yang tidak termasuk angkatan kerja adalah penduduk berusia kerja yang masih sekolah, mengurus rumah tangga, atau fokus pada kegiatan lainnya.

### **3.2 Faktor-faktor yang Dapat Mempengaruhi Lama Waktu Mencari Kerja**

Lama waktu mencari kerja dapat dipengaruhi oleh berbagai faktor, beberapa di antaranya adalah:

- a. Jenis kelamin.

Perbedaan antara laki-laki dan perempuan dalam bentuk, sifat, dan fungsi biologisnya menentukan peran mereka dalam reproduksi. Gender, yang sering kali disamakan dengan jenis kelamin, sebenarnya memiliki makna yang berbeda. Gender merujuk pada pembagian peran dan tanggung jawab antara laki-laki dan perempuan yang umumnya ditentukan oleh konstruksi sosial dalam masyarakat dan budaya. Di sisi lain, seks atau jenis kelamin adalah sesuatu yang ditetapkan oleh Tuhan dan tidak dapat diubah (Dr. Ikhlasiah Dalimoenthe, 2020).

Peningkatan partisipasi pekerjaan pada laki-laki selalu lebih tinggi daripada partisipasi pekerjaan perempuan karena laki-laki dianggap sebagai tulang punggung ekonomi keluarga, yang berakibat pada keterlibatan mereka yang lebih aktif dalam mencari pekerjaan yang sesuai dengan aspirasi mereka, termasuk pendapatan dan posisi yang diinginkan. Hampir semua laki-laki yang mencapai usia kerja terlibat dalam kegiatan ekonomi, karena mereka berperan sebagai pencari nafkah utama dalam keluarga (Simanjuntak, 1998).

- b. Usia.

Usia adalah periode waktu sejak kelahiran seseorang dan dapat diukur dalam satuan waktu berdasarkan urutan kronologisnya. Pada individu



normal, kita dapat melihat bahwa tingkat perkembangan anatomi dan fisiologis mereka sejalan (Nuswantari, 1998).

Durasi pencarian pekerjaan bervariasi antara kelompok dalam populasi tenaga kerja, dan semakin bertambah usia, durasi tersebut semakin lama. Tingkat pengangguran yang tinggi di kalangan generasi muda adalah fenomena yang merupakan bagian tak terpisahkan dari struktur sosial, dan tidak dapat dihindari ketika para pemuda lulus sekolah dan mencari pekerjaan di pasar yang penuh dengan persaingan tenaga kerja. Menurut penafsiran ini, hanya tingkat pengangguran yang tinggi di kalangan kelompok usia tua yang memiliki potensi mengakibatkan masalah atau bahaya, karena hal ini mencerminkan ketidakmampuan ekonomi untuk menyerap tenaga kerja inti dalam masyarakat (Simanjuntak, 1998).

Usia bagi para tenaga kerja produktif berada pada rentang usia 20 – 40 tahun. Pada usia di bawah 20 tahun, sebagian besar individu belum memiliki kemampuan yang memadai bahkan masih menempuh pendidikan. Sementara pada individu yang berusia di atas 40 tahun, rata-rata telah mengalami penurunan kemampuan fisik sehingga produktivitasnya menurun (Priyono & Yasin, 2016).

c. Hubungan dengan Kepala Rumah Tangga.

BPS membagi anggota rumah tangga berdasarkan hubungannya dengan kepala rumah tangga ke dalam 11 kelompok, yaitu di antaranya (Badan Pusat Statistik, 2022):

- a) Kepala rumah tangga
- b) Istri/suami
- c) Anak kandung
- d) Anak tiri/angkat
- e) Menantu
- f) Cucu
- g) Orang tua/mertua
- h) Famili lain
- i) Pembantu rumah tangga

j) Sopir/tukang kebun

k) Lainnya (tidak memiliki hubungan dengan kepala rumah tangga)

d. Status Perkawinan.

Menurut Yulianti, dkk (2011), status perkawinan mengacu pada posisi seseorang dalam hubungan perkawinan. Status perkawinan dapat terdiri dari:

a) Belum kawin: Status untuk mereka yang belum menikah pada saat pencatatan.

b) Kawin: Status bagi mereka yang terikat dalam perkawinan pada saat pencatatan, baik tinggal bersama atau terpisah. Dalam hal ini, tidak hanya mencakup mereka yang sah secara hukum, tetapi juga mereka yang dianggap suami istri oleh masyarakat sekitar.

c) Cerai hidup: Status bagi mereka yang hidup terpisah sebagai suami istri karena telah bercerai dan belum menikah lagi. Mereka yang mengklaim bercerai, meskipun belum resmi secara hukum, dianggap bercerai. Sebaliknya, mereka yang sementara tinggal terpisah namun belum resmi bercerai, misalnya karena suami atau istri meninggalkan pasangannya untuk sekolah, bekerja, mencari pekerjaan, atau sedang dalam konflik.

d) Cerai mati: Status bagi mereka yang suami atau istri telah meninggal dunia dan belum menikah lagi (Yulianti, 2011).

Status perkawinan menjadi faktor yang menyebabkan perbedaan antara partisipasi tenaga kerja pria dan wanita. Wanita yang telah menikah sering kali memiliki tanggung jawab rumah tangga yang harus dijalankan, dan pasar kerja terbatas bagi mereka yang belum menikah atau berstatus lajang. Oleh karena itu, status perkawinan memiliki pengaruh terhadap ketersediaan tenaga kerja, karena status perkawinan terkait dengan status saat memasuki pasar kerja (Astuti, 2013).

e. Tingkat Pendidikan.

Tingkat pendidikan merujuk pada tingkat pendidikan yang ditentukan berdasarkan perkembangan peserta didik, tujuan yang ingin dicapai, dan

kemampuan yang dikembangkan. Tingkat pendidikan memiliki pengaruh terhadap perubahan sikap dan perilaku yang mendukung gaya hidup sehat. Tingkat pendidikan yang lebih tinggi mempermudah individu atau masyarakat untuk menerima informasi dan menerapkannya dalam kehidupan sehari-hari, terutama dalam hal kesehatan (Suhardjo, 2007).

Tren peningkatan tingkat pengangguran di kalangan tenaga kerja berpendidikan menjadi permasalahan yang serius. Hal ini dapat dijelaskan dengan fakta bahwa semakin tinggi tingkat pendidikan seseorang, semakin tinggi pula harapannya untuk mendapatkan pekerjaan yang sesuai. Kelompok pencari kerja yang berpendidikan lebih cenderung menghabiskan waktu yang lebih lama dalam mencari pekerjaan, karena mereka memiliki pengetahuan yang lebih baik tentang perkembangan informasi di pasar kerja serta kemampuan untuk memilih pekerjaan yang diinginkan dan menolak pekerjaan yang tidak diinginkan (Simanjuntak, 1998).

f. Pengalaman Kerja.

Dalam perkiraan, pengalaman kerja dapat meningkatkan kemampuan pencari kerja untuk mendapatkan pekerjaan yang sesuai, karena pengalaman kerja mencerminkan pengetahuan tentang pasar kerja. Dengan memiliki kombinasi pengalaman kerja dan tingkat pendidikan yang tinggi, tenaga kerja akan memiliki peluang yang lebih banyak untuk mendapatkan pekerjaan (Simanjuntak, 1998).

g. Pelatihan Kerja.

Pelatihan bertujuan untuk meningkatkan keterampilan karyawan baik secara horizontal maupun vertikal. Peningkatan horizontal berarti memperluas pengetahuan dan keterampilan dalam berbagai aspek atau jenis pekerjaan. Peningkatan vertikal berarti mendalami pemahaman dalam bidang tertentu (Simanjuntak, 1998).

h. Klasifikasi Tempat Tinggal.

Tempat tinggal merupakan faktor yang mempengaruhi tingkat pengangguran, terutama dalam konteks perbedaan antara desa dan kota.

Di perkotaan, intensitas lapangan kerja dapat bervariasi, dan ketidakseimbangan antara permintaan dan penawaran tenaga kerja dapat menyebabkan tingkat pengangguran yang tinggi. Calon tenaga kerja di perkotaan cenderung menjadi lebih selektif dalam memilih pekerjaan yang sesuai dengan minat mereka, sehingga jika pekerjaan yang tersedia tidak sesuai dengan minat mereka, mereka lebih cenderung menjadi pengangguran. Di sisi lain, di desa, calon tenaga kerja cenderung kurang selektif, sehingga tingkat pengangguran tidak sebesar di kota (Adyaksa, 2020).

Hubungan antara tempat tinggal dan tingkat pengangguran terdidik terlihat pada kedua wilayah, yaitu desa dan kota. Banyaknya penduduk yang bermigrasi ke kota untuk mencari pekerjaan yang diimpikan mereka menyebabkan ketidakseimbangan antara permintaan dan penawaran tenaga kerja di kota. Penawaran tenaga kerja jauh lebih banyak daripada permintaan, yang mengakibatkan tingkat pengangguran yang tinggi di kota (Pratomo, 2017).

### **3.3 Statistika Deskriptif**

Statistika deskriptif adalah sebuah jenis analisis statistik terpopuler yang digunakan dalam penyajian data. Biasanya statistika deskriptif juga digunakan untuk menyajikan data dalam bentuk visual dan memberikan gambaran karakteristik data secara umum. Visualisasi data sebagai bentuk deskriptif dapat tersajikan dalam berbagai bentuk, seperti diagram lingkaran, diagram titik, diagram batang, dan sebagainya (Martias, 2021).

#### **3.3.1 Diagram Batang**

Diagram batang biasanya digunakan untuk menyajikan data berskala nominal dengan menampilkan frekuensi dari masing-masing kategori yang berbeda. Kategori dinyatakan oleh sumbu X, sementara frekuensi dinyatakan oleh sumbu Y (Fatimah, 2021).

### 3.4 Analisis Survival

Analisis survival adalah serangkaian prosedur statistika yang digunakan untuk menganalisis data waktu terkait dengan kejadian tertentu. Data yang digunakan dalam analisis ini mencakup interval waktu antara kejadian serta data yang terkait dengan waktu, dimulai dari waktu awal (time origin) hingga terjadinya peristiwa spesifik. Peristiwa spesifik tersebut dapat berupa kegagalan, kematian, kekambuhan penyakit, respons dari percobaan, atau peristiwa lain yang dipilih berdasarkan kepentingan peneliti. Selain itu, peristiwa spesifik juga dapat mencakup kejadian positif seperti kelahiran, kelulusan sekolah, kesembuhan dari penyakit, atau pengalaman positif lainnya (Kleinbaum & Klein, 2005).

Analisis survival sering digunakan dalam berbagai bidang, seperti biologi, kedokteran, kesehatan umum, sosiologi, teknik, ekonomi, demografi, dan epidemiologi. Contohnya, analisis survival dapat digunakan untuk mempelajari daya tahan hidup pasien kanker paru-paru dalam bidang kedokteran. Di bidang teknik, analisis survival dapat digunakan untuk menganalisis masa hidup lampu pijar. Selain itu, dalam bidang ekonomi, demografi, dan epidemiologi, analisis survival juga dapat diterapkan untuk berbagai keperluan penelitian (Collet, 2003).

Pemodelan analisis survival memiliki 3 jenis model atau bentuk, yaitu parametrik, semiparametrik, dan nonparametrik. Model survival parametrik merupakan model dengan sebaran data waktu survival yang diasumsikan mengikuti suatu sebaran distribusi parametrik, seperti di antaranya distribusi Eksponensial, Weibull, dan Log-logistik (Kleinbaum & Klein, 2005).

Sedangkan model survival semiparametrik adalah model survival yang tidak didasari oleh suatu bentuk distribusi data, yang mana salah satunya adalah model regresi cox *proportional hazard* (Chandra & Rohmaniah, 2019).

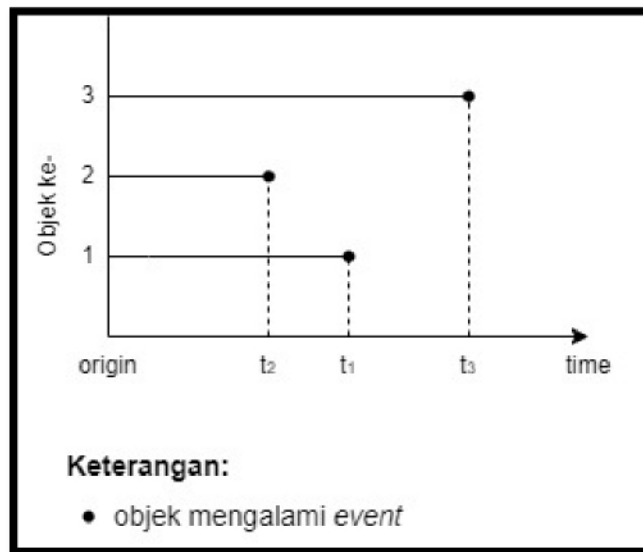
Sementara itu estimasi nonparametrik merupakan estimasi yang dalam perhitungannya dilakukan secara empiris atau tidak mengikuti suatu distribusi data. Dengan kata lain, estimasi dilakukan dengan hanya mengikuti pola data yang diberikan oleh data sampel (Abdullah, 2022). Contoh dari metode estimasi nonparametrik adalah metode Kaplan-Meier dan Nelson-Aalen.

### 3.4.1 Data Survival

Analisis survival melibatkan penggunaan data yang mencatat waktu antara kejadian-kejadian tertentu. Waktu survival dapat didefinisikan sebagai jangka waktu antara titik awal pengamatan hingga terjadinya peristiwa kegagalan, yang dapat diukur dalam satuan hari, bulan, atau tahun. Titik awal, yang juga dikenal sebagai *time origin* atau *start-point*, merujuk pada waktu saat terjadinya kejadian awal yang menjadi fokus analisis, seperti saat seseorang didiagnosis menderita kanker atau saat pemberian perlakuan tertentu. Sementara itu, waktu kegagalan, yang juga disebut sebagai *failure time* atau *end-point*, mengacu pada waktu terjadinya peristiwa akhir, seperti kematian, kesembuhan, atau kejadian lainnya (Collet, 2003).

Dalam analisis survival, terdapat tiga elemen penting yang perlu diperhatikan dalam menentukan waktu survival (Abdullah, 2022):

1. Waktu awal pencatatan (*time origin* atau *start point*): Merupakan waktu saat pengamatan dimulai atau saat terjadi kejadian awal yang menjadi fokus analisis.
2. Waktu akhir pencatatan (*failure event* atau *end point*): Merupakan waktu saat terjadi peristiwa kegagalan atau akhir pengamatan, seperti kematian atau kejadian lain yang menjadi fokus analisis.
3. Skala pengukuran: Skala pengukuran waktu harus jelas, baik dalam satuan hari, bulan, atau tahun, untuk membatasi periode kejadian yang diamati.



**Gambar 3.1** Ilustrasi data survival

### 3.4.2 Teori Data Tersensor

Dalam melakukan observasi data survival, seringkali ditemukan adanya data tersensor. Data tersensor merujuk pada data yang tidak lengkap karena adanya individu yang hilang atau tidak dapat diamati sampai akhir penelitian, sehingga data mereka tidak dapat diambil. Dalam konteks ini, pada akhir periode pengamatan, individu-individu tersebut belum mengalami peristiwa kegagalan (*failure event*). Namun, jika pada akhir pengamatan individu tersebut telah mengalami peristiwa kegagalan, maka mereka dianggap tidak tersensor. Dalam kedua kasus tersebut, pengamatan terhadap individu yang tersensor atau tidak tersensor memiliki implikasi penting dalam analisis survival (Collet, 2003).

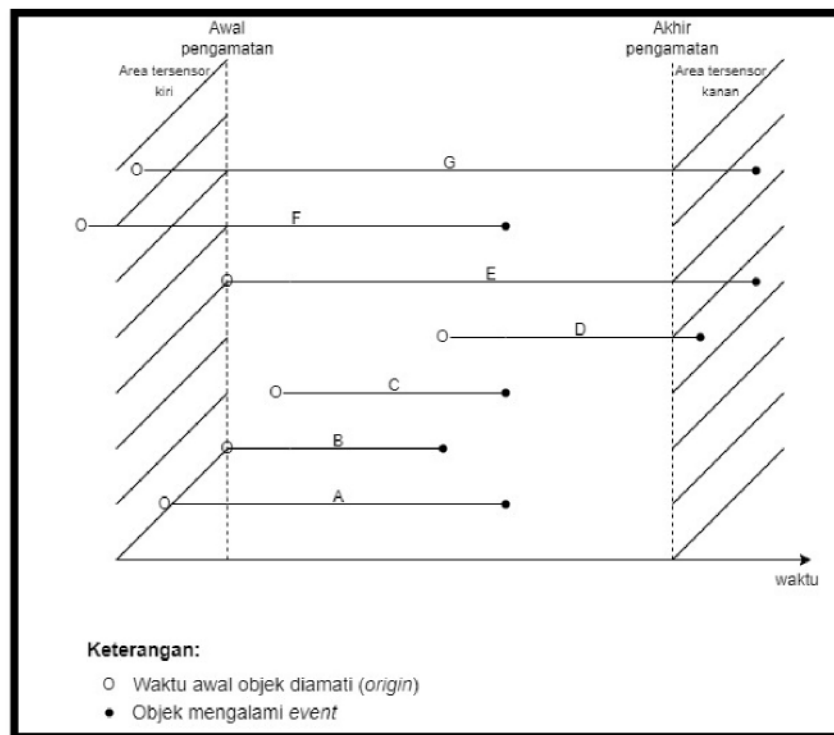
Dalam mendapatkan data survival yang mencakup seluruh peristiwa kejadian individu hingga terjadinya kejadian tersebut, waktu yang diperlukan dapat menjadi sangat lama, menyebabkan pengamatan menjadi tidak efektif dan biaya pengamatan yang tinggi. Oleh karena itu, penyensoran data menjadi suatu kebutuhan penting untuk mengatasi hal tersebut. Konsep penyensoran ini membedakan analisis survival dari ilmu statistik lainnya (Kleinbaum & Klein, 2005).

Terdapat empat jenis penyensoran data survival, yaitu:

1. Penyensoran kanan (*right censor*), yaitu penyensoran yang terjadi jika objek yang diamati mengalami *event* setelah melewati suatu titik waktu tertentu yang tidak teramati. Sehingga waktu tepatnya terjadi event tersebut tidak diketahui secara pasti.
2. Penyensoran kiri (*left censor*), yaitu penyensoran yang terjadi jika objek yang diamati mengalami *origin* sebelum titik waktu tertentu yang menjadi awal pengamatan. Sehingga waktu tepatnya terjadi event tersebut tidak diketahui secara pasti.
3. Penyensoran ganda, yaitu penyensoran yang terjadi jika pada suatu data penelitian terdapat data yang tersensor kiri dan tersensor kanan.
4. Penyensoran interval, yaitu penyensoran yang terjadi jika objek yang diteliti mengalami *event* dalam selang antara dua titik waktu dan tidak diketahui secara pasti (Abdullah, 2022).

Penyensoran data survival dapat diilustrasikan seperti terlampir pada

**Gambar 3.2.**



**Gambar 3.2** Ilustrasi penyensoran data



Pada gambar tersebut terlihat bahwa observasi data survival memiliki dua batas waktu yakni waktu awal pengamatan dan akhir pengamatan. Ketika suatu objek penelitian memiliki waktu awal yang diamati (*origin*) sebelum pengamatan dimulai, maka akan ada sebagian waktu di mana objek tersebut tidak teramati sehingga data tersebut menjadi data yang tersensor kiri. Jika suatu objek penelitian mengalami kejadian (*event*) setelah melewati waktu akhir pengamatan, maka akan ada juga sebagian waktu di mana objek tersebut tidak teramati sehingga data tersebut menjadi data tersensor kanan. Pada gambar tersebut terlihat bahwa objek B dan C merupakan objek yang akan menjadi data terobservasi lengkap atau tidak tersensor. Sementara objek A, F, dan G akan menjadi data yang tersensor kiri, lalu objek D, E, dan G akan menjadi data tersensor kanan, dan objek G akan menjadi data yang tersensor ganda.

