



UNIVERSITAS INDONESIA

**MODEL PARAMETRIK WEIBULL
*ACCELERATED FAILURE TIME (AFT)***

SKRIPSI

**CINDY AYUNI SAFITRI
0806325453**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
PROGRAM STUDI SARJANA MATEMATIKA
DEPOK
JUNI 2012**



UNIVERSITAS INDONESIA

**MODEL PARAMETRIK WEIBULL
*ACCELERATED FAILURE TIME (AFT)***

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana sains

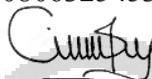
**CINDY AYUNI SAFITRI
0806325453**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
PROGRAM STUDI SARJANA MATEMATIKA
DEPOK
JUNI 2012**

HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri,
dan semua sumber baik yang dikutip maupun dirujuk
telah saya nyatakan dengan benar.



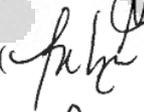
Nama : Cindy Ayuni Safitri
NPM : 0806325453
Tanda Tangan : 
Tanggal : 19 Juni 2012

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :
Nama : Cindy Ayuni Safitri
NPM : 0806325453
Program Studi : Matematika
Judul Skripsi : Model Parametrik Weibull *Accelerated Failure Time*
(AFT)

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia

DEWAN PENGUJI

Pembimbing : Sarini Abdullah, S.Si., M.Stats. ()
Pembimbing : Mila Novita, S.Si., M.Si. ()
Penguji : Fevi Novkaniza, S.Si., M.Si. ()
Penguji : Dra. Siti Nurrohmah, M.Si. ()
Penguji : Dr. Dian Lestari ()

Ditetapkan di : Depok
Tanggal : 19 Juni 2012

KATA PENGANTAR

Alhamdulillah, puji syukur penulis panjatkan kepada Allah SWT, karena atas berkat dan rahmat-Nya, penulis dapat menyelesaikan skripsi ini. Penulisan skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Sains Jurusan Matematika pada Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia.

Penulis menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi penulis untuk menyelesaikan skripsi ini. Oleh karena itu, penulis mengucapkan terima kasih kepada berbagai pihak antara lain:

1. Pembimbing tugas akhir penulis, Sarini Abdullah, S.Si., M.Stats. dan Mila Novita, S.Si., M.Si., yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan serta membimbing penulis dalam penyusunan skripsi ini. Terima kasih juga untuk kesabaran, nasehat, doa, dan dukungan yang telah diberikan selama penyusunan skripsi ini.
2. Dr. Sri Mardiyati, M. Kom. selaku pembimbing akademik penulis yang telah memberikan bantuan, nasehat, dan motivasi, serta meluangkan waktunya untuk berbagi cerita selama 4 tahun masa perkuliahan penulis.
3. Dr. Yudi Satria M.T. dan Rahmi Rusin, S.Si., M.ScTech., selaku ketua dan sekretaris Departemen Matematika UI, atas segala bantuan dan dukungan yang telah diberikan.
4. Seluruh staf pengajar di Departemen Matematika UI, terima kasih atas semua ilmu yang telah diajarkan.
5. Seluruh staf Tata Usaha (terutama Mba Santi yang sering direpotkan), staf Perpustakaan, serta karyawan Departemen Matematika, terima kasih atas segala bantuannya.
6. Ibu dan almarhum ayah tersayang, serta seluruh keluarga besar, yang selalu memberikan doa, nasihat, dukungan, dan motivasi kepada penulis sehingga dapat menjalani perkuliahan sampai menyelesaikan skripsi ini dengan baik.

7. Numa, mamih Sita, Ines, Risya, Tuti, QQ, Ade, Dhila, Hindun, Dheni, bang Andy, Arief, dan Adhi, terima kasih untuk persahabatan, canda tawa, curhat, “mecin”, dukungan, bantuan, dan kebersamaannya selama perkuliahan sampai sama-sama berjuang menyelesaikan skripsi kita ini.
8. Keluarga Matematika 2008, oma Icha dan Eka (teman seperjuangan aktuarial yang bertahan sampai skripsi sama bunda), Emy, Luthfa, Umbu, Awe, Resti dan Siwi (teman sepembimbingan, semangat!), Ega, Dhea, Uchi, Agnes, Agy, Anisah, Arkies, Arman, uni Aci, Aya, Bowo, Citra, Danis, Dede, Dhewe, Dian, Dini, Fani, Kohen, Ifah, Janu, Juni, Lian, Maimun, Masykur, Maul, May, Mei, Mela, Nadia, Nita, Nora, Oline, mas Puput, Purwo, Uci, Vika, Wulan, Yulial, dan Zee, terima kasih untuk semua semangat, bantuan, dan kebersamaannya selama 4 tahun ini, *Math UI 2008 One and Inseparable!*
9. Sahabat gelembung tersayang “Klan 13”, Uthie, Debby, Andin, Ijul, Chanchan, Debo, bubun Dece, Echa, Mbe’, Jujud, Epo, dan Saddam.
10. Kakak-kakak rusuh, Kak Anggie, Kak Farah, Kak Arif, Kak Wiwi, Kak Alfa, Kak Indah, dan Kak Nita, terima kasih buat semua cerita-ceritanya, semangatnya, dan waktunya buat mendengarkan keluh kesah penulis.
11. Kak Ajat (*special thanks to Mr. Classic Question*), bang Andy, dan Umbu yang sering direpotkan dalam pengerjaan skripsi ini.
12. Seluruh keluarga besar Matematika UI angkatan 2005-2011 (“kakak asuh” Kak Lena, Kak Ita, Kak Nedi, Kak Manda, Kak Sav, Kak Widya, Kak Lois, Kak Toto, Kak Stefi, Kak Farah, Kak Winda, Kak Gamar, Kak Anjar, Kak Anggun, Eja, Ai, Sita, Cepi, Tika, Mike, “adik asuh” Dwi, Aid, Yuza, Fariz, Bayu, dll) terima kasih untuk semua bantuan dan dukungannya selama masa perkuliahan.
13. Semua pihak yang telah memberikan doa dan semangat kepada penulis yang tidak bisa disebutkan satu per satu.