### 3.4.3 Fungsi Survival

Fungsi survival adalah suatu fungsi probabilitas dari suatu objek yang diamati telah berhasil melewati suatu waktu tertentu sebelum terjadinya *event* pada objek tersebut. Misalkan, waktu tersebut dikatakan sebagai  $t$ , maka *event* yang terjadi pada objek tersebut akan terjadi pada saat  $T > t$ . Fungsi survival dinyatakan sebagai  $S(t)$ . Jika suatu fungsi kepadatan peluang dari  $T$  dinyatakan sebagai  $f(t)$ , fungsi kepadatan peluang kumulatif sebagai  $F(t)$ , maka hubungan antara  $S(t)$ ,  $f(t)$ , dan  $F(t)$  adalah sebagai berikut (Abdullah, 2022):

$$\begin{aligned} S(t) &= P(T > t) \\ &= 1 - P(T \leq t) \\ &= 1 - F(t) \end{aligned} \tag{3.1}$$

Sementara itu

$$F(t) = \int_0^t f(t)dt \tag{3.2}$$

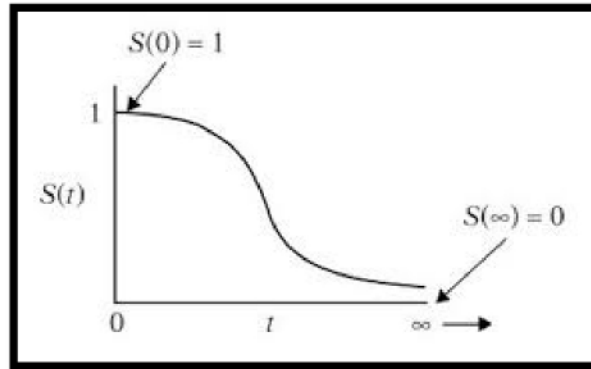
Dengan:

$f(t)$ : Fungsi kepadatan peluang dari  $T$

$F(t)$ : Fungsi kepadatan peluang kumulatif dari  $T$

$S(t)$ : Fungsi peluang ketahanan suatu objek terhadap *event* sampai dengan  $t$

Fungsi survival merupakan fungsi yang bersifat monoton turun, dimana ketika  $t = 0$  maka  $S(t)$  akan bernilai 1. Sementara ketika  $t \rightarrow \infty$  maka  $S(t)$  akan bernilai 0. Hal ini menunjukkan bahwa semakin lama berjalannya waktu, maka ketahanan suatu objek terhadap suatu kejadian akan semakin menurun. Sebagai ilustrasi ditampilkan contoh bentuk fungsi survival pada **Gambar 3.3**.



**Gambar 3.3** Contoh bentuk fungsi survival

#### 3.4.4 Fungsi Hazard dan Hazard Kumulatif

Fungsi *hazard* adalah sebuah fungsi yang menyatakan seberapa besar tingkat bahaya atau resiko terjadinya *event* pada suatu waktu tertentu setelah bertahan melewati waktu tersebut (misal  $t$ ), dan fungsi ini bukan merupakan fungsi probabilitas. Fungsi *hazard* dinyatakan sebagai  $h(t)$ , dengan penjabaran sebagai berikut (Abdullah, 2022):

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t < T \leq t + \Delta t | T > t)}{\Delta t} \quad (3.3)$$

Sementara itu fungsi *hazard* kumulatif dinyatakan sebagai  $H(t)$  yaitu:

$$H(t) = \int_0^t h(u) du, t > 0 \quad (3.4)$$

Hubungan antara fungsi *hazard* dan fungsi survival dapat dinyatakan sebagai berikut:

$$h(t) = \frac{f(t)}{S(t)} \quad (3.5)$$

Dengan:

$h(t)$  : Fungsi *hazard*

$H(t)$  : Fungsi *hazard* kumulatif

### 3.5 Estimasi dan Kurva Kaplan-Meier

Kaplan-Meier adalah sebuah metode nonparametrik yang digunakan untuk mengestimasi fungsi *survival* dan dapat menangani data yang tersensor, serta dapat digunakan pada data sampel berukuran kecil (Kaplan & Meier, 1958). Kaplan-Meier merupakan salah satu teknik statistik yang sangat berguna dalam mengestimasi probabilitas kelangsungan hidup seseorang dalam periode waktu tertentu. Biasanya, metode ini digunakan untuk menyimpulkan dan meringkas pengalaman bertahan hidup. Persamaan model Kaplan-Meier dapat dituliskan sebagai berikut (Harlan, 2017):

$$S(t) = \frac{n_i - d_i}{n_i} \quad (3.6)$$

Dengan:

$S(t)$  : Peluang suatu objek *survive* sampai waktu ke- $t$

$n_i$  : Banyaknya objek yang berisiko namun masih bertahan pada waktu  $i$

$d_i$  : Banyaknya objek yang mengalami *event* pada waktu  $i$

Hasil dari estimasi Kaplan-Meier kemudian dapat digambarkan menjadi sebuah kurva. Pada setiap titik waktu ( $t$ ), kurva Kaplan Meier menunjukkan proporsi kumulatif individu yang masih bertahan setelah mengalami peristiwa tertentu. Notasi yang digunakan untuk menyatakan nilai kurva Kaplan Meier pada waktu ( $t$ ) adalah  $S(t)$ , dan dapat dihitung menggunakan persamaan berikut (Salsabila, 2020):

$$S(t) = S(t - 1) \times \left(1 - \frac{d_i}{n_i}\right) \quad (3.7)$$

Dengan:

$d_i$  : Jumlah individu yang mengalami peristiwa pada waktu  $i$

$n_i$  : Jumlah individu yang masih bertahan pada waktu  $i$

### 3.6 Uji Log-Rank

Uji *Log-Rank* merupakan metode statistik yang digunakan untuk membandingkan dua kelompok atau lebih dalam analisis survival, dan akan sangat berguna ketika mengamati dan menginterpretasi kurva survival *Kaplan-Meier*. Tujuan dari uji *Log-Rank* adalah untuk menguji apakah terdapat perbedaan yang

signifikan dalam fungsi survival antara kelompok-kelompok yang dibandingkan. Langkah-langkah dalam melakukan uji *Log-Rank* adalah sebagai berikut (Salsabila, 2020):

1. Menentukan frekuensi terjadinya *event* pada setiap kelompok di interval waktu tertentu.
2. Menghitung nilai sumbu kumulatif absolut ( $n_{1i}$  dan  $n_{2i}$ ) pada setiap waktu  $i$  untuk masing-masing kelompok.
3. Menghitung nilai harapan ( $e_{1i}$  dan  $e_{2i}$ ) pada setiap waktu  $i$  untuk masing-masing kelompok dengan rumus:

$$e_{1i} = \frac{n_{1i}}{n_{1i}+n_{2i}} \times d_i \text{ dan } e_{2i} = \frac{n_{2i}}{n_{1i}+n_{2i}} \times d_i \quad (3.8)$$

Dengan:

$n_{.i}$ : Banyaknya objek yang bertahan sampai dengan waktu ke- $i$

$e_{.i}$ : Ekspektasi banyaknya terjadi *event* pada waktu ke- $i$

4. Menghitung nilai total harapan ( $E_1$  dan  $E_2$ ) dengan menjumlahkan nilai harapan ( $e_{1i}$  dan  $e_{2i}$ ) untuk seluruh interval waktu.

Selanjutnya dilakukan langkah-langkah untuk menguji hipotesis sebagai berikut (Riyandianci, 2017):

- i. Hipotesis:

$H_0$ : Tidak ada perbedaan signifikan dalam fungsi survival antara kelompok-kelompok yang dibandingkan.

$H_1$ : Terdapat perbedaan signifikan dalam fungsi survival antara kelompok-kelompok yang dibandingkan.

- ii. Statistik Uji:

Statistik uji yang digunakan dalam uji Log Rank adalah Chi-Square dan dihitung menggunakan rumus berikut:

$$\chi^2 = \sum_{i=1}^G \frac{(O_i - E_i)^2}{E_i} \quad (3.9)$$

Dengan:

$O_i$ : Nilai observasi objek kelompok ke- $i$

$E_i$ : Nilai ekspektasi objek kelompok ke- $i$

$G$ : Banyaknya kelompok

iii. Kriteria Penolakan:

$H_0$  akan ditolak jika nilai  $\chi^2_{hitung} > \chi^2_{\alpha, G-1}$  atau  $p\text{-value} < \alpha$ .

### 3.7 Regresi

Regresi adalah sebuah metode untuk mengukur hubungan antara dua variabel atau lebih dengan hasil berupa suatu fungsi. Dengan pemisahan yang jelas antara variabel bebas ( $x$ ) dengan variabel terikat ( $y$ ), dimana ada ketergantungan di antara keduanya. Kedua variabel tersebut biasanya bersifat saling berpengaruh. Sehingga dapat disimpulkan bahwa regresi adalah sebuah bentuk fungsi yang dapat dinyatakan dengan  $y = f(x)$  (Kurniawan & Yuniarto, 2016).

Regresi digunakan untuk menentukan sifat – sifat dan kekuatan hubungan antara dua variabel serta memprediksi nilai dari suatu variabel yang belum diketahui dengan didasarkan pada observasi masa lalu terhadap variabel tersebut dan variabel-variabel lainnya. Selanjutnya analisis regresi ditujukan untuk:

1. Membuat estimasi rata-rata dan nilai variabel tergantung dengan didasarkan pada nilai variabel bebas.
2. Menguji hipotesis karakteristik dependensi
3. Untuk meramalkan nilai rata-rata variabel bebas dengan didasarkan pada nilai variabel bebas diluar jangkauan sampel (Juharsah, et al., 2020).

#### 3.7.1 Regresi Cox *Proportional Hazard*

Metode ini disebut sebagai regresi cox *proportional hazard* karena memiliki asumsi yang harus terpenuhi, yaitu asumsi *proportional hazard* atau rasio dari fungsi *hazard* dua individu yang berbeda adalah konstan. Persamaan regresi cox adalah model yang berdistribusi semiparametrik, karena dalam persamaan cox tidak diperlukan informasi terkait distribusi khusus yang mendasari waktu survival dan dalam melakukan estimasi parameternya tidak harus ditentukan fungsi *hazard* dasarnya (Susenati, 2015).

Hubungan antara variabel bebas (independen) terhadap variabel terikat (dependen) yang merupakan waktu survival dapat terlihat dari model cox yang dihasilkan, yaitu berupa fungsi *hazard*.  $X$  diasumsikan sebagai variabel bebas yang tidak dipengaruhi oleh waktu, dengan melambangkakan  $x_1, x_2, \dots, x_p$  adalah nilai

dari variabel bebas  $X_1, X_2, \dots, X_p$  yang dapat mempengaruhi risiko kematian individu pada waktu tertentu. Bentuk model regresi cox yang akan dihasilkan adalah (Kleinbaum & Klein, 2005):

$$\begin{aligned} h(t, x) &= h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p) \\ &= h_0(t) \exp(\sum_j^p \beta_j x_j) \end{aligned} \quad (3.10)$$

Dengan:

$h_0(t)$  : Fungsi dasar *hazard*

$\beta_1, \beta_2, \dots, \beta_p$  : Parameter regresi

$x_1, x_2, \dots, x_p$  : Nilai dari variabel bebas  $X_1, X_2, \dots, X_p$ .

Model cox *proportional hazard* yang dihasilkan tetap dapat memberikan informasi yang berguna yaitu berupa rasio *hazard* yang tidak bergantung pada nilai  $h_0(t)$  yang mana merupakan fungsi *hazard* dasar, meskipun bentuk fungsional dari fungsi *hazard* dasarnya tidak diketahui. Namun jika  $h_0(t)$  diketahui maka perhitungan nilai *hazard* akan mengikuti sebaran distribusi data waktu tertentu sehingga akan menjadikannya sebagai persamaan parametrik. Sementara itu *hazard ratio* dalam regresi cox dihasilkan dengan menghitung rasio melalui pembagian nilai *hazard rate* satu individu terhadap *hazard rate* individu lainnya. Sebagai penjabaran akan diberikan contoh seperti berikut:

Misalnya terdapat suatu individu A yang memiliki *hazard rate*  $h_A(t, X^A)$ , dengan  $X^A = (X_1^A, X_2^A, \dots, X_j^A)$  dan suatu individu B yang memiliki *hazard rate*  $h_B(t, X^B)$ , dengan  $X^B = (X_1^B, X_2^B, \dots, X_j^B)$ . Dengan menghitung rasio *hazard* diantara keduanya akan menghasilkan:

$$\begin{aligned} HR &= \frac{h_A(t, X^A)}{h_B(t, X^B)} = \frac{h_0(t) \exp(\sum_j^p \beta_j X_j^A)}{h_0(t) \exp(\sum_j^p \beta_j X_j^B)} \\ &= \exp(\sum_j^p \beta_j X_j^A - \sum_j^p \beta_j X_j^B) \\ &= \exp[\sum_j^p \beta_j (X_j^A - X_j^B)] \end{aligned} \quad (3.11)$$

Apabila nilai *hazard ratio* (HR) selalu bernilai konstan, maka  $X_1, X_2, \dots, X_j$  akan memenuhi asumsi *proportional hazard*. Sehingga koefisien regresi ( $\beta$ ) tetap dapat ditaksir meskipun fungsi  $h_0(t)$  tidak diketahui bentuknya.

Nilai *hazard ratio* akan dapat menunjukkan adanya peningkatan maupun penurunan resiko dari individu yang mendapat perlakuan tertentu. Interpretasi nilai *hazard ratio* berdasarkan hasil perhitungan dalam persamaan 3.11 adalah sebagai berikut:

1. Jika nilai  $HR > 1$ , maka individu A memiliki resiko yang lebih tinggi dibanding individu B.
2. Jika nilai  $HR < 1$ , maka individu A memiliki resiko yang lebih rendah dibanding individu B.
3. Jika  $HR = 1$ , maka individu A dan B memiliki tingkat resiko yang sama.

### 3.7.2 Estimasi Parameter

Parameter dalam persamaan yang terbentuk dari regresi *cox proportional hazard* didapatkan melalui penggunaan metode *Maximum Partial Likelihood* (MPLE). Estimasi nilai dari  $\beta$  adalah nilai ketika fungsi *partial likelihood* maksimum. Fungsi *partial likelihood* secara umum adalah sebagai berikut:

$$L(\beta) = \prod_{j \in D} \frac{\exp(\beta x_j)}{\sum_{j \in R_k} \exp(\beta x_j)} \quad (3.12)$$

Dengan  $x$  adalah variabel penjelas,  $\beta$  adalah parameter regresi,  $D$  adalah himpunan indeks  $j$  dari seluruh waktu kejadian, dan  $R_k$  adalah himpunan resiko dari semua individu yang belum mengalami kejadian pada saat waktu tertentu. Kemudian untuk mempermudah estimasi nilai maksimum  $L(\beta)$ , persamaan di atas dapat ditransformasikan dengan penggunaan  $\ln$ , sehingga dapat dituliskan sebagai berikut:

$$\begin{aligned} L(\beta) &= \ln \left( \prod_{j \in D} \frac{\exp(\beta x_j)}{\sum_{j \in R_k} \exp(\beta x_j)} \right) \\ &= \sum_{j=1}^p \left( \ln \frac{\exp(\beta x_j)}{\sum_{j \in R_k} \exp(\beta x_j)} \right) \\ &= \sum_{j=1}^p [\beta x_j - \ln \sum_{j \in R_k} \exp(\beta x_j)] \end{aligned} \quad (3.13)$$

Nilai estimasi dari  $\beta$  bisa didapatkan dengan memaksimalkan fungsi *log partial likelihood*, yaitu dengan mencari solusi dari persamaan:

$$\frac{\partial \ln L(\beta)}{\partial \beta} = 0 \quad (3.14)$$

Metode yang digunakan untuk pemaksimuman tersebut adalah prosedur iterasi *Newton-Raphson*.

### 3.7.3 Ties dalam Partial Likelihood

*Ties* dalam analisis survival merupakan terjadinya *failure event* yang bersamaan di antara dua individu atau lebih. Hal ini akan menyebabkan adanya permasalahan dalam *partial likelihood* yaitu pada saat menentukan anggota dari suatu himpunan resikonya. Sebagai contoh kasus ditampilkan tabel sebagai berikut:

**Tabel 3.1** Contoh Kasus Data *Ties*

$i$	1	2	3	4	5
$t_i$	2	4	6	6	14

Misal  $i$  adalah individu ke- $i$  dan  $t_i$  adalah waktu terjadinya *event* pada individu ke- $i$ , dengan  $t_1 < t_2 < t_3 < t_4 < t_5$  adalah waktu kejadian yang telah diurutkan. Pada saat  $t = 6$ , terdapat dua individu yang mengalami *event* dan tidak diketahui mana di antara kedua individu tersebut yang terlebih dahulu mengalami *event*. Kejadian bersama tersebut akan menimbulkan masalah dalam melakukan estimasi parameter yang berkaitan dengan penentuan anggota dari himpunan resiko. Salah satu metode alternatif yang dapat digunakan sebagai pendekatan untuk mengatasi adanya data *ties* adalah metode *efron*.

Secara umum, metode pendekatan *efron* memiliki persamaan *partial likelihood* sebagai berikut (Klein & Moeschberger, 2003):

$$L(B_{Efron}) = \prod_{j \in D} \frac{\exp(\beta S_k)}{\prod_{j=1}^{d_k} [\sum_{i \in R(t_j)} \exp(x_i \beta) - \frac{j-1}{d_k} \sum_{i \in D(t_j)} \exp(x_i \beta)]} \quad (3.15)$$

Dengan  $S_k$  adalah jumlah kovariat  $x$  pada kasus terjadinya kejadian bersama (*ties*),  $d_k$  banyaknya *ties* yang terjadi pada waktu  $t_j$ , dan  $R(t_j)$  adalah himpunan resiko.

### 3.7.4 Uji Signifikansi Parameter

Dalam regresi Cox *proportional hazard*, beberapa pengujian signifikansi parameter diperlukan untuk mengetahui apakah variabel bebas memiliki pengaruh yang signifikan terhadap persamaan Cox yang terbentuk. Beberapa pengujian yang digunakan adalah sebagai berikut (Susenati, 2015):

- a. Uji *Partial Likelihood* Rasio.



Uji ini digunakan untuk mengetahui apakah satu atau beberapa parameter  $\beta_j$  sama dengan nol, yang menunjukkan bahwa parameter-parameter tersebut tidak berpengaruh secara signifikan terhadap persamaan Cox yang terbentuk. Statistik uji yang digunakan mengikuti distribusi chi-square dengan derajat kebebasan  $p$ . Berikut adalah rangkaian pengujian hipotesis untuk uji partial likelihood rasio:

i) Hipotesis:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1: \text{minimal ada satu } \beta_i \neq 0, \text{ dengan } i = 1, 2, \dots, k$$

ii) Tingkat Signifikansi:

$$\alpha = 10\% = 0.1$$

iii) Statistik Uji:

$$G = -2[\ln L_R - \ln L_f] \tag{3.16}$$

di mana  $L_R$  adalah partial likelihood model awal, dan  $L_f$  adalah partial likelihood model akhir, mengacu pada persamaan 3.12.

iv) Daerah Penolakan:

$H_0$  ditolak jika nilai  $G \geq \chi^2_{(\alpha; db=p)}$  atau nilai p-value  $\leq \alpha$ , di mana  $p$  adalah banyaknya variabel bebas.

v) Kesimpulan:

Jika  $H_0$  ditolak, maka  $\beta_i \neq 0$  yang mengindikasikan bahwa baik satu ataupun beberapa variabel bebas memberikan pengaruh secara signifikan terhadap waktu survival (variabel dependen).

b. Uji Wald.

Uji Wald digunakan untuk mengetahui variabel-variabel bebas mana yang berpengaruh secara signifikan. Statistik uji pada uji ini dinotasikan dengan  $z$  dan juga mengikuti distribusi chi-square dengan derajat kebebasan  $p$ . Pengujian dilakukan untuk masing-masing  $\beta_j$  dengan  $j = 1, 2, \dots, k$ , dan berikut adalah rangkaian pengujian hipotesisnya:

i) Hipotesis:

$H_0: \beta_i = 0$  (Variabel bebas  $i$  tidak berpengaruh terhadap waktu survival)

$H_1: \beta_i \neq 0$  (Variabel bebas  $i$  berpengaruh terhadap waktu survival)

ii) Tingkat Signifikansi:

$$\alpha = 10\% = 0.1$$

iii) Statistik Uji:

$$Z^2 = \left( \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \right)^2 \quad (3.17)$$

iv) Daerah Penolakan:

$H_0$  ditolak jika nilai  $Z^2 \geq \chi^2_{(\alpha; db=p)}$  atau nilai p-value  $\leq \alpha$

v) Kesimpulan:

Jika  $H_0$  ditolak, maka  $\beta_i \neq 0$ , yang mengindikasikan bahwa kovariat yang diuji berpengaruh secara signifikan terhadap waktu survival (variabel dependen).

Dalam pengujian kontribusi variabel dalam regresi cox *proportional hazard*, baik secara serentak maupun secara parsial, keduanya menggunakan statistik uji yang mengikuti distribusi chi-kwadrat. Namun, dalam hasil perhitungan menggunakan perangkat lunak RStudio, penentuan kontribusi variabel penjelas didasarkan pada nilai p-value yang dihitung menggunakan statistik Z yang mengikuti distribusi normal. Hal ini terkait dengan ukuran sampel yang cukup besar sehingga menyebabkan sebaran yang ada mengikuti distribusi normal (Bain & Engelhardt, 1991).

### 3.7.5 Asumsi *Proportional Hazard*

Klein dan Kleinbaum (2005) menjelaskan bahwa ada dua pendekatan untuk memeriksa asumsi proporsional hazard dalam analisis regresi Cox, yaitu dengan pendekatan grafik menggunakan plot log minus-log survival dan menggunakan metode Goodness of Fit.

a. Pendekatan Grafik Log Minus-Log Survival.

Pendekatan grafik ini menggunakan plot log minus-log survival, yang juga dikenal sebagai plot log kumulatif hazard. Grafik ini

didasarkan pada fungsi survival yang diperkirakan menggunakan metode Kaplan-Meier. Fungsi survival Kaplan-Meier dinyatakan sebagai berikut:

$$S(t, X) = S_0(t) \exp(\sum_{i=1}^p \beta_i X_i) \quad (3.18)$$

Dalam fungsi tersebut,  $S(t, X)$  adalah fungsi survival pada waktu  $t$  untuk kelompok dengan nilai peubah penjelas  $X$ .  $S_0(t)$  adalah fungsi survival baseline pada waktu  $t$ . Dalam analisis ini, fungsi log minus-log survival diperlukan untuk memeriksa apakah asumsi proporsional hazard terpenuhi. Fungsi ini dinyatakan sebagai berikut:

$$-\ln S(t, X) = \sum_{i=1}^p \beta_i X_i + \ln[-\ln S_0(t)] \quad (3.19)$$

Jika asumsi proporsional hazard terpenuhi, kurva kelangsungan hidup dari dua kelompok yang berbeda akan sejajar dan paralel. Namun, metode ini memiliki beberapa kelemahan, terutama dalam pengelompokan peubah kontinu, karena jika kelompok menjadi terlalu kecil, perbedaan antara kurva menjadi sulit untuk ditentukan.

b. Metode Goodness of Fit (GOF).

Metode Goodness of Fit menggunakan statistik uji untuk mengevaluasi asumsi proporsional hazard secara lebih objektif daripada pendekatan grafik. Statistik uji yang digunakan dalam metode ini adalah residual Schoenfeld. Residual Schoenfeld merupakan kumpulan nilai untuk setiap individu pada setiap kovariat dalam model regresi Cox proporsional hazard. Nilai residual Schoenfeld untuk kovariat ke- $j$  pada individu ke- $i$  didefinisikan sebagai berikut:

$$r_{ji} = \delta_i (x_{ji} - a_{ji}) \quad (3.20)$$

Di sini,  $\delta_i$  adalah variabel dummy yang menunjukkan status individu (0 jika tersensor, 1 jika terjadi peristiwa),  $x_{ji}$  adalah nilai kovariat ke- $j$  untuk individu ke- $i$ , dan  $a_{ji}$  adalah rata-rata terbobot kovariat ke- $j$  untuk individu pada risiko pada saat waktu  $t_i$ .

Jika asumsi proporsional hazard terpenuhi, residual Schoenfeld untuk setiap kovariat tidak akan berkorelasi dengan peringkat waktu ketahanan. Untuk menguji asumsi ini, dilakukan uji korelasi antara residual Schoenfeld dan peringkat waktu ketahanan. Jika nilai  $p$  dari uji korelasi tersebut lebih kecil dari 0.05, maka dapat disimpulkan bahwa kovariat yang diuji tidak memenuhi asumsi proporsional hazard.

Dalam kesimpulannya, nilai  $p$  yang kurang dari 0.05 menunjukkan bahwa asumsi proporsional hazard tidak terpenuhi untuk kovariat yang diuji.

## BAB 4 METODOLOGI PENELITIAN

### 4.1 Populasi dan Sampel

Populasi dalam penelitian ini adalah seluruh angkatan kerja di Daerah Istimewa Yogyakarta. Sampel yang digunakan adalah pekerja di Daerah Istimewa Yogyakarta yang mulai bekerja dalam waktu 1 tahun terakhir pada bulan Agustus 2022 dan para pencari kerja yang belum mendapatkan pekerjaan, dengan data yang bersumber dari Sakernas Agustus 2022 oleh Badan Pusat Statistik Provinsi Daerah Istimewa Yogyakarta.

### 4.2 Data dan Sumber Data

Penelitian ini menggunakan data sekunder yang didapatkan dari Badan Pusat Statistik Provinsi Daerah Istimewa Yogyakarta dan diambil pada tanggal 27 Januari 2023.

### 4.3 Variabel dan Definisi Operasional

Variabel-variabel yang digunakan dalam penelitian ini adalah sebagai berikut:

**Tabel 4.1** Definisi operasional variabel

Variabel	Definisi	Keterangan
Lama Waktu Mencari Pekerjaan (Y)	Lama waktu mencari pekerjaan adalah waktu yang dihabiskan oleh para pencari kerja dalam melakukan usaha untuk mendapatkan suatu pekerjaan, terhitung sejak responden mulai mencari pekerjaan sampai mendapatkan suatu pekerjaan	Variabel numerik dengan satuan bulan
Status (Penyensoran)	Status dalam hal ini merupakan penyensoran terhadap data lama waktu mencari pekerjaan. Suatu data akan tergolong data tersensor ketika	Variabel kategorik dengan: 0 = data tersensor 1 = data lengkap

	<p>seorang pencari kerja memperoleh pekerjaannya setelah dilakukannya pencacahan oleh BPS. Dengan kata lain, responden tersebut masih dalam proses mencari pekerjaan pada saat BPS melakukan pencacahan untuk kebutuhan data Sakernas Agustus 2022. Apabila seorang responden telah memperoleh pekerjaan pada saat dilakukan pencacahan oleh BPS, maka data responden tersebut akan tergolong sebagai data lengkap (tidak tersensor).</p>	
Jenis Kelamin ( $X_1$ )	Variabel jenis kelamin menyatakan jenis kelamin dari responden yang tergolong sebagai variabel kategorik.	Variabel kategorik dengan: 1 = Laki-laki 2 = Perempuan
Usia ( $X_2$ )	Usia merupakan umur dari responden yang diukur dalam satuan tahun.	<ul style="list-style-type: none"> <li>• Variabel numerik dalam regresi <i>cox proportional hazard</i> dengan satuan tahun.</li> <li>• Variabel kategorik dalam kurva Kaplan-Meier dan Uji <i>Log-Rank</i> dengan: 1 = &lt;20 tahun 2 = 20 – 40 tahun</li> </ul>

		3 = >40 tahun
Hubungan dengan Kepala Rumah Tangga ( $X_3$ )	Hubungan dengan kepala rumah tangga dalam hal ini merupakan hubungan antara responden dengan kepala rumah tangganya.	Variabel kategorik dengan: 1 = kepala rumah tangga 2 = istri/suami 3 = anak 4 = menantu 5 = cucu 6 = orang tua/mertua 7 = lainnya.
Status Perkawinan ( $X_4$ )	Status perkawinan merupakan status perkawinan dari responden.	Variabel kategorik dengan: 1 = belum kawin 2 = kawin 3 = cerai
Tingkat Pendidikan ( $X_5$ )	Tingkat pendidikan terakhir yang telah diselesaikan oleh responden.	Variabel kategorik dengan: 1 = tidak/belum tamat SD 2 = SD/ sederajat 3 = SMP/ sederajat 4 = SMA/ sederajat 5 = Diploma I/II/III 6 = S1/D4 7 = S2
Pengalaman Kerja ( $X_6$ )	Pengalaman kerja menyatakan apakah responden telah memiliki pengalaman bekerja ataupun tidak.	Variabel kategorik dengan: 1 = ya 2 = tidak

Pelatihan Kerja ( $X_7$ )	Pelatihan kerja menyatakan apakah responden telah memiliki pengalaman bekerja ataupun tidak.	Variabel kategorik dengan: 1 = ya 2 = tidak
Klasifikasi Tempat Tinggal ( $X_8$ )	Klasifikasi tempat tinggal menyatakan klasifikasi daerah tempat tinggal dimana menjadi domisili asal responden.	Variabel kategorik dengan: 1 = daerah perkotaan 2 = daerah pedesaan

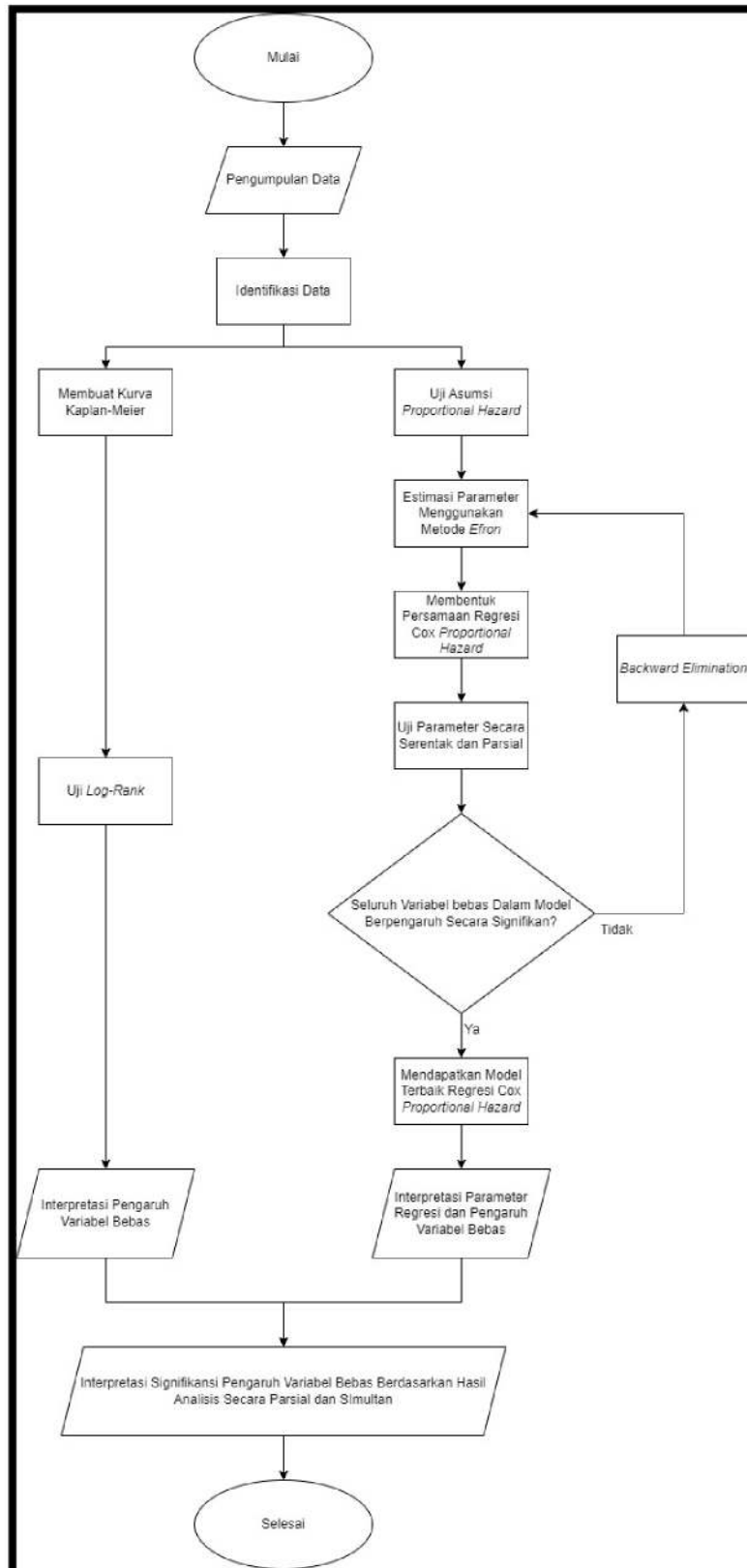
#### 4.4 Metode Analisis Data

Metode yang digunakan dalam menganalisis data penelitian ini adalah menggunakan uji *log-rank* yang disertai dengan kurva kaplan-meier dan regresi cox *proportional hazard* dengan pendekatan metode *efron* untuk mengatasi data yang mengalami kejadian bersama (*ties*). Kedua metode tersebut akan dibandingkan hasilnya untuk melihat faktor-faktor apa saja yang mempengaruhi lama waktu mencari pekerjaan berdasarkan hasil dari diterapkannya kedua metode tersebut.

#### 4.5 Alur Penelitian

Tahapan-tahapan yang dilakukan dalam penelitian ini adalah sebagai berikut:





**Gambar 4.1** Diagram Alir Penelitian

## BAB 5 HASIL DAN PEMBAHASAN

Bab ini akan membahas terkait penerapan analisis regresi cox *proportional hazard* dengan pendekatan *efron* untuk menganalisis faktor-faktor yang mempengaruhi lama waktu mencari pekerjaan secara simultan dan juga penerapan uji *log-rank* beserta kurva kaplan-meier untuk analisis secara parsial. Kedua metode tersebut akan diterapkan untuk melihat bagaimana hasil analisis secara parsial dan secara simultan.

### 5.1 Pengumpulan Data

Data yang digunakan pada penelitian ini adalah data lama waktu mencari pekerjaan di Daerah Istimewa Yogyakarta beserta faktor-faktornya yang dapat mempengaruhi. Data yang digunakan merupakan data sekunder yang didapatkan dari hasil Survei Angkatan Kerja Nasional (Sakernas) Agustus 2022 yang dilakukan oleh Badan Pusat Statistik (BPS) Provinsi Daerah Istimewa Yogyakarta.

### 5.2 Identifikasi Data

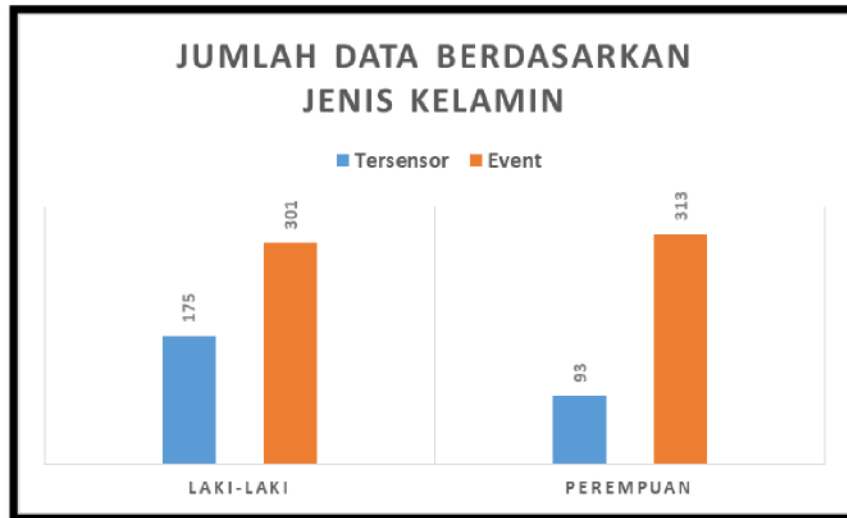
Penelitian ini menggunakan metode analisis survival dengan waktu awal yang diamati (*origin*) adalah waktu dimana responden mulai mencari pekerjaan, dan waktu akhir yang diamati (*event*) adalah waktu dimana responden mendapatkan pekerjaan. Sensor data yang digunakan dalam penelitian ini adalah sensor kanan (*right censor*), dimana suatu data akan tersensor apabila seorang responden belum mendapatkan pekerjaan pada saat dilakukan pencacahan oleh BPS pada bulan Agustus 2022. Dalam penelitian ini, digunakan sebanyak 8 variabel bebas yang dianggap dapat menjadi faktor-faktor yang mempengaruhi lama waktu mencari pekerjaan, di antaranya yaitu jenis kelamin, usia, hubungan dengan kepala rumah tangga, status perkawinan, tingkat pendidikan, pengalaman kerja, pelatihan kerja, dan klasifikasi tempat tinggal. Banyaknya data sampel yang digunakan dalam penelitian ini adalah sebanyak 882 data, dengan pembagian masing-masing variabel bebas berdasarkan status penyensorannya dapat dilihat pada **Tabel 5.1**.

**Tabel 5.1** Jumlah data berdasarkan status penyensoran

Variabel	Tersensor	Event	
<b>Jenis Kelamin</b>	Laki-laki	175	301
	Perempuan	93	313
<b>Usia</b>	20-40 tahun	166	334
	<20 atau >40 tahun	102	280
<b>Hubungan dengan kepala rumah tangga</b>	KRT	90	214
	Istri/suami	30	138
	Anak	128	216
	Menantu	5	17
	Cucu	6	8
	Orang tua/mertua	3	7
	Lainnya	6	14
<b>Status Perkawinan</b>	Belum Kawin	146	220
	Kawin	107	349
	Cerai Hidup/Mati	15	45
<b>Tingkat Pendidikan</b>	Tidak/belum tamat SD	12	41
	SD/Sederajat	19	89
	SMP/Sederajat	32	110
	SMA/SMK/Sederajat	143	283
	Diploma I/II/III	9	19
	S1/D4	50	69
	S2	3	3
<b>Pengalaman Kerja</b>	Ya	164	484
	Tidak	104	130
<b>Pelatihan Kerja</b>	Ya	91	181
	Tidak	177	433
<b>Klasifikasi Tempat Tinggal</b>	Perkotaan	204	378
	Pedesaan	64	236

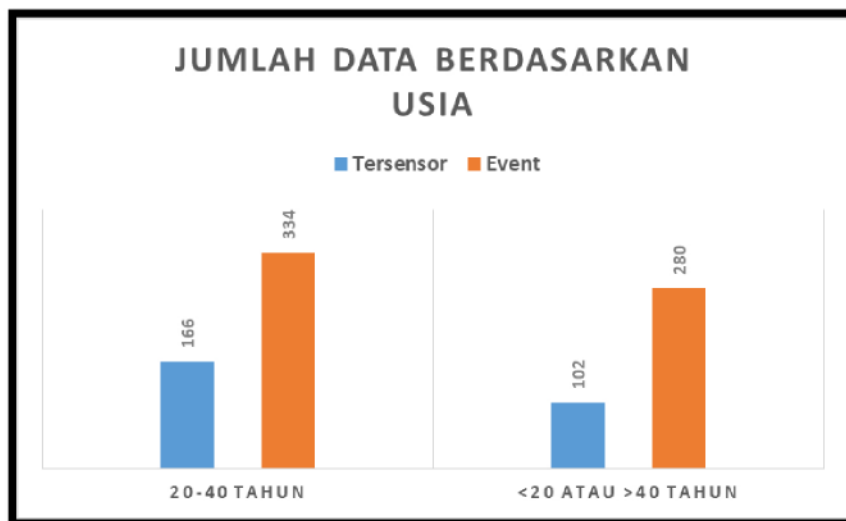
Berdasarkan tabel di atas dapat diketahui persebaran jumlah data dari setiap kategori berdasarkan variabel dan status penyensorannya.