Akhir kata, penulis berharap Allah SWT berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga skripsi ini membawa manfaat bagi pengembangan ilmu pengetahuan.

Penulis

2012

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Indonesia, penulis yang bertanda tangan di bawah ini:

Nama : Cindy Ayuni Safitri
NPM : 0806325453
Program Studi : Sarjana
Departemen : Matematika
Fakultas : MIPA (Matematika dan Ilmu Pengetahuan Alam)
Jenis karya : Skripsi

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Indonesia Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah penulis yang berjudul :
Model Parametrik Weibull *Accelerated Failure Time* (AFT)

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Indonesia berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan memublikasikan tugas akhir penulis selama tetap mencantumkan nama penulis sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini penulis buat dengan sebenarnya.

Dibuat di : Depok
Pada tanggal : 19 Juni 2012
Yang menyatakan


(Cindy Ayuni Safitri)

ABSTRAK

Nama : Cindy Ayuni Safitri
Program Studi : Matematika
Judul : Model Parametrik Weibull *Accelerated Failure Time* (AFT)

Model parametrik Weibull digunakan ketika waktu *survival* diketahui berdistribusi Weibull. Dengan asumsi *accelerated failure time* (AFT), model parametrik Weibull AFT dibentuk dengan meregresikan kovariat secara linier terhadap log waktu. Koefisien regresi pada model parametrik Weibull AFT ditaksir dengan metode maksimum *likelihood*. Sebagai contoh penerapan digunakan data berupa waktu sampai meninggal untuk seseorang yang mengidap penyakit leukemia, dengan awal pengamatan saat pasien diberi suatu perlakuan. Selain itu, dilakukan juga simulasi data dengan men-*generate* data dari distribusi Weibull dan non-Weibull. Dengan pengecekan plot dari Cox-Snell residual, diperoleh hasil bahwa jika asumsi distribusinya tepat maka model Weibull AFT lebih baik dibandingkan model Cox PH yang tidak menggunakan asumsi distribusi waktu *survival*, dan sebaliknya jika asumsi tidak terpenuhi.

Kata Kunci : model parametrik, model *accelerated failure time*, distribusi weibull, model *cox proportional hazard*, *cox-snell residual*.
xiv+104 halaman; 17 gambar; 15 tabel
Daftar Pustaka : 7 (1995-2008)

ABSTRACT

Name : Cindy Ayuni Safitri
Program Study : Mathematics
Title : Weibull Accelerated Failure Time (AFT) Parametric Model