Untuk mempermudah, persebaran jumlah data di atas dapat dibuat visualisasi ke dalam bentuk *bar chart* atau diagram batang.



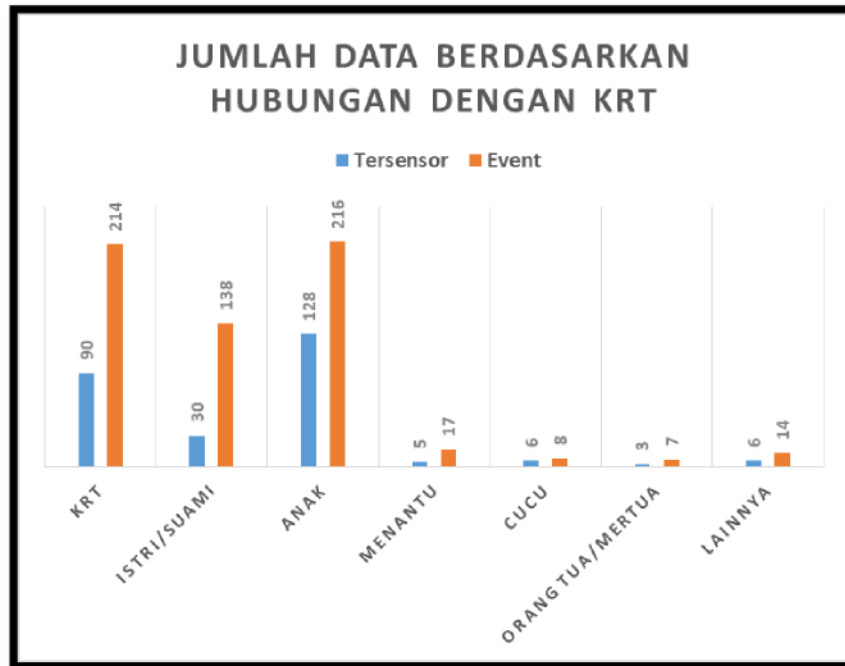
**Gambar 5.1** Jumlah data berdasarkan jenis kelamin

Terlihat bahwa pada kategori laki-laki terdapat sebanyak 175 data tersensor dan 301 data mengalami *event* atau tidak tersensor, sementara pada kategori perempuan terdapat sebanyak 93 data tersensor dan 313 data mengalami *event*.



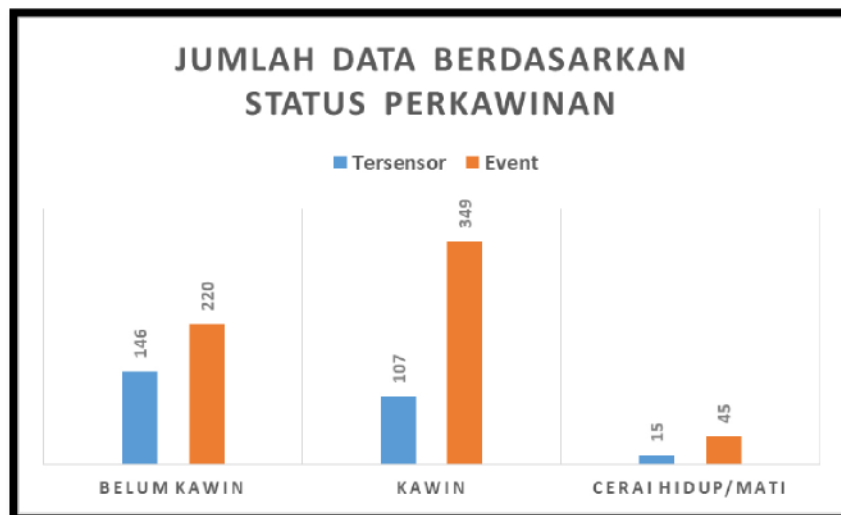
**Gambar 5.2** Jumlah data berdasarkan usia

Terlihat bahwa pada kategori usia 20 – 40 tahun terdapat sebanyak 166 data tersensor dan 334 data mengalami *event* atau tidak tersensor, sementara pada kategori < 20 atau > 40 tahun terdapat sebanyak 102 data tersensor dan 280 data mengalami *event*.



**Gambar 5.3** Jumlah data berdasarkan hubungan dengan KRT

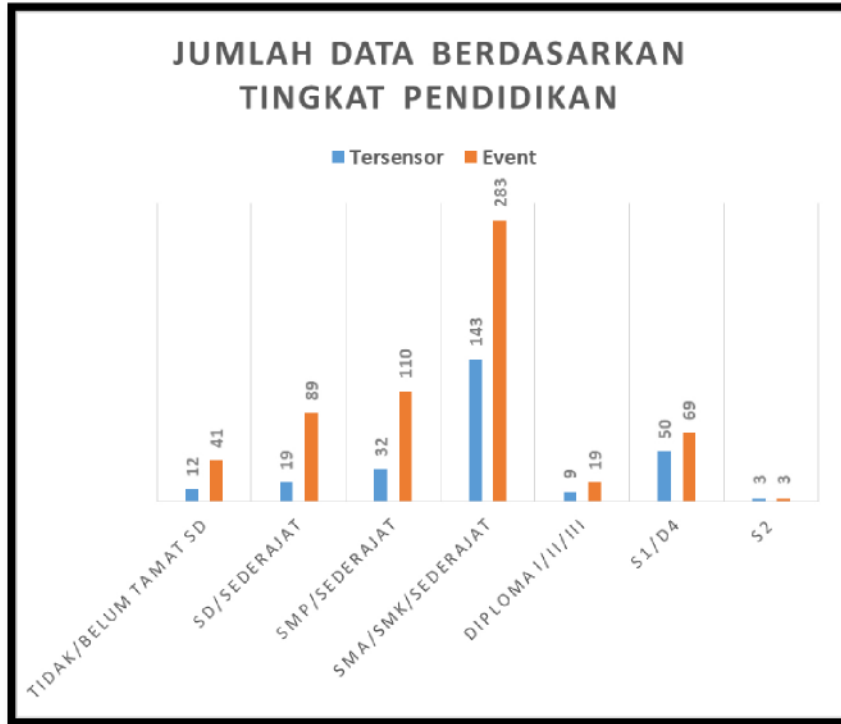
Terlihat bahwa pada kategori kepala rumah tangga (KRT) terdapat sebanyak 90 data tersensor dan 214 data mengalami *event*, pada kategori istri/suami sebanyak 30 data tersensor dan 138 data mengalami *event*, dan seterusnya.



**Gambar 5.4** Jumlah data berdasarkan status perkawinan

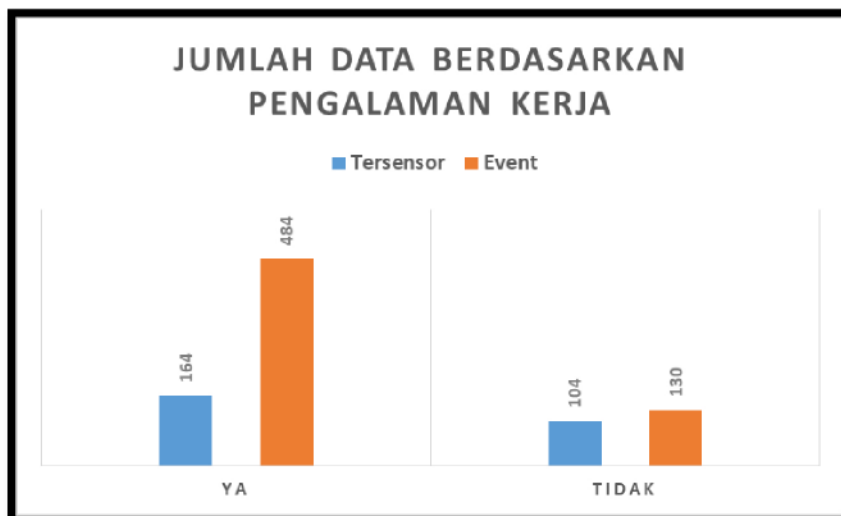
Terlihat bahwa pada kategori belum kawin terdapat sebanyak 146 data tersensor dan 220 data mengalami *event*, pada kategori kawin sebanyak 107 data tersensor

dan 349 data mengalami *event*, dan pada kategori cerai sebanyak 15 data tersensor dan 45 data mengalami *event*.



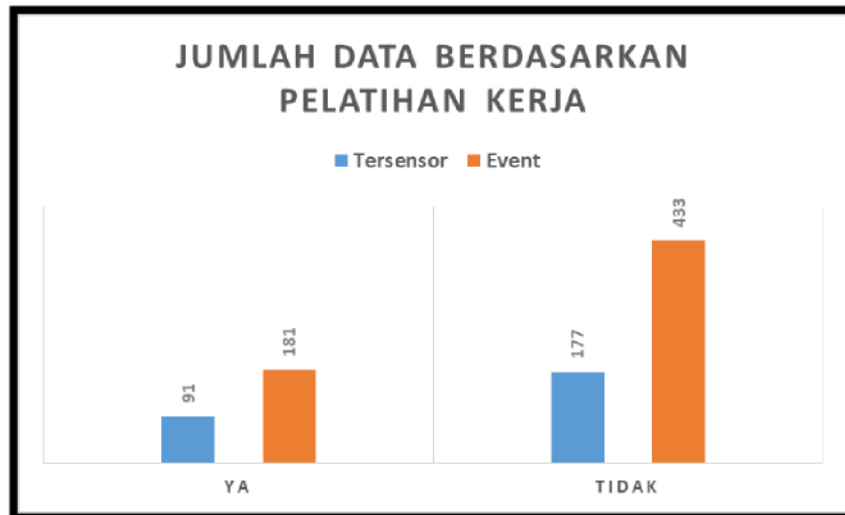
**Gambar 5.5** Jumlah data berdasarkan tingkat pendidikan

Terlihat bahwa pada kategori tidak/belum tamat SD terdapat sebanyak 12 data tersensor dan 41 data mengalami *event*, pada kategori SD/Sederajat sebanyak 19 data tersensor dan 89 data mengalami *event*, dan seterusnya.



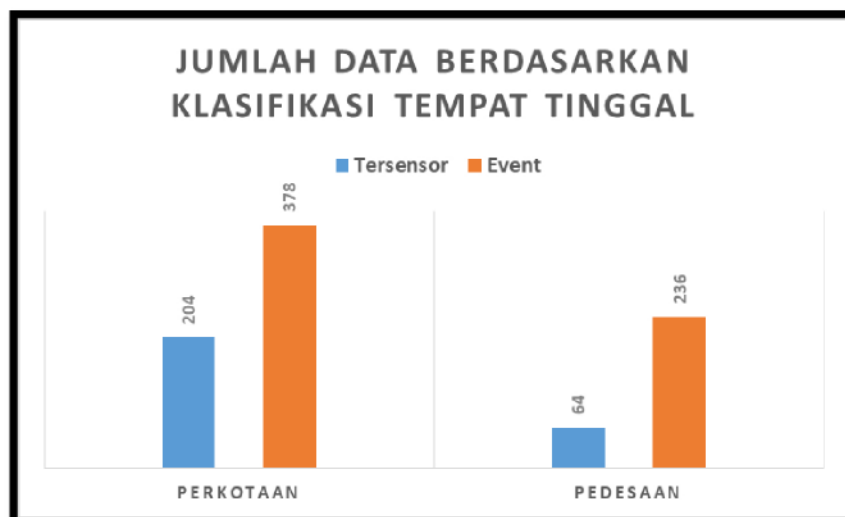
**Gambar 5.6** Jumlah data berdasarkan pengalaman kerja

Terlihat bahwa pada kategori berpengalaman terdapat sebanyak 164 data tersensor dan 484 data mengalami *event*, sementara pada kategori tidak berpengalaman terdapat sebanyak 104 data tersensor dan 130 data mengalami *event*.



**Gambar 5.7** Jumlah data berdasarkan pelatihan kerja

Terlihat bahwa pada kategori melakukan pelatihan terdapat sebanyak 91 data tersensor dan 181 data mengalami *event*, sementara pada kategori tidak melakukan pelatihan terdapat sebanyak 177 data tersensor dan 433 data mengalami *event*.



**Gambar 5.8** Jumlah data berdasarkan tempat tinggal

Terlihat bahwa pada kategori perkotaan terdapat sebanyak 204 data tersensor dan 378 data mengalami *event*, sementara pada kategori pedesaan terdapat sebanyak 64 data tersensor dan 236 data mengalami *event*.

Berdasarkan diagram batang yang terbentuk dari masing-masing variabel bebas, diketahui bahwa data yang ada didominasi oleh data yang mengalami event atau tidak tersensor. Hal ini menunjukkan bahwa sebagian besar individu yang diteliti telah mendapatkan pekerjaan. Lalu hal ini juga akan membuat estimasi yang dilakukan menjadi lebih baik dan akurat.

### **5.3 Kurva Kaplan-Meier dan Uji *Log-Rank***

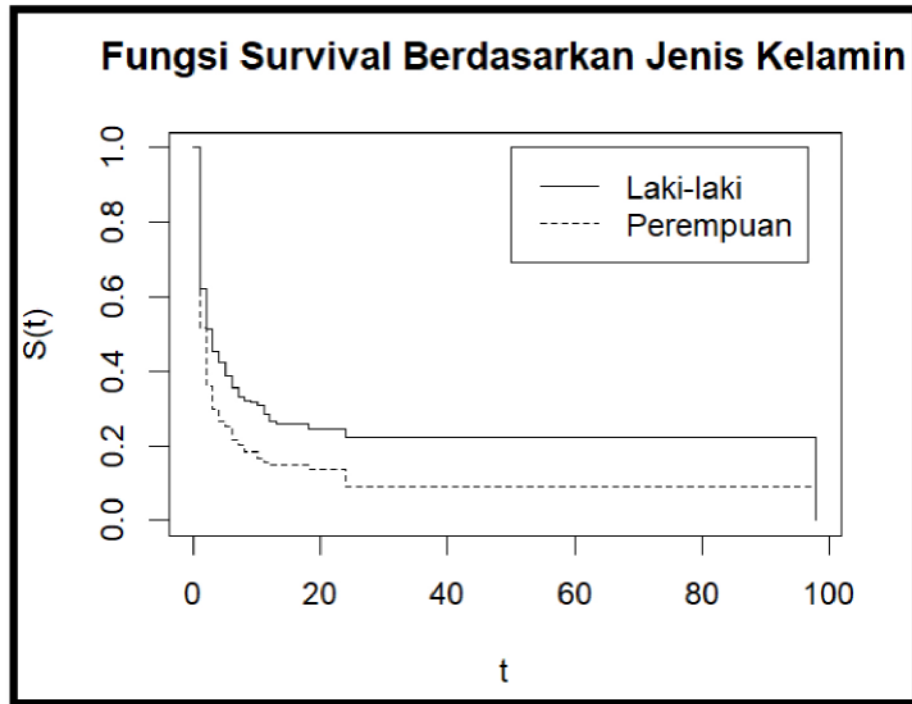
Untuk melakukan analisis pengaruh variabel-variabel bebas secara parsial, dapat digunakan kurva Kaplan-Meier sebagai visualisasi dan uji *log-rank* untuk melakukan uji statistik.

#### **5.3.1 Membuat Kurva Kaplan-Meier dan Melakukan Uji *Log-Rank***

##### **5.3.1.1 Kurva Kaplan-Meier dan Uji *Log-Rank* Variabel $X_1$ (Jenis Kelamin)**

Setelah dilakukan proses pada program *RStudio*, dihasilkan output kurva Kaplan-Meier dari variabel  $X_1$  (jenis kelamin) yang dapat dilihat pada **Gambar 5.9**.





**Gambar 5.9** Kurva Kaplan-Meier variabel  $X_1$

Keterangan:

$t$  : Waktu (dalam bulan)

$S(t)$  : Peluang belum mendapatkan pekerjaan sampai dengan waktu  $t$

Berdasarkan kurva hasil estimasi kaplan-meier, terlihat bahwa antara laki-laki dan perempuan memiliki perbedaan yang signifikan dalam lama waktu yang dibutuhkan untuk mendapatkan pekerjaan. Didapatkan hasil bahwa seorang individu berjenis kelamin perempuan lebih cepat dalam mendapatkan pekerjaan dibanding individu berjenis kelamin laki-laki. Menurut BPS, pada bulan Agustus 2022 tingkat pekerja paruh waktu bagi perempuan lebih tinggi dari laki-laki yaitu sebesar 36,65%, sedangkan tingkat pekerja paruh waktu bagi laki-laki adalah sebesar 22,19%. Data ini menunjukkan bahwa laki-laki yang berperan sebagai tulang punggung keluarga akan berusaha untuk meningkatkan penghasilannya dengan pekerjaan waktu penuh (BPS Provinsi D.I. Yogyakarta, 2022). Hal inilah yang kemudian bisa menyebabkan lama waktu mencari pekerjaan bagi perempuan cenderung lebih cepat daripada laki-laki. Dijelaskan dalam artikel yang dirilis oleh

GoodStats, Partisipasi perempuan yang dominan ini dipengaruhi oleh beragam faktor, termasuk salah satunya adalah keuntungan dari pekerjaan paruh waktu yang memberikan peluang kepada perempuan untuk mencapai keselarasan antara kehidupan profesional dan kehidupan pribadi. Hal ini terutama bermanfaat bagi perempuan yang memiliki tanggung jawab keluarga (Armavillia, 2023).

Untuk memastikan adanya pengaruh jenis kelamin terhadap lama mencari kerja, selanjutnya akan dilakukan uji *log-rank* dengan tahapan uji hipotesis sebagai berikut:

i) Hipotesis

$H_0$ : Tidak ada perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

$H_1$ : Terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

ii) Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

iii) Daerah Kritis

$H_0$  ditolak jika  $\chi_{hitung}^2 \geq \chi_{(0.05;1)}^2$  atau  $p\text{-value} < \alpha = 0.05$

iv) Statistik Uji dan Keputusan

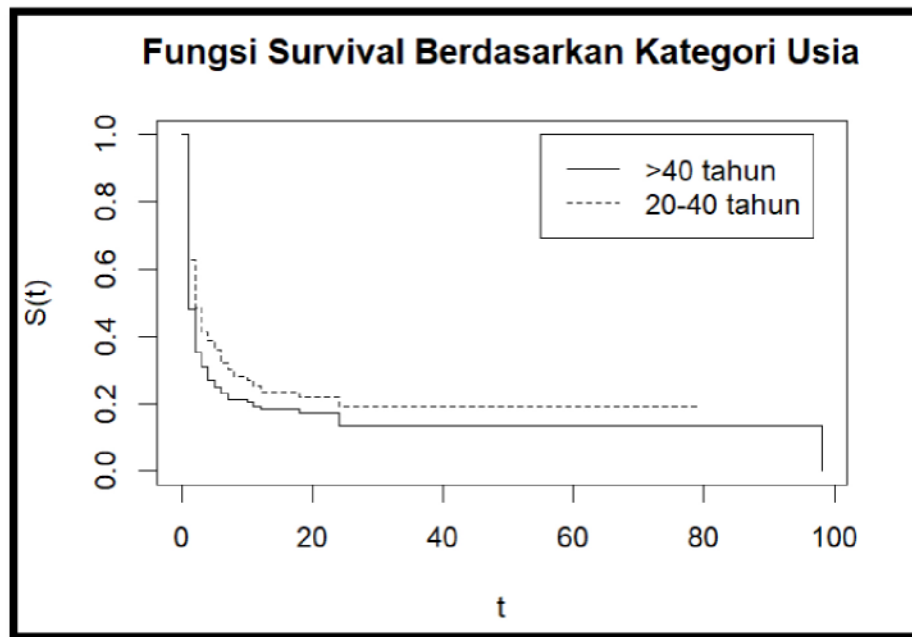
Berdasarkan hasil penghitungan pada program *RStudio*, didapatkan  $\chi_{hitung}^2 = 22.3 > \chi_{(0.05;1)}^2 = 3.8415$  dan  $p\text{-value} = 2 \times 10^{-6} < 0.05$  maka tolak  $H_0$ .

v) Kesimpulan

Dengan tingkat kepercayaan 95%, didapatkan hasil tolak  $H_0$  sehingga dapat dikatakan bahwa terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan. Oleh karena itu dapat disimpulkan bahwa variabel jenis kelamin mempengaruhi lama waktu mencari kerja secara signifikan.

### 5.3.1.2 Kurva Kaplan-Meier dan Uji *Log-Rank* Variabel $X_2$ (Usia)

Setelah dilakukan proses pada program *RStudio*, dihasilkan output kurva Kaplan-Meier dari variabel  $X_2$  (usia) dengan menampilkan dua kategori yang memiliki perbedaan signifikan dan dapat dilihat pada **Gambar 5.10**.



**Gambar 5.10** Kurva Kaplan-Meier variabel  $X_2$

Keterangan:

$t$  : Waktu (dalam bulan)

$S(t)$  : Peluang belum mendapatkan pekerjaan sampai dengan waktu  $t$

Berdasarkan kurva hasil estimasi Kaplan-Meier, terlihat bahwa antara kategori usia memiliki perbedaan yang cukup signifikan dalam lama waktu yang dibutuhkan untuk mendapatkan pekerjaan. Didapatkan hasil bahwa individu berusia  $> 40$  tahun lebih cepat dalam mendapatkan pekerjaan dibanding usia 20 – 40 tahun. Hal ini bisa disebabkan oleh pengalaman kerja yang akan semakin banyak didapatkan oleh seseorang seiring dengan bertambahnya umur. Bagi individu yang memiliki pengalaman kerja, terdapat hubungan negatif antara usia dan lama waktu mencari pekerjaan. Ini berarti semakin tua usia seseorang, semakin singkat waktu yang diperlukan untuk menemukan pekerjaan (Zahra, Fuad, & Dianah, 2021).

Untuk memastikan adanya pengaruh usia terhadap lama mencari kerja, selanjutnya akan dilakukan uji *log-rank* dengan tahapan uji hipotesis sebagai berikut:

i) Hipotesis

$H_0$ : Tidak ada perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

$H_1$ : Terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

ii) Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

iii) Daerah Kritis

$$H_0 \text{ ditolak jika } \chi_{hitung}^2 \geq \chi_{(0.05;2)}^2 \text{ atau } p\text{-value} < \alpha = 0.05$$

iv) Statistik Uji dan Keputusan

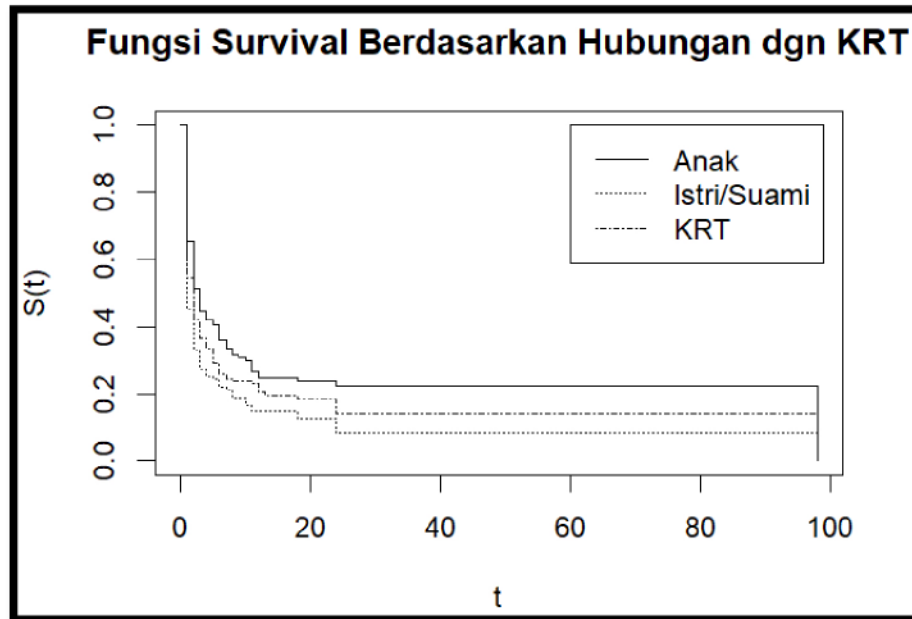
Berdasarkan hasil penghitungan pada program *RStudio*, didapatkan  $\chi_{hitung}^2 = 10.5 \geq \chi_{(0.05;2)}^2 = 5.9915$  dan  $p\text{-value} = 0.005 < 0.05$  maka tolak  $H_0$ .

v) Kesimpulan

Dengan tingkat kepercayaan 95%, didapatkan hasil tolak  $H_0$  sehingga dapat dikatakan bahwa terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan. Oleh karena itu dapat disimpulkan bahwa variabel usia mempengaruhi lama waktu mencari kerja secara signifikan.

### 5.3.1.3 Kurva Kaplan-Meier dan Uji *Log-Rank* Variabel $X_3$ (Hubungan dengan Kepala Rumah Tangga)

Pada variabel  $X_3$  (hubungan dengan kepala rumah tangga) terdapat 3 kategori yang memiliki perbedaan yang signifikan antara satu sama lain. Di antaranya yaitu anak, istri/suami, dan kepala rumah tangga. Plot dari ketiga kategori tersebut digambarkan dengan hasil output oleh program *RStudio* yang dapat dilihat pada **Gambar 5.11**.



**Gambar 5.11** Kurva Kaplan-Meier variabel  $X_3$

Keterangan:

$t$  : Waktu (dalam bulan)

$S(t)$  : Peluang belum mendapatkan pekerjaan sampai dengan waktu  $t$

Berdasarkan kurva hasil estimasi Kaplan-Meier, didapatkan bahwa ketiga kategori tersebut memiliki perbedaan yang cukup signifikan dalam lama waktu yang dibutuhkan untuk mencari pekerjaan. Terlihat bahwa individu yang berstatus sebagai istri/suami dalam rumah tangga memiliki waktu yang cenderung lebih cepat dibanding kedua kategori lainnya, lalu diikuti oleh kategori kepala rumah tangga yang memiliki waktu mendapatkan pekerjaan tercepat kedua, dan kategori anak menjadi kategori yang cenderung paling lambat dalam mendapatkan pekerjaan. Menurut data yang digunakan sebagai sampel dalam penelitian ini, 167 dari 168 individu yang berstatus sebagai seorang istri/suami dalam rumah tangga adalah seorang perempuan. Sehingga penjelasan terkait hasil ini telah dijelaskan pada subsubsubbab **5.3.1.1**.

Untuk memastikan adanya pengaruh status hubungan dengan kepala rumah tangga terhadap lama waktu mencari kerja, selanjutnya akan dilakukan uji *log-rank* dengan tahapan uji hipotesis sebagai berikut:

i) Hipotesis

$H_0$ : Tidak ada perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

$H_1$ : Terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

ii) Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

iii) Daerah Kritis

$H_0$  ditolak jika  $\chi^2_{hitung} \geq \chi^2_{(0.05;6)}$  atau  $p\text{-value} < \alpha = 0.05$

iv) Statistik Uji dan Keputusan

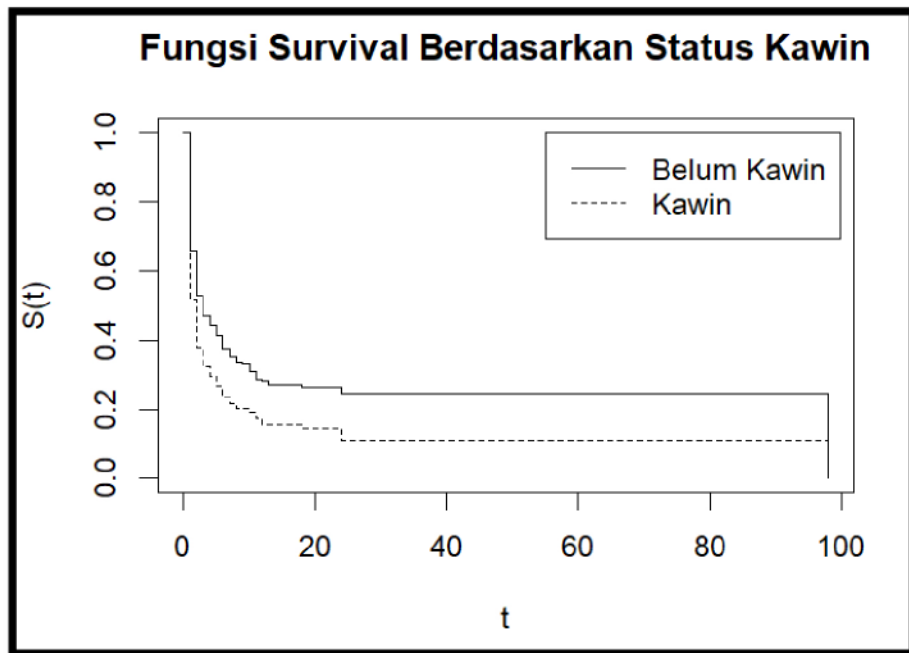
Berdasarkan hasil penghitungan pada program *RStudio*, didapatkan  $\chi^2_{hitung} = 18.4 \geq \chi^2_{(0.05;6)} = 12.5916$  dan  $p\text{-value} = 0.005 < 0.05$  maka tolak  $H_0$ .

v) Kesimpulan

Dengan tingkat kepercayaan 95%, didapatkan hasil tolak  $H_0$  sehingga dapat dikatakan bahwa terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan. Oleh karena itu dapat disimpulkan bahwa variabel hubungan dengan kepala rumah tangga mempengaruhi lama waktu mencari kerja secara signifikan.

#### 5.3.1.4 Kurva Kaplan-Meier dan Uji *Log-Rank* Variabel $X_4$ (Status Perkawinan)

Pada variabel  $X_4$  (status perkawinan) terdapat 2 kategori yang memiliki perbedaan yang signifikan antara satu sama lain. Di antaranya yaitu belum kawin dan kawin. Plot dari kedua kategori tersebut digambarkan dengan hasil output oleh program *RStudio* yang dapat dilihat pada **Gambar 5.12**.



**Gambar 5.12** Kurva Kaplan-Meier variabel  $X_4$

Keterangan:

$t$  : Waktu (dalam bulan)

$S(t)$  : Peluang belum mendapatkan pekerjaan sampai dengan waktu  $t$

Berdasarkan kurva hasil estimasi Kaplan-Meier, didapatkan bahwa kedua kategori tersebut memiliki perbedaan yang signifikan dalam lama waktu yang dibutuhkan untuk mencari pekerjaan. Terlihat bahwa individu yang berstatus sudah kawin memiliki waktu yang lebih cepat dibanding individu yang berstatus belum kawin dalam lama waktu mendapatkan pekerjaan. Menurut hasil penelitian yang dilakukan oleh Khan & Yousaf (2013) dalam (Hartoko, 2019), ditemukan bahwa individu yang telah menikah memiliki lama waktu menganggur yang lebih singkat daripada individu yang belum menikah. Hal ini dikarenakan individu yang telah menikah memiliki tanggung jawab terhadap keluarga mereka, sehingga mendorong mereka untuk menerima pekerjaan meskipun dengan upah yang rendah (Hartoko, 2019).

Hal ini juga sejalan dengan ajaran Islam bahwa dengan menikah, maka pintu-pintu rezeki akan semakin terbuka dan akan dimampukan oleh Allah dengan Karunia-Nya. Seperti yang tertulis dalam dalil pada **Lampiran 6**.

Untuk memastikan adanya status perkawinan terhadap lama waktu mencari kerja, selanjutnya akan dilakukan uji *log-rank* dengan tahapan uji hipotesis sebagai berikut:

i) Hipotesis

$H_0$ : Tidak ada perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

$H_1$ : Terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

ii) Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

iii) Daerah Kritis

$$H_0 \text{ ditolak jika } \chi_{hitung}^2 \geq \chi_{(0.05;2)}^2 \text{ atau } p\text{-value} < \alpha = 0.05$$

iv) Statistik Uji dan Keputusan

Berdasarkan hasil penghitungan pada program *RStudio*, didapatkan  $\chi_{hitung}^2 = 22.4 \geq \chi_{(0.05;2)}^2 = 5.9915$  dan  $p\text{-value} = 1 \times 10^{-5} < 0.05$  maka tolak  $H_0$ .

v) Kesimpulan

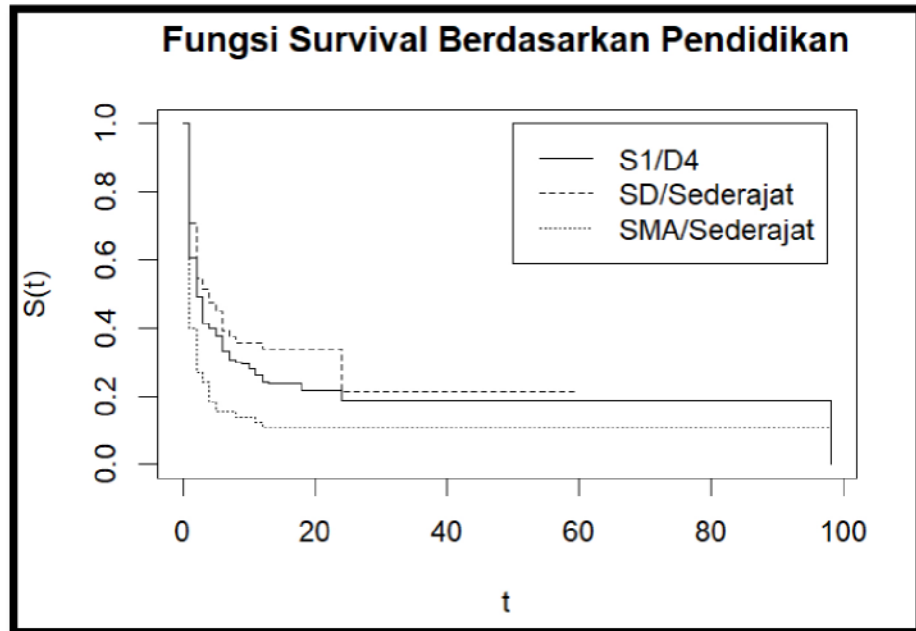
Dengan tingkat kepercayaan 95%, didapatkan hasil tolak  $H_0$  sehingga dapat dikatakan bahwa terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan. Oleh karena itu dapat disimpulkan bahwa variabel status perkawinan mempengaruhi lama waktu mencari kerja secara signifikan.

### **5.3.1.5 Kurva Kaplan-Meier dan Uji Log-Rank Variabel $X_5$ (Tingkat Pendidikan)**

Pada variabel  $X_5$  (tingkat pendidikan) terdapat 3 kategori yang memiliki perbedaan yang signifikan antara satu sama lain. Di antaranya yaitu S1/D4,



SD/Sederajat, dan SMA/Sederajat. Plot dari ketiga kategori tersebut digambarkan dengan hasil output oleh program *RStudio* yang dapat dilihat pada **Gambar 5.13**.



**Gambar 5.13** Kurva Kaplan-Meier variabel  $X_5$

Keterangan:

$t$  : Waktu (dalam bulan)

$S(t)$  : Peluang belum mendapatkan pekerjaan sampai dengan waktu  $t$

Berdasarkan kurva hasil estimasi Kaplan-Meier, didapatkan bahwa ketiga kategori tersebut memiliki perbedaan yang signifikan dalam lama waktu yang dibutuhkan untuk mencari pekerjaan. Terlihat bahwa individu yang memiliki tingkat pendidikan SD/Sederajat memiliki waktu yang cenderung lebih cepat dibanding kedua kategori lainnya, lalu diikuti oleh kategori SMA/Sederajat yang memiliki waktu mendapatkan pekerjaan tercepat kedua, dan kategori S1/D4 menjadi kategori yang cenderung paling lambat dalam mendapatkan pekerjaan. Menurut data yang disajikan oleh BPS D.I. Yogyakarta, penduduk bekerja di DIY didominasi oleh pekerjaan kegiatan informal yaitu sebanyak 53,38% dari penduduk bekerja. Kegiatan informal adalah kegiatan bekerja yang merupakan usaha sendiri, usaha dibantu buruh tidak tetap/tidak dibayar, pekerja bebas, dan pekerja keluarga/tidak

dibayar (BPS Provinsi D.I. Yogyakarta, 2022). Hal inilah yang kemudian bisa membuat rendahnya tingkat pendidikan di DIY justru membuat lama waktu mencari kerja menjadi lebih cepat, karena seorang individu yang berpendidikan rendah akan cenderung mencari atau lebih mudah diterima pada pekerjaan dengan kegiatan informal.

Untuk memastikan adanya pengaruh tingkat pendidikan terhadap lama waktu mencari kerja, selanjutnya akan dilakukan uji *log-rank* dengan tahapan uji hipotesis sebagai berikut:

i) Hipotesis

$H_0$ : Tidak ada perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

$H_1$ : Terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

ii) Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

iii) Daerah Kritis

$$H_0 \text{ ditolak jika } \chi_{hitung}^2 \geq \chi_{(0.05;6)}^2 \text{ atau } p\text{-value} < \alpha = 0.05$$

iv) Statistik Uji dan Keputusan

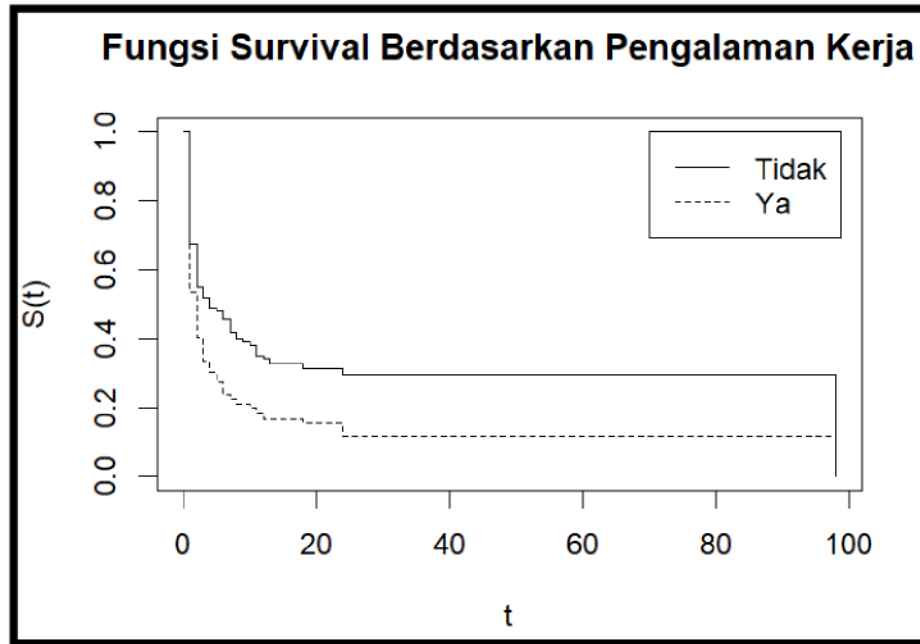
Berdasarkan hasil penghitungan pada program *RStudio*, didapatkan  $\chi_{hitung}^2 = 34.3 \geq \chi_{(0.05;6)}^2 = 12.5916$  dan  $p\text{-value} = 6 \times 10^{-6} < 0.05$  maka tolak  $H_0$ .

v) Kesimpulan

Dengan tingkat kepercayaan 95%, didapatkan hasil tolak  $H_0$  sehingga dapat dikatakan bahwa terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan. Oleh karena itu dapat disimpulkan bahwa variabel tingkat pendidikan mempengaruhi lama waktu mencari kerja secara signifikan.

### 5.3.1.6 Kurva Kaplan-Meier dan Uji *Log-Rank* Variabel $X_6$ (Pengalaman Kerja)

Setelah dilakukan proses pada program *RStudio*, dihasilkan output kurva Kaplan-Meier dari variabel  $X_6$  (pengalaman kerja) yang dapat dilihat pada Gambar 5.14.



Gambar 5.14 Kurva Kaplan-Meier variabel  $X_6$

Keterangan:

$t$  : Waktu (dalam bulan)

$S(t)$  : Peluang belum mendapatkan pekerjaan sampai dengan waktu  $t$

Berdasarkan kurva hasil estimasi kaplan-meier, terlihat bahwa individu yang tidak berpengalaman kerja dan berpengalaman kerja memiliki perbedaan yang signifikan dalam lama waktu yang dibutuhkan untuk mendapatkan pekerjaan. Didapatkan hasil bahwa individu berpengalaman kerja lebih cepat dalam mendapatkan pekerjaan dibanding individu yang tidak berpengalaman kerja.

Untuk memastikan adanya pengaruh pengalaman kerja terhadap lama mencari kerja, selanjutnya akan dilakukan uji *log-rank* dengan tahapan uji hipotesis sebagai berikut:

i) Hipotesis

$H_0$ : Tidak ada perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

$H_1$ : Terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

ii) Tingkat Signifikansi

$$\alpha = 10\% = 0.05$$

iii) Daerah Kritis

$H_0$  ditolak jika  $\chi_{hitung}^2 \geq \chi_{(0.05;1)}^2$  atau  $p\text{-value} < \alpha = 0.05$

iv) Statistik Uji dan Keputusan

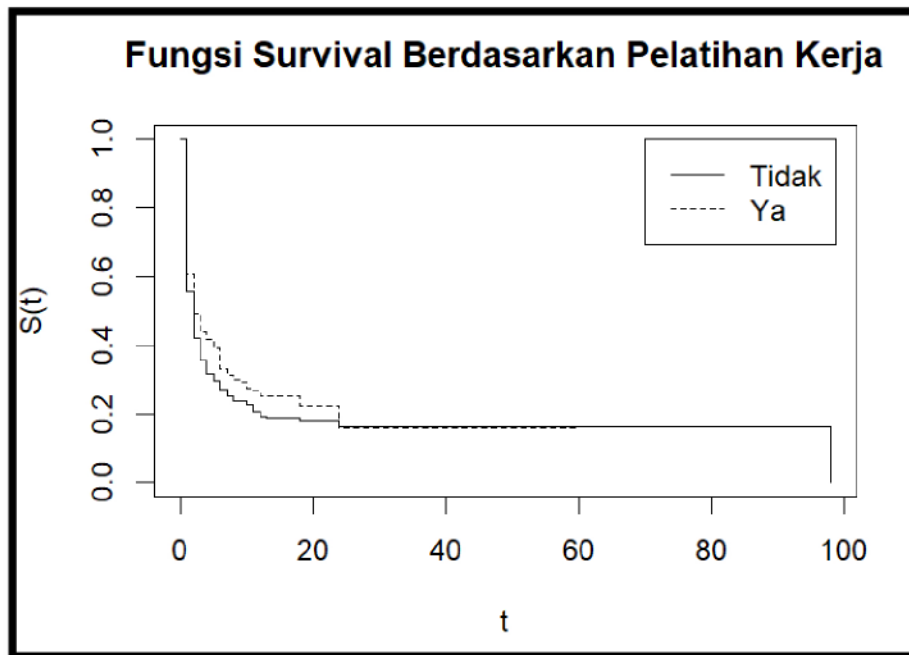
Berdasarkan hasil penghitungan pada program *RStudio*, didapatkan  $\chi_{hitung}^2 = 26.2 \geq \chi_{(0.05;1)}^2 = 3.8415$  dan  $p\text{-value} = 3 \times 10^{-7} < 0.05$  maka tolak  $H_0$ .

v) Kesimpulan

Dengan tingkat kepercayaan 95%, didapatkan hasil tolak  $H_0$  sehingga dapat dikatakan bahwa terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan. Oleh karena itu dapat disimpulkan bahwa variabel pengalaman kerja mempengaruhi lama waktu mencari kerja secara signifikan.

### 5.3.1.7 Kurva Kaplan-Meier dan Uji *Log-Rank* Variabel $X_7$ (Pelatihan Kerja)

Setelah dilakukan proses pada program *RStudio*, dihasilkan output kurva Kaplan-Meier dari variabel  $X_7$  (pelatihan kerja) yang dapat dilihat pada **Gambar 5.15**.



**Gambar 5.15** Kurva Kaplan-Meier variabel  $X_7$

Keterangan:

$t$  : Waktu (dalam bulan)

$S(t)$  : Peluang belum mendapatkan pekerjaan sampai dengan waktu  $t$

Berdasarkan kurva hasil estimasi kaplan-meier, terlihat bahwa kedua kategori yang ada hanya memiliki perbedaan yang tipis dalam hal lama waktu mencari pekerjaan dengan selisih yang tidak terlalu signifikan.

Untuk memastikan adanya pengaruh pelatihan kerja terhadap lama mencari kerja, selanjutnya akan dilakukan uji *log-rank* dengan tahapan uji hipotesis sebagai berikut:

i) Hipotesis

$H_0$ : Tidak ada perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

$H_1$ : Terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

ii) Tingkat Signifikansi

$\alpha = 5\% = 0.05$

iii) Daerah Kritis

$H_0$  ditolak jika  $\chi^2_{hitung} \geq \chi^2_{(0.05;1)}$  atau  $p\text{-value} < \alpha = 0.05$

iv) Statistik Uji dan Keputusan

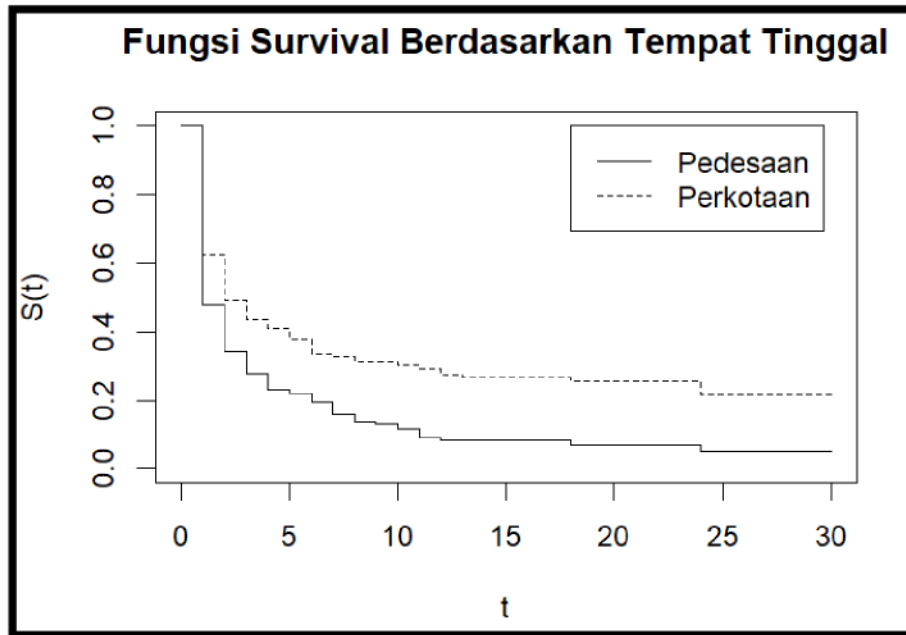
Berdasarkan hasil penghitungan pada program *RStudio*, didapatkan  $\chi^2_{hitung} = 3.4 \geq \chi^2_{(0.05;1)} = 3.8415$  dan  $p\text{-value} = 0.07 > 0.05$  maka gagal tolak  $H_0$ .

v) Kesimpulan

Dengan tingkat kepercayaan 95%, didapatkan hasil gagal tolak  $H_0$  sehingga dapat dikatakan bahwa tidak terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan. Oleh karena itu dapat disimpulkan bahwa variabel pelatihan kerja tidak mempengaruhi lama waktu mencari kerja secara signifikan.

#### **5.3.1.8 Kurva Kaplan-Meier dan Uji *Log-Rank* Variabel $X_8$ (Klasifikasi Tempat Tinggal)**

Setelah dilakukan proses pada program *RStudio*, dihasilkan output kurva Kaplan-Meier dari variabel  $X_8$  (klasifikasi tempat tinggal) yang dapat dilihat pada **Gambar 5.16**.



**Gambar 5.16** Kurva Kaplan-Meier variabel  $X_8$

Keterangan:

$t$  : Waktu (dalam bulan)

$S(t)$  : Peluang belum mendapatkan pekerjaan sampai dengan waktu  $t$

Berdasarkan kurva hasil estimasi kaplan-meier, terlihat bahwa antara individu yang tinggal di pedesaan dan perkotaan memiliki perbedaan yang signifikan dalam hal lama waktu yang dibutuhkan untuk mendapatkan pekerjaan. Didapatkan hasil bahwa individu yang tinggal di pedesaan lebih cepat dalam mendapatkan pekerjaan dibanding individu yang tinggal di perkotaan. Menurut Badan Pusat Statistik (BPS), fenomena ini terjadi karena daerah perkotaan memiliki lebih banyak sektor formal dibandingkan dengan daerah pedesaan. Seperti yang kita tahu, sektor formal lebih sulit diakses oleh angkatan kerja karena membutuhkan keterampilan khusus atau persyaratan tertentu dibandingkan dengan sektor informal. Terlebih lagi, hal ini juga disebabkan oleh kenyataan bahwa penduduk pedesaan umumnya tidak sangat selektif dalam memilih pekerjaan, sehingga mereka melakukan berbagai jenis kegiatan, baik sebagai pekerja keluarga maupun pekerja bebas di bidang pertanian (BPS Provinsi D.I. Yogyakarta, 2022).

Untuk memastikan adanya pengaruh klasifikasi tempat tinggal terhadap lama mencari kerja, selanjutnya akan dilakukan uji *log-rank* dengan tahapan uji hipotesis sebagai berikut:

i) Hipotesis

$H_0$ : Tidak ada perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

$H_1$ : Terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

ii) Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

iii) Daerah Kritis

$H_0$  ditolak jika  $\chi_{hitung}^2 \geq \chi_{(0.05;1)}^2$  atau  $p\text{-value} < \alpha = 0.05$

iv) Statistik Uji dan Keputusan

Berdasarkan hasil penghitungan pada program *RStudio*, didapatkan  $\chi_{hitung}^2 = 33.8 \geq \chi_{(0.05;1)}^2 = 3.8415$  dan  $p\text{-value} = 6 \times 10^{-9} < 0.05$  maka tolak  $H_0$ .

v) Kesimpulan

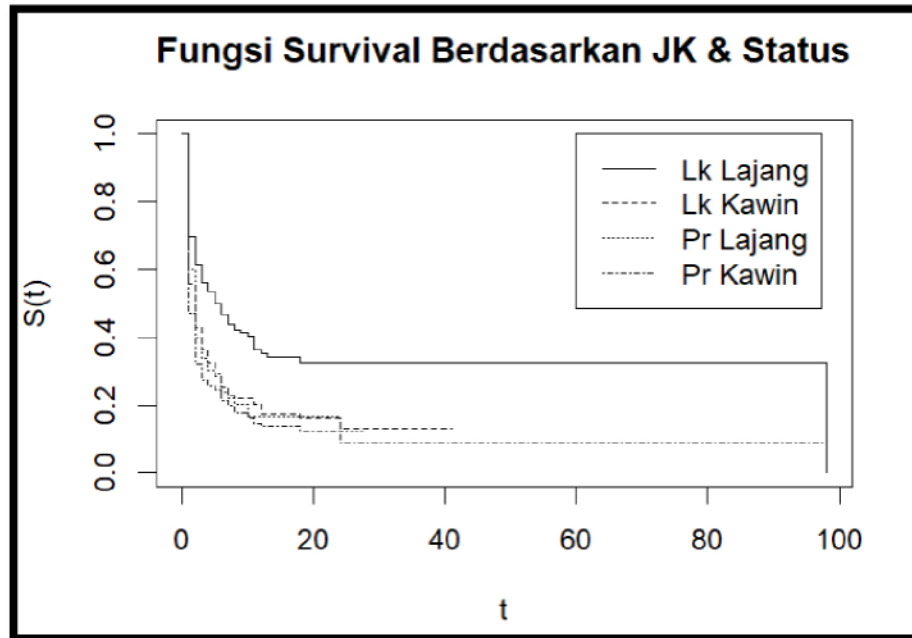
Dengan tingkat kepercayaan 95%, didapatkan hasil tolak  $H_0$  sehingga dapat dikatakan bahwa terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan. Oleh karena itu dapat disimpulkan bahwa variabel klasifikasi tempat tinggal mempengaruhi lama waktu mencari kerja secara signifikan.

### **5.3.1.9 Kurva Kaplan-Meier dan Uji Log-Rank Kombinasi Variabel $X_1$ (Jenis Kelamin) dan $X_4$ (Status Perkawinan)**

Setelah melihat hasil kurva kaplan-meier dan uji *log-rank* dari variabel jenis kelamin dan status perkawinan, akan menarik untuk melihat bagaimana hasil analisis dari gabungan kedua variabel ini. Kurva kaplan-meier perbandingan fungsi survival berdasarkan jenis kelamin dan status perkawinan dihasilkan setelah



melakukan proses pada program *Rstudio* yang hasilnya dapat dilihat pada **Gambar 5.17**.



**Gambar 5.17** Kurva kaplan-meier variabel gabungan  $X_1$  dan  $X_4$

Keterangan:

$t$  : Waktu (dalam bulan)

$S(t)$  : Peluang belum mendapatkan pekerjaan sampai dengan waktu  $t$

Berdasarkan kurva kaplan-meier yang dihasilkan, didapatkan bahwa jenis kelamin dan status perkawinan ternyata sangat mempengaruhi lama waktu mencari pekerjaan. Terlihat bahwa individu sebagai seorang laki-laki yang belum menikah cenderung akan memiliki lama waktu mencari pekerjaan yang jauh lebih lama dibandingkan dengan laki-laki yang sudah menikah dan perempuan baik yang belum maupun sudah menikah. Sementara itu pada ketiga kategori yaitu laki-laki sudah menikah, perempuan belum menikah, dan perempuan sudah menikah terlihat tidak memiliki perbedaan yang terlalu signifikan. Untuk memastikan pengaruh dari variabel ini terhadap lama waktu mencari kerja, akan dilakukan uji *log-rank* dengan langkah-langkah sebagai berikut:

- i) Hipotesis

$H_0$ : Tidak ada perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

$H_1$ : Terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan.

ii) Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

iii) Daerah Kritis

$$H_0 \text{ ditolak jika } \chi_{hitung}^2 \geq \chi_{(0.05;5)}^2 \text{ atau } p\text{-value} < \alpha = 0.05$$

iv) Statistik Uji dan Keputusan

Berdasarkan hasil penghitungan pada program *RStudio*, didapatkan  $\chi_{hitung}^2 = 44.9 \geq \chi_{(0.05;5)}^2 = 11.0705$  dan  $p\text{-value} = 2 \times 10^{-8} < 0.05$  maka tolak  $H_0$ .

v) Kesimpulan

Dengan tingkat kepercayaan 95%, didapatkan hasil tolak  $H_0$  sehingga dapat dikatakan bahwa terdapat perbedaan yang signifikan dalam fungsi survival di antara kelompok-kelompok yang dibandingkan. Oleh karena itu dapat disimpulkan bahwa jenis kelamin dan status perkawinan mempengaruhi lama waktu mencari kerja secara signifikan.

### 5.3.2 Interpretasi Pengaruh Variabel Bebas Berdasarkan Kurva Kaplan-Meier dan Uji *Log-Rank*

Setelah dilakukan pembuatan kurva Kaplan-Meier dan uji *log-rank*, telah diketahui signifikansi dari variabel-variabel bebas yang digunakan berdasarkan hasil uji *log-rank* seperti yang dapat dilihat pada **Tabel 5.2**.

**Tabel 5.2** Interpretasi Pengaruh Variabel Bebas Berdasarkan Uji *Log-Rank*

Variabel	<i>p-value</i>	Keputusan	Kesimpulan
$X_1$ (jenis kelamin)	$2 \times 10^{-6}$	Tolak $H_0$	Variabel mempengaruhi secara signifikan
$X_2$ (usia)	0.005	Tolak $H_0$	Variabel mempengaruhi secara signifikan

Variabel	<i>p-value</i>	Keputusan	Kesimpulan
$X_3$ (hub. dgn KRT)	0.005	Tolak $H_0$	Variabel mempengaruhi secara signifikan
$X_4$ (status kawin)	$1 \times 10^{-5}$	Tolak $H_0$	Variabel mempengaruhi secara signifikan
$X_5$ (pendidikan)	$6 \times 10^{-6}$	Tolak $H_0$	Variabel mempengaruhi secara signifikan
$X_6$ (pengalaman kerja)	$3 \times 10^{-7}$	Tolak $H_0$	Variabel mempengaruhi secara signifikan
$X_7$ (pelatihan kerja)	0.07	Gagal Tolak $H_0$	Variabel tidak mempengaruhi secara signifikan
$X_8$ (tempat tinggal)	$6 \times 10^{-9}$	Tolak $H_0$	Variabel mempengaruhi secara signifikan

#### 5.4 Regresi Cox *Proportional Hazard*

Setelah melakukan analisis secara parsial, dalam subbab ini akan dilakukan penerapan regresi *cox proportional hazard* dengan metode *efron* pada data yang digunakan dalam penelitian ini. regresi *cox proportional hazard* bertujuan untuk melihat faktor-faktor yang mempengaruhi lama waktu mencari kerja berdasarkan hasil analisis secara simultan, sekaligus membentuk model regresi yang dapat digunakan untuk menghitung *hazard rate* seseorang mendapatkan pekerjaan pada suatu waktu tertentu berdasarkan faktor-faktor yang diamati.

##### 5.4.1 Uji Asumsi *Proportional Hazard*

Sebelum melakukan estimasi parameter regresi *cox proportional hazard*, asumsi *proportional hazard* harus terpenuhi terlebih dahulu untuk setiap variabel bebas yang akan digunakan sebagai syarat untuk dapat dilakukannya regresi *cox PH*. Berikut adalah rangkaian uji hipotesis yang dilakukan:

i) Hipotesis

$H_0: \rho = 0$  (asumsi *proportional hazard* terpenuhi).

$H_1: \rho \neq 0$  (asumsi *proportional hazard* tidak terpenuhi).

ii) Tingkat Signifikansi

$\alpha = 5\% = 0.05$

iii) Daerah Kritis

$H_0$  ditolak jika  $p\text{-value} < \alpha = 0.05$

iv) Statistik Uji dan Keputusan

Berdasarkan perhitungan yang dilakukan menggunakan program *RStudio*, dihasilkan output hasil uji yang terlihat pada **Tabel 5.3**.

**Tabel 5.3** Hasil uji asumsi *proportional hazard*

Variabel	chisq	df	p-value	Keputusan
X1	0.0231	1	0.88	Gagal Tolak $H_0$
X2	2.1409	1	0.14	Gagal Tolak $H_0$
X3	6.2567	6	0.4	Gagal Tolak $H_0$
X4	0.4596	2	0.79	Gagal Tolak $H_0$
X5	7.7175	6	0.26	Gagal Tolak $H_0$
X6	0.1337	1	0.71	Gagal Tolak $H_0$
X7	0.0428	1	0.84	Gagal Tolak $H_0$
X8	2.2766	1	0.13	Gagal Tolak $H_0$
GLOBAL	21.9118	19	0.29	Gagal Tolak $H_0$

v) Kesimpulan

Dengan tingkat kepercayaan 95%, didapatkan hasil bahwa seluruh variabel bebas memenuhi asumsi proporsionalitas baik secara keseluruhan (global) maupun secara parsial.

**5.4.2 Estimasi Parameter**

Estimasi parameter regresi *cox proportional hazard* dilakukan dengan penghitungan menggunakan program *RStudio* dan menghasilkan output dengan hasil seperti terlihat pada **Tabel 5.4**.

**Tabel 5.4** Estimasi parameter model awal

Variabel	Coef	Exp(coef)	Se(coef)	Z	P-value	Keputusan
X1Perempuan	0.572615	1.772897	0.115208	4.97	6.69E-07	Tolak $H_0$
X2	0.00452	1.004531	0.004714	0.959	0.33757	Gagal tolak $H_0$
X3Anak	-0.020641	0.979571	0.155481	-0.133	0.89439	Gagal tolak $H_0$
X3Cucu	-0.076901	0.925981	0.393657	-0.195	0.84512	Gagal tolak $H_0$
X3Istri/suami	-0.368495	0.691775	0.156004	-2.362	0.01817	Tolak $H_0$
X3Lainnya	0.02733	1.027706	0.302369	0.09	0.92798	Gagal tolak $H_0$
X3Menantu	-0.164624	0.848213	0.275164	-0.598	0.54966	Gagal tolak $H_0$
X3Orang tua/mertua	-0.62235	0.536682	0.412089	-1.51	0.13098	Gagal tolak $H_0$
X4Cerai Hidup/Mati	-0.100822	0.904094	0.212215	-0.475	0.63472	Gagal tolak $H_0$
X4Kawin	0.190819	1.210241	0.153301	1.245	0.21323	Gagal tolak $H_0$
X5Diploma I/II/III	-0.153765	0.857474	0.239518	-0.642	0.52089	Gagal tolak $H_0$
X5S1/D4	-0.311552	0.732309	0.14382	-2.166	0.03029	Tolak $H_0$
X5S2	-0.177582	0.837293	0.587133	-0.302	0.7623	Gagal tolak $H_0$
X5SD/Sederajat	0.260055	1.297001	0.14243	1.826	0.06787	Gagal tolak $H_0$
X5SMP/Sederajat	0.253145	1.288071	0.119821	2.113	0.03463	Tolak $H_0$

Variabel	Coef	Exp(coef)	Se(coef)	Z	P-value	Keputusan
X5Tidak/belum tamat SD	0.027055	1.027424	0.203838	0.133	0.89441	Gagal tolak H0
X6Ya	0.324336	1.383112	0.111878	2.899	0.00374	Tolak H0
X7Ya	-0.055005	0.94648	0.094125	-0.584	0.55896	Gagal tolak H0
X8Perkotaan	-0.364375	0.69463	0.089186	-4.086	4.40E-05	Tolak H0

Berdasarkan output hasil estimasi pada **Tabel 5.4**, dapat dituliskan model persamaan regresi cox sebagai berikut:

$$\begin{aligned}
h(t|x) = h_0 \exp & \left( 0.5726X_1 (\text{perempuan}) + 0.0045X_2 - 0.0206X_3 (\text{anak}) - \right. \\
& 0.0769X_3 (\text{cucu}) - 0.3685X_3 (\text{istri|suami}) + 0.0273X_3 (\text{lainnya}) - \\
& 0.1646X_3 (\text{menantu}) - 0.6223X_3 (\text{orang tua|mertua}) - 0.1008X_4 (\text{cerai}) + \\
& 0.1908X_4 (\text{kawin}) - 0.1538X_5 (\text{Diploma}) - 0.3116X_5 (\text{S1|D4}) - 0.1776X_5 (\text{S2}) + \\
& 0.2601X_5 (\text{SD|sederajat}) + 0.2531X_5 (\text{SMP|sederajat}) + \\
& 0.0271X_5 (\text{tidak|belum tamat SD}) + 0.3243X_6 (\text{ya}) - 0.055X_7 (\text{ya}) - \\
& \left. 0.3644X_8 (\text{perkotaan}) \right) \tag{5.1}
\end{aligned}$$

Estimasi ini menghasilkan nilai *Concordance* sebesar 0.65, yang artinya model memiliki kemampuan yang cukup baik dalam membandingkan urutan waktu kejadian di antara pasangan observasi.

### 5.4.3 Uji Signifikansi Parameter

Untuk melihat apakah variabel-variabel bebas dalam model memiliki pengaruh yang signifikan terhadap persamaan regresi cox yang telah terbentuk, perlu untuk dilakukan uji signifikansi baik secara serentak maupun parsial. Uji signifikansi yang dilakukan di antaranya:

1. Uji secara serentak.

Untuk mengetahui apakah seluruh variabel bebas secara bersama-sama signifikan terhadap model, dilakukan tahapan uji sebagai berikut:

- i) Hipotesis

$H_0: \beta_1 = \beta_2 = \dots = \beta_j = 0$  (variabel bebas secara bersama-sama tidak signifikan dalam mempengaruhi model).

$H_0$ : terdapat setidaknya satu  $\beta_i \neq 0$ , dengan  $i = 1, 2, \dots, j$  (variabel bebas secara bersama-sama signifikan dalam mempengaruhi model).

ii) Tingkat Signifikansi

$$\alpha = 10\% = 0.1$$

iii) Daerah Kritis

$$H_0 \text{ ditolak jika } G \geq \chi^2_{(0.1;19)} \text{ atau } p\text{-value} < \alpha = 0.1$$

iv) Statistik Uji dan Keputusan

Berdasarkan output pada program *RStudio*, didapatkan *likelihood ratio* ( $G$ ) sebesar  $104.5 \geq \chi^2_{(0.05;19)} = 30.1435$  dengan  $p\text{-value} = 8 \times 10^{-14} < \alpha = 0.05$  maka tolak  $H_0$ .

v) Kesimpulan

Dengan tingkat kepercayaan 90%, didapatkan bahwa seluruh variabel bebas secara bersama-sama signifikan dalam mempengaruhi model.

2. Uji secara parsial.

Setelah mendapatkan hasil bahwa seluruh variabel bebas signifikan dalam mempengaruhi model secara bersama-sama, penting untuk melakukan uji secara parsial untuk mengetahui signifikansi dari masing-masing variabel dalam mempengaruhi model.

i) Hipotesis

$H_0: \beta_i = 0$  (variabel bebas  $\beta_i$  tidak signifikan dalam model).

$H_0: \beta_i \neq 0$  (variabel bebas  $\beta_i$  signifikan dalam model).

ii) Tingkat Signifikansi

$$\alpha = 10\% = 0.1$$

iii) Daerah Kritis

$$H_0 \text{ ditolak jika } p\text{-value} < \alpha = 0.1$$

iv) Statistik Uji dan Keputusan

Berdasarkan perhitungan yang dilakukan menggunakan program *RStudio*, dihasilkan output hasil uji yang terlihat pada **Tabel 5.4**.

**Tabel 5.5** Hasil uji parsial

<b>Variabel</b>	<b>Z</b>	<b>P-value</b>	<b>Keputusan</b>
X1Perempuan	4.97	6.69E-07	Tolak H0
X2	0.959	0.33757	Gagal tolak H0
X3Anak	-0.133	0.89439	Gagal tolak H0
X3Cucu	-0.195	0.84512	Gagal tolak H0
X3Istri/suami	-2.362	0.01817	Tolak H0
X3Lainnya	0.09	0.92798	Gagal tolak H0
X3Menantu	-0.598	0.54966	Gagal tolak H0
X3Orang tua/mertua	-1.51	0.13098	Gagal tolak H0
X4Cerai Hidup/Mati	-0.475	0.63472	Gagal tolak H0
X4Kawin	1.245	0.21323	Gagal tolak H0
X5Diploma I/II/III	-0.642	0.52089	Gagal tolak H0
X5S1/D4	-2.166	0.03029	Tolak H0
X5S2	-0.302	0.7623	Gagal tolak H0
X5SD/Sederajat	1.826	0.06787	Gagal tolak H0
X5SMP/Sederajat	2.113	0.03463	Tolak H0
X5Tidak/belum tamat SD	0.133	0.89441	Gagal tolak H0
X6Ya	2.899	0.00374	Tolak H0
X7Ya	-0.584	0.55896	Gagal tolak H0
X8Perkotaan	-4.086	4.40E-05	Tolak H0

v) Kesimpulan

Dengan tingkat kepercayaan 95%, didapatkan bahwa variabel bebas  $X_1$ ,  $X_3$ ,  $X_5$ ,  $X_6$ , dan  $X_8$  masing-masing memiliki kategori yang signifikan terhadap model, sehingga keempat variabel tersebut dapat dianggap mempengaruhi model secara signifikan. Sementara variabel  $X_2$ ,  $X_4$ , dan  $X_7$  tidak memiliki kategori yang signifikan terhadap model, sehingga keempat variabel tersebut tidak mempengaruhi model secara signifikan.

**5.4.4 Pembentukan Model Alternatif**

Setelah membentuk model regresi cox seperti pada subsubbab 5.4.3, didapatkan bahwa tidak seluruh variabel bebas mempengaruhi model secara signifikan, sehingga dapat dibentuk model alternatif yang hanya berisi variabel-variabel bebas yang berpengaruh secara signifikan. Untuk itu perlu dilakukan eliminasi terhadap variabel bebas yang tidak mempengaruhi secara signifikan

dengan menggunakan metode *backwards*. Setelah dilakukan penghitungan dengan menggunakan program *RStudio*, didapatkan output hasil estimasi model terbaik seperti terlihat pada **Tabel 5.6**.

**Tabel 5.6** Hasil estimasi model alternatif

Variabel	Coef	Exp(coef)	Se(coef)	Z	P-value	Keputusan
X1Perempuan	0.32637	1.38593	0.08123	4.018	5.88E-05	Tolak H0
X6Ya	0.4721	1.60336	0.09936	4.751	2.02E-06	Tolak H0
X8Perkotaan	-0.47772	0.6202	0.08394	-5.691	1.26E-08	Tolak H0

Sehingga dapat dituliskan sebuah persamaan model terbaik regresi cox yang terbentuk adalah sebagai berikut:

$$h(t|x) = h_0 \exp(0.3264X_{1(\text{perempuan})} + 0.4721X_{6(\text{ya})} - 0.4777X_{8(\text{perkotaan})}) \quad (5.2)$$

Estimasi ini menghasilkan nilai *Concordance* sebesar 0.619, yang artinya model memiliki kemampuan yang masih cukup baik dalam membandingkan urutan waktu kejadian di antara pasangan observasi.

#### 5.4.5 Interpretasi Parameter Regresi dan Pengaruh Variabel Bebas

Setelah mendapatkan model terbaik, selanjutnya akan diinterpretasikan parameter regresi beserta pengaruhnya terhadap variabel bebas.

Berdasarkan pada persamaan 5.2, telah diketahui bahwa variabel-variabel bebas yang mempengaruhi lama waktu mencari kerja berdasarkan hasil analisis secara simultan dengan metode regresi cox PH di antaranya yaitu  $X_1$  (jenis kelamin),  $X_5$  (tingkat pendidikan),  $X_6$  (pengalaman kerja), dan  $X_8$  (klasifikasi tempat tinggal). Kemudian parameter regresi yang dihasilkan dapat diinterpretasikan sebagai berikut:

- a. Dengan jenis kelamin laki-laki sebagai referensi, didapatkan hasil bahwa  $X_{1(\text{perempuan})}$  yang mewakili individu sebagai seseorang berjenis kelamin perempuan berpengaruh secara positif terhadap nilai *hazard* untuk mendapatkan pekerjaan. Dimana individu yang berjenis kelamin perempuan memiliki kesempatan untuk mendapatkan



pekerjaan yang lebih besar yaitu 1.44822 kali dibandingkan individu berjenis kelamin laki-laki.

- b. Dengan tingkat pendidikan SMA/Sederajat sebagai referensi, didapatkan hasil seperti terlihat pada **Tabel 5.7**.

**Tabel 5.7** Interpretasi Parameter Variabel  $X_5$

Kategori	Pengaruh terhadap <i>hazard</i>	Rasio <i>hazard</i> (dibanding tingkat pendidikan SMA)
Diploma I/II/III	Negatif	0.89551 (lebih kecil)
S1/D4	Negatif	0.773 (lebih kecil)
S2	Negatif	0.95482 (lebih kecil)
SD/sederajat	Positif	1.41579 (lebih besar)
SMP/Sederajat	Positif	1.30468 (lebih besar)
Tidak/belum tamat SD	Positif	1.06382 (lebih besar)

- c. Dengan tidak berpengalaman bekerja sebagai referensi, didapatkan hasil bahwa  $X_6$  (*ya*) yang mewakili sebagai individu berpengalaman kerja memiliki pengaruh positif terhadap nilai *hazard* dalam mendapatkan pekerjaan. Dimana individu yang memiliki pengalaman kerja memiliki kesempatan lebih besar untuk mendapatkan pekerjaan yaitu 1.48927 kali dibandingkan individu yang tidak memiliki pengalaman kerja.
- d. Dengan tempat tinggal pedesaan sebagai referensi, didapatkan hasil bahwa  $X_8$  (*perkotaan*) yang mewakili sebagai individu yang tinggal di kawasan perkotaan memiliki pengaruh negatif terhadap nilai *hazard* dalam mendapatkan pekerjaan. Dimana individu yang tinggal di perkotaan memiliki kesempatan lebih kecil untuk mendapatkan pekerjaan yaitu 0.68139 kali dibandingkan individu yang tinggal di pedesaan.

### 5.5 Signifikansi Pengaruh Variabel Bebas Berdasarkan Hasil Uji *Log-Rank* dan Regresi *Cox Proportional Hazard*

Uji *log-rank* digunakan untuk melakukan analisis variabel bebas secara parsial, sedangkan regresi *cox proportional hazard* digunakan untuk melakukan analisis pengaruh variabel bebas secara simultan. Hasil dari kedua metode tersebut dalam mengidentifikasi pengaruh dari variabel bebas dapat dilihat pada **Tabel 5.8**.

**Tabel 5.8** Hasil identifikasi pengaruh variabel bebas

Variabel	Signifikan Berdasarkan Metode	
	Uji <i>Log-Rank</i>	Regresi Cox PH
Jenis Kelamin ( $X_1$ )	√	√
Usia ( $X_2$ )	√	X
Hubungan dengan KRT ( $X_3$ )	√	X
Status Perkawinan ( $X_4$ )	√	X
Tingkat Pendidikan ( $X_5$ )	√	√
Pengalaman Kerja ( $X_6$ )	√	√
Pelatihan Kerja ( $X_7$ )	X	X
Klasifikasi Tempat Tinggal ( $X_8$ )	√	√

Berdasarkan pada **Tabel 5.8**, didapatkan bahwa hanya variabel pelatihan kerja ( $X_7$ ) yang tidak mempengaruhi lama waktu mencari kerja secara signifikan berdasarkan analisis secara parsial melalui kurva kaplan-meier dan uji *log-rank*. Sedangkan hasil analisis secara simultan yang dilakukan melalui regresi *cox PH* menunjukkan bahwa hanya empat dari delapan variabel bebas yang mempengaruhi lama waktu mencari kerja secara signifikan, yaitu di antaranya jenis kelamin ( $X_1$ ), tingkat pendidikan ( $X_5$ ), pengalaman kerja ( $X_6$ ), dan klasifikasi tempat tinggal ( $X_8$ ). Hal ini menandakan bahwa analisis secara parsial lebih baik dalam mengidentifikasi pengaruh dari masing-masing variabel bebas terhadap variabel terikat dibanding analisis secara simultan. Hal ini bisa disebabkan oleh adanya variabel-variabel yang lebih menonjol dalam mempengaruhi model regresi *cox proportional hazard*, sehingga pengaruh dari variabel bebas lainnya menjadi kurang signifikan dalam model.

## BAB 6 PENUTUP

### 6.1 Kesimpulan

Berdasarkan hasil dari analisis dan pembahasan yang telah dilakukan, kesimpulan yang dapat diambil adalah:

1. Berdasarkan hasil diagram batang dari semua variabel bebas yang digunakan, didapatkan bahwa jumlah data yang digunakan sebagai sampel dalam penelitian ini didominasi oleh data yang mengalami *event* (tidak tersensor). Maka dapat disimpulkan bahwa sebagian besar individu yang diteliti telah mendapatkan pekerjaan.
2. Berdasarkan model akhir persamaan regresi cox yang dihasilkan, diketahui bahwa kesempatan bagi seorang individu untuk mendapatkan pekerjaan sebagai perempuan lebih besar yaitu sebesar 1.3859 kali dari individu sebagai laki-laki. Kemudian dalam hal pengalaman kerja, seorang individu yang memiliki pengalaman kerja memiliki kesempatan yang lebih besar dalam mendapatkan pekerjaan yakni sebesar 1.6034 kali dibanding individu yang tidak memiliki pengalaman kerja. Lalu dalam hal tempat tinggal, seorang individu yang tinggal di wilayah perkotaan memiliki kesempatan mendapatkan pekerjaan yang lebih rendah yakni 0.6202 kali dibanding individu yang tinggal di pedesaan.
3. Hasil analisis secara parsial menggunakan kurva kaplan-meier dan uji *log-rank* menunjukkan bahwa tujuh dari delapan faktor-faktor yang dilibatkan dalam penelitian ini mempengaruhi lama waktu mencari kerja secara signifikan, yaitu meliputi jenis kelamin, usia, hubungan dengan kepala rumah tangga, status perkawinan, tingkat pendidikan, pengalaman kerja, dan klasifikasi tempat tinggal. Sementara berdasarkan hasil analisis secara simultan menggunakan regresi cox *proportional hazard*, dihasilkan bahwa faktor-faktor yang mempengaruhi lama waktu mencari kerja secara signifikan adalah jenis kelamin, pengalaman kerja, dan klasifikasi tempat tinggal. Hal ini mengindikasikan bahwa analisis secara parsial dapat berperan lebih baik dalam mengidentifikasi adanya

pengaruh dari masing-masing variabel bebas terhadap variabel terikat. Hal ini bisa disebabkan oleh adanya variabel-variabel bebas yang lebih menonjol dalam mempengaruhi model sehingga pengaruh dari variabel-variabel bebas lainnya menjadi kurang signifikan terhadap model.

## **6.2 Saran**

Setelah melakukan penelitian ini, peneliti memiliki beberapa saran yang dapat diberikan, yaitu di antaranya:

1. Jika terdapat penelitian serupa di kemudian hari, peneliti menyarankan kepada peneliti selanjutnya untuk mengutamakan penggunaan metode parametrik dibanding semiparametrik maupun nonparametrik. Sebab penggunaan metode parametrik akan menghasilkan kemampuan prediksi yang lebih baik, serta lebih efisien dalam melakukan estimasi dan perkiraannya lebih akurat.
2. Pemerintah daerah dapat membuat kebijakan yang dapat membantu para pencari kerja untuk mendapatkan pekerjaan dengan mempermudah pencari kerja mendapatkan pengalaman kerja, yaitu dengan cara membuat program magang pemerintah, atau menjalin kerjasama dengan perusahaan-perusahaan swasta untuk lebih membuka peluang magang.

## DAFTAR PUSTAKA

- Abdullah, S. (2022). *Analisis Survival: Konsep dan Aplikasi dengan R*. Jakarta Timur: Bumi Aksara.
- Adyaksa, F. F. (2020). Analisis Faktor-faktor yang Mempengaruhi Pengangguran Terdidik di Indonesia Tahun 2018.
- Armavillia, K. E. (2023, Juli 26). *Pekerja Paruh Waktu Di Indonesia Didominasi Oleh Perempuan, Mengapa?* Diambil kembali dari GoodStats: <https://goodstats.id/article/pekerja-paruh-waktu-di-indonesia-didominasi-oleh-perempuan-u5Sow>
- Arrozi, F., & Sutrisna, K. (2018). ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI LAMA Mencari Kerja Bagi Tenaga Kerja Terdidik di Kota Denpasar. *E-JURNAL EKONOMI PEMBANGUNAN UNIVERSITAS UDAYANA*, Vol. 7. No. 12 Desember 2018, 2732-2763.
- Astuti, M. (2013). Faktor-faktor yang Mempengaruhi Lama Menganggur Bagi Pekerja di Industri Perkebunan Kelapa Sawit Ogan Komering Ulu. *Jurnal Ilmiah STIE MDP Vol. 2 No. 2 Maret 2013*.
- Badan Pusat Statistik. (2022). *Keadaan Angkatan Kerja di Indonesia Agustus 2022*. Jakarta: Badan Pusat Statistik.
- Bain, L. J., & Engelhardt, M. (1991). *Introduction to Probability and Mathematical Statistics Second Edition*. California: Belmont Duxbury.
- BPS Provinsi D.I. Yogyakarta. (2020). *Statistik Ketenagakerjaan Daerah Istimewa Yogyakarta 2019-2020*. Yogyakarta: BPS Provinsi D.I. Yogyakarta.
- BPS Provinsi D.I. Yogyakarta. (2022). *Keadaan Ketenagakerjaan D.I. Yogyakarta Agustus 2022*. Yogyakarta: BPS Provinsi D.I. Yogyakarta.
- Chandra, N. E., & Rohmaniah, S. A. (2019). Analisis Survival Model Regresi Semiparametrik Pada Lama Studi Mahasiswa. *Jurnal Ilmiah Teknosains*, Vol. V No.2 November 2019, 94-98.
- Collet, D. (2003). *Text in Statistical Science: Modelling Survival Data in Medical Research Second Edition*. USA: Chapman & Hall.

- Dr. Ikhlasiah Dalimoenthe, M. (2020). *Sosiologi Gender*. Jakarta Timur: Bumi Aksara.
- Fatimah, F. (2021). *Analisis dan Interpretasi Data Riset*. Tangerang Selatan: Universitas Terbuka.
- Harlan, J. (2017). *Analisis Survival*. Depok: Gunadarma.
- Hartoko, Y. (2019). PENGARUH PENDIDIKAN, PELATIHAN, JENIS KELAMIN, UMUR, STATUS PERKAWINAN, DAN DAERAH TEMPAT TINGGAL TERHADAP LAMA MENCARI KERJA TENAGA KERJA TERDIDIK DI INDONESIA. *Jurnal Pendidikan dan Ekonomi, Volume 8, Nomor 3, Tahun 2019*, 201-207.
- Iskandar, B. M., Kusumawati, R., & Subekti, R. (2015). Model Cox Proportional Hazard pada Kejadian Bersama. *Jurnal UNY. Vol. 4(2)*, 23-31.
- Juharsah, Madi, R. A., Hartini, Hamid, W., Yusuf, & Hastuti. (2020). Pelatihan Pengolahan Data Untuk Analisis Regresi Menggunakan Software Statistical Product And Service Solution (SPSS). *Jurnal Pengabdian Masyarakat Ilmu Terapan, Vol. 2, No. 1, April 2020*, 128.
- Kaplan, E. L., & Meier, P. (1958). Nonparametric Estimation from Incomplete Observation. *Journal of The American Statistical Association*, 457-481.
- Klein, J. P., & Moeschberger, M. L. (2003). *Survival Analysis Techniques for Censored and Truncated Data Second Edition*. New York: Springer-Verlag.
- Kleinbaum, D. G., & Klein, M. (2005). *Survival Analysis: A Self-Learning Text*. New York: Springer.
- Kleinbaum, D. G., & Klein, M. (2005). *Survival Analysis: A Self-Learning Text Second Edition*. New York: Springer.
- Kurniawan, R., & Yuniarto, B. (2016). *Analisis Regresi: Dasar dan Penerapannya dengan R Edisi Pertama*. Jakarta: KENCANA.
- Martias, L. D. (2021). Statistika Deskriptif Sebagai Kumpulan Informasi. *FIHRIS: Jurnal Ilmu Perpustakaan dan Informasi Vol. 16, No. 1, Januari-Juni 2021*, 40-59.
- Nuswantari, D. (1998). *Kamus Kedokteran Dorland edisi 25*. Jakarta: EGC.

- Pratomo, D. S. (2017). Fenomena Pengangguran Terdidik di Indonesia.
- Priyono, J., & Yasin, M. (2016). Analisis usia, gaji dan beban tanggungan terhadap produksi home industry sepatu di Sidoarjo (Studi Kasus di Kecamatan Krian). *Jurnal Ekonomi dan Bisnis. 1 (1)*, 95-120.
- Rahmadeni, & Ranti, S. (2016). Perbandingan Model Cox Menggunakan Estimasi Efron Partial Likelihood dan Breslow Partial Likelihood. *Seminar Nasional Teknologi Informasi Komunikasi dan Industri*.
- Riangkaryaman. (2017). Model Semiparametrik Cox PH dan Parametrik (Weibull dan Lognormal) dalam Analisis Survival. *Skripsi*.
- Riyandianci, N. (2017). Analisis Survival pada Pasien Penderita Kanker Serviks di RSUD Dr. Soetomo Surabaya Menggunakan Stratified Cox dan Extended Cox. *Tugas Akhir*.
- Salsabila, A. (2020). Analisis Survival Terhadap Pasien Lupus Menggunakan Metode Kaplan Meier dan Uji Log Rank. *Tugas Akhir*.
- Siddiq, F., & Dokhi, M. (2022). Survival Analisis Durasi Menganggur Angkatan Kerja Disabilitas yang Mengalami Berhenti Bekerja Akibat Pandemi Covid-19. *Jurnal Statistika dan Aplikasinya Vol. 6 No. 2, December 2022*, 326-340.
- Simanjuntak, P. J. (1998). *Pengantar Ekonomi Sumber Daya Manusia*. Jakarta: FEUI.
- Sudana, I. A., Suciptawati, N. P., & Harini, L. I. (2013). PENERAPAN REGRESI COX PROPORTIONAL HAZARD UNTUK MENDUGA FAKTOR-FAKTOR YANG MEMENGARUHI LAMA MENCARI KERJA. *E-Jurnal Matematika Vol. 2, No.3, Agustus 2013*, 7-10.
- Suhardjo, D. (2007). *Definisi Tingkat Pendidikan*. Bandung: Refika Aditama.
- Susenati, M. N. (2015). Analisis Lama Waktu Mencari Kerja dengan Pendekatan Regresi Cox Proportional Hazard. *Tugas Akhir*.
- Yulianti, d. (2011). Pengaruh Karakteristik Kependudukan Terhadap Pengangguran Di Sumatera Barat. *Jurnal Manajemen dan Kewirausahaan, Vol. 2, No. 2, Mei 2011*.

Zahra, Fuad, Z., & Dianah, A. (2021). Analisis Faktor yang Mempengaruhi Lama Mencari Kerja Dalam Perspektif Ekonomi Islam. *JIMEBIS – Scientific Journal of Students Islamic Economics and Business Volume 2 Nomor 1 Mei 2021*, 1-7.





## Lampiran 2 Sintaks Regresi Cox PH

```
# uji asumsi proportional hazard
```

```
cox.zph(coxph(Surv(Y, Status) ~ X1+X2+X3+X4+X5+X6+X7+X8, data =  
carikerja))
```

```
# regresi cox ph
```

```
m.cox1<-coxph(Surv(Y,Status)~X1+X2+X3+X4+X5+X6+X7+X8,data =  
carikerja, method = "efron")
```

```
summary(m.cox1)
```

```
m.cox2<-coxph(Surv(Y,Status)~X1+X2+X3+X4+X5+X6+X8,data = carikerja,  
method = "efron")
```

```
summary(m.cox2)
```

```
m.cox3<-coxph(Surv(Y,Status)~X1+X3+X4+X5+X6+X8,data = carikerja,  
method = "efron")
```

```
summary(m.cox3)
```

```
m.cox4<-coxph(Surv(Y,Status)~X1+X3+X5+X6+X8,data = carikerja, method =  
"efron")
```

```
summary(m.cox4)
```

```
m.cox5<-coxph(Surv(Y,Status)~X1+X5+X6+X8,data = carikerja, method =  
"efron")
```

```
summary(m.cox5)
```

```
m.cox6<-coxph(Surv(Y,Status)~X1+X6+X8,data = carikerja, method =  
"efron")
```

```
summary(m.cox6)
```

### Lampiran 3 Sintaks Kurva Kaplan-Meier & Uji *Log-Rank*

```
#kaplan-meier & uji logrank
```

```
#X1
```

```
KM_JK=survfit(Surv(Y,Status)~X1,data=carikerja)
```

```
summary(KM_JK)
```

```
plot(KM_JK[1],conf.int=F, main = "Fungsi Survival Berdasarkan Jenis  
Kelamin", xlab = "t", ylab = "S(t)") #Laki-laki
```

```
lines(KM_JK[2],conf.int=F,lty=2)
```

```
legend(50,1,lty=c(1,2),c("Laki-laki","Perempuan"))
```

```
survdiff(Surv(Y,Status)~X1,data=carikerja)
```

```
#X2
```

```
KM_AGE=survfit(Surv(Y,Status)~X2a,data=carikerja)
```

```
summary(KM_AGE)
```

```
plot(KM_AGE[1],conf.int=F, main = "Fungsi Survival Berdasarkan Kategori  
Usia", xlab = "t", ylab = "S(t)") #<20 atau >40
```

```
lines(KM_AGE[2],conf.int=F,lty=2)
```

```
legend(50,1,lty=c(1,2),c("<20 atau >40","20-40 tahun"))
```

```
survdiff(Surv(Y,Status)~X2a,data=carikerja)
```

```
#X3
```

```
KM_HKRT=survfit(Surv(Y,Status)~X3,data=carikerja)
```

```
summary(KM_HKRT)
```

```

plot(KM_HKRT[1],conf.int=F, main = "Fungsi Survival Berdasarkan
Hubungan dgn KRT", xlab = "t", ylab = "S(t)") #KRT

lines(KM_HKRT[2],conf.int=F,lty=2) #Anak

lines(KM_HKRT[4],conf.int=F,lty=3) #Istri/suami

legend(60,1,lty=c(1,2,3),c("KRT","Anak","Istri/suami"))

survdif(Surv(Y,Status)~X3,data=carikerja)

#X4

KM_MRG=survfit(Surv(Y,Status)~X4,data=carikerja)

summary(KM_MRG)

plot(KM_MRG[1],conf.int=F, main = "Fungsi Survival Berdasarkan Status
Kawin", xlab = "t", ylab = "S(t)") #Belum Kawin

lines(KM_MRG[3],conf.int=F,lty=2) #Kawin

legend(55,1,lty=c(1,2),c("Belum Kawin","Kawin"))

survdif(Surv(Y,Status)~X4,data=carikerja)

#X5

KM_EDU=survfit(Surv(Y,Status)~X5,data=carikerja)

summary(KM_EDU)

plot(KM_EDU[1],conf.int=F,lty=1, main = "Fungsi Survival Berdasarkan
Pendidikan", xlab = "t", ylab = "S(t)") #SMA/Sederajat

lines(KM_EDU[3],conf.int=F,lty=2) #S1/D4

lines(KM_EDU[5],conf.int=F,lty=3) #SD/Sederajat

legend(50,1,lty=c(1,2,3),c("S1/D4","SD/Sederajat","SMA/Sederajat"))

```

```
survdif(Surv(Y,Status)~X5,data=carikerja)
```

```
#X6
```

```
KM_WORK=survfit(Surv(Y,Status)~X6,data=carikerja)
```

```
summary(KM_WORK)
```

```
plot(KM_WORK[1],conf.int=F, main = "Fungsi Survival Berdasarkan  
Pengalaman Kerja", xlab = "t", ylab = "S(t)") #Tidak
```

```
lines(KM_WORK[2],conf.int=F,lty=2)
```

```
legend(70,1,lty=c(1,2),c("Tidak","Ya"))
```

```
survdif(Surv(Y,Status)~X6,data=carikerja)
```

```
#X7
```

```
KM_TRA=survfit(Surv(Y,Status)~X7,data=carikerja)
```

```
summary(KM_TRA)
```

```
plot(KM_TRA[1],conf.int=F, main = "Fungsi Survival Berdasarkan Pelatihan  
Kerja", xlab = "t", ylab = "S(t)") #Tidak
```

```
lines(KM_TRA[2],conf.int=F,lty=2)
```

```
legend(70,1,lty=c(1,2),c("Tidak","Ya"))
```

```
survdif(Surv(Y,Status)~X7,data=carikerja)
```

```
#X8
```

```
KM_LIF=survfit(Surv(Y,Status)~X8,data=carikerja)
```

```
summary(KM_LIF)
```

```

plot(KM_LIF[1],conf.int=F, main = "Fungsi Survival Berdasarkan Tempat
Tinggal", xlab = "t", ylab = "S(t)") #Pedesaan

lines(KM_LIF[2],conf.int=F,lty=2)

legend(18,1,lty=c(1,2),c("Pedesaan","Perkotaan"))

survdif(Surv(Y,Status)~X8,data=carikerja)

#X1.X4

KM_X1.X4=survfit(Surv(Y,Status)~X1.X4,data=carikerja)

summary(KM_X1.X4)

plot(KM_X1.X4[1],conf.int=F, main = "Fungsi Survival Berdasarkan JK &
Status", xlab = "t", ylab = "S(t)")

lines(KM_X1.X4[3],conf.int=F,lty=2)

lines(KM_X1.X4[4],conf.int=F,lty=3)

lines(KM_X1.X4[6],conf.int=F,lty=4)

legend(60,1,lty=c(1,2,3,4),c("Lk Lajang","Lk Kawin","Pr Lajang","Pr Kawin"))

survdif(Surv(Y,Status)~X1.X4,data=carikerja)

```

#### **Lampiran 4** Output Regresi Cox PH

```

> m.cox6<-coxph(Surv(Y,Status)~X1+X6+X8,data = carikerja,
method = "efron")

> summary(m.cox6)

Call:
coxph(formula = Surv(Y, Status) ~ X1 + X6 + X8, data =
carikerja,
      method = "efron")

n= 882, number of events= 614

```

	coef	exp(coef)	se(coef)	z	Pr(> z )	
X1Perempuan	0.32637	1.38593	0.08123	4.018	5.88e-05	***
X6Ya	0.47210	1.60336	0.09936	4.751	2.02e-06	***
X8Perkotaan	-0.47772	0.62020	0.08394	-5.691	1.26e-08	***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
X1Perempuan	1.3859	0.7215	1.1819	1.6251
X6Ya	1.6034	0.6237	1.3196	1.9481
X8Perkotaan	0.6202	1.6124	0.5261	0.7311

Concordance= 0.619 (se = 0.015 )

Likelihood ratio test= 75.54 on 3 df, p=3e-16

Wald test = 75.16 on 3 df, p=3e-16

Score (logrank) test = 76.5 on 3 df, p=<2e-16

**Lampiran 5** Output Uji *Log-Rank*

```
> survdiff(Surv(Y,Status)~X1,data=carikerja)
```

Call:

```
survdiff(formula = Surv(Y, Status) ~ X1, data = carikerja)
```

	N	Observed	Expected	(O-E)^2/E	(O-E)^2/V
X1=Laki-laki	476	301	348	6.46	22.3
X1=Perempuan	406	313	266	8.48	22.3

Chisq= 22.3 on 1 degrees of freedom, p= 2e-06

```
> survdiff(Surv(Y,Status)~X2b,data=carikerja)
```

Call:

```
survdiff(formula = Surv(Y, Status) ~ X2b, data = carikerja)
```

	N	Observed	Expected	(O-E) <sup>2</sup> /E	(O-E) <sup>2</sup> /V
X2b=<20 tahun	86	56	55.8	0.000472	0.000799
X2b=>40 tahun	296	224	194.3	4.543908	9.991603
X2b=20-40 tahun	500	334	363.9	2.452769	9.038747

Chisq= 10.5 on 2 degrees of freedom, p= 0.005

```
> survdiff(Surv(Y,Status)~X3,data=carikerja)
```

Call:

```
survdiff(formula = Surv(Y, Status) ~ X3, data = carikerja)
```

	N	Observed	Expected	(O-E) <sup>2</sup> /E	(O-E) <sup>2</sup> /V
X3=A. KRT	304	214	204.06	0.4839	1.0790
X3=Anak	344	216	252.82	5.3624	13.4659
X3=Cucu	14	8	8.63	0.0453	0.0696
X3=Istri/suami	168	138	111.82	6.1273	11.2259
X3=Lainnya	20	14	13.57	0.0134	0.0202
X3=Menantu	22	17	15.17	0.2217	0.3355
X3=Orang tua/mertua	10	7	7.93	0.1087	0.1580

Chisq= 18.4 on 6 degrees of freedom, p= 0.005

```
> survdiff(Surv(Y,Status)~X4,data=carikerja)
```



Call:

```
survdiff(formula = Surv(Y, Status) ~ X4, data = carikerja)
```

	N	Observed	Expected	(O-E) <sup>2</sup> /E	(O-E) <sup>2</sup> /V
X4=Belum kawin	366	220	267.7	8.510	22.326
X4=Cerai Hidup/Mati	60	45	40.9	0.404	0.638
X4=Kawin	456	349	305.3	6.245	18.421

Chisq= 22.4 on 2 degrees of freedom, p= 1e-05

```
> survdiff(Surv(Y,Status)~X5,data=carikerja)
```

Call:

```
survdiff(formula = Surv(Y, Status) ~ X5, data = carikerja)
```

	N	Observed	Expected	(O-E) <sup>2</sup> /E	(O-E) <sup>2</sup> /V
X5=A. SMA/Sederajat	426	283	308.26	2.0700	6.177
X5=Diploma I/II/III	28	19	20.37	0.0916	0.140
X5=S1/D4	119	69	91.64	5.5953	9.505
X5=S2	6	3	4.05	0.2716	0.408
X5=SD/Sederajat	108	89	65.30	8.6003	15.033
X5=SMP/Sederajat	142	110	89.05	4.9277	8.770
X5=Tidak/belum tamat SD	53	41	35.33	0.9110	1.435

Chisq= 34.3 on 6 degrees of freedom, p= 6e-06

```
> survdiff(Surv(Y,Status)~X6,data=carikerja)
```

Call:

```
survdiff(formula = Surv(Y, Status) ~ X6, data = carikerja)
```

	N	Observed	Expected	(O-E) <sup>2</sup> /E	(O-E) <sup>2</sup> /V
X6=Tidak	234	130	177	12.67	26.2
X6=Ya	648	484	437	5.15	26.2

Chisq= 26.2 on 1 degrees of freedom, p= 3e-07

```
> survdiff(Surv(Y,Status)~X7,data=carikerja)
```

Call:

```
survdiff(formula = Surv(Y, Status) ~ X7, data = carikerja)
```

	N	Observed	Expected	(O-E) <sup>2</sup> /E	(O-E) <sup>2</sup> /V
X7=Tidak	610	433	415	0.744	3.38
X7=Ya	272	181	199	1.556	3.38

Chisq= 3.4 on 1 degrees of freedom, p= 0.07

```
> survdiff(Surv(Y,Status)~X8,data=carikerja)
```

Call:

```
survdiff(formula = Surv(Y, Status) ~ X8, data = carikerja)
```

	N	Observed	Expected	(O-E) <sup>2</sup> /E	(O-E) <sup>2</sup> /V
X8=Pedesaan	300	236	183	15.45	33.8
X8=Perkotaan	582	378	431	6.55	33.8

Chisq= 33.8 on 1 degrees of freedom, p= 6e-09

```
> survdiff(Surv(Y,Status)~X1.X4,data=carikerja)
```

Call:

```
survdiff(formula = Surv(Y, Status) ~ X1.X4, data =  
carikerja)
```

(O-E) <sup>2</sup> /V	N	Observed	Expected	(O-E) <sup>2</sup> /E
X1.X4=Laki-laki belum kawin 34.774	214	116	169.9	17.086
X1.X4=Laki-laki cerai 2.699	15	7	11.7	1.870
X1.X4=Laki-laki sudah kawin 1.505	247	178	166.9	0.739
X1.X4=Perempuan belum kawin 0.693	152	104	97.9	0.385
X1.X4=Perempuan cerai 4.106	45	38	29.3	2.609
X1.X4=Perempuan sudah kawin 14.804	209	171	138.4	7.659

Chisq= 44.9 on 5 degrees of freedom, p= 2e-08

### Lampiran 6 Dalil Tentang Pernikahan

QS. An-Nur ayat 32:

وَأَنْكِحُوا الْأَيَامَىٰ مِنْكُمْ وَالصَّالِحِينَ مِنْ عِبَادِكُمْ وَإِمَائِكُمْ ۚ إِنْ يَكُونُوا فُقَرَاءَ يُغْنِهِمُ اللَّهُ مِنْ فَضْلِهِ ۗ وَاللَّهُ وَاسِعٌ عَلِيمٌ

Artinya:

“Dan nikahkanlah orang-orang yang masih membujang di antara kamu, dan juga orang-orang yang layak (menikah) dari hamba-hamba sahayamu yang laki-laki dan perempuan. Jika mereka miskin, Allah akan memberi kemampuan kepada mereka dengan karunia-Nya. Dan Allah Mahaluas (pemberian-Nya), Maha Mengetahui.” (QS An-Nur ayat 32)