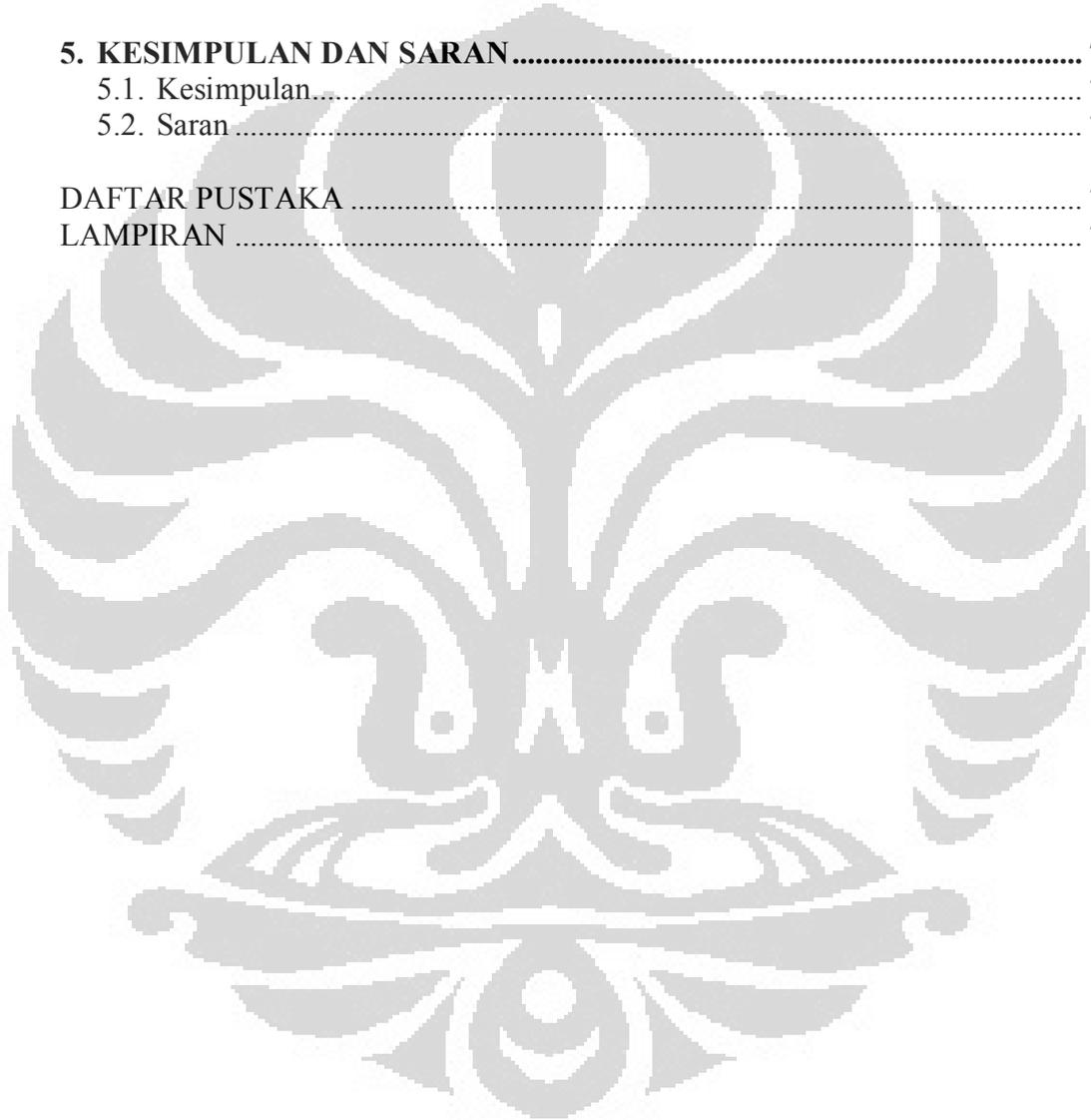
Weibull parametric model is used when the survival time follows a Weibull distribution. Under the assumption of accelerated failure time (AFT) model, the response in the model, i.e. the log of survival time, is modeled as a linear combination of the covariates. Regression coefficients are estimated using maximum likelihood method. As an example, data of time to death event for leukemia patients, with the beginning of the observation when the patient was given a treatment. In addition, simulations are also performed by generating data from the Weibull and non-Weibull distribution. By checking the plot of the Cox-Snell residuals, the results show that Weibull AFT model is better than Cox PH model when the assumption is met, while the Cox PH is better when the assumption is violated.

Keywords : parametric model, accelerated failure time model, weibull distribution, cox proportional hazard model, cox-snell residual
xiv+104 pages ; 17 pictures; 15 tables
Bibliography : 7 (1995-2008)

DAFTAR ISI

HALAMAN JUDUL.....	ii
HALAMAN PERNYATAAN ORISINALITAS.....	iii
HALAMAN PENGESAHAN.....	iv
KATA PENGANTAR.....	v
LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH.....	vii
ABSTRAK.....	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xii
DAFTAR TABEL.....	xiii
DAFTAR LAMPIRAN.....	xiv
1. PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah dan Ruang Lingkup Masalah.....	3
1.3 Jenis Penelitian dan Metode yang Digunakan.....	4
1.4 Tujuan Penulisan.....	4
2. LANDASAN TEORI.....	5
2.1 <i>Survival Time</i>	5
2.2 Kuantitas Dasar pada Analisis <i>Survival</i>	5
2.2.1 Fungsi <i>Survival</i>	6
2.2.2 Fungsi <i>Hazard</i>	8
2.3 Data Tersensor.....	11
2.3.1 Data Tersensor Kiri.....	12
2.3.2 Data Tersensor Kanan.....	12
2.3.3 Data Tersensor Interval.....	14
2.4 Model Cox <i>Proportional Hazard</i>	15
2.5 Estimasi Kaplan-Meier.....	17
2.6 Distribusi Weibull.....	17
2.7 Model <i>Accelerated Failure Time</i> (AFT).....	19
2.8 Metode <i>Maximum Likelihood</i>	20
2.9 Metode <i>Newton-Raphson</i>	21
3. MODEL PARAMETRIK WEIBULL AFT.....	23
3.1 Model Parametrik Weibull.....	23
3.1.1 Model Weibull PH.....	26
3.1.2 Model Weibull AFT.....	29
3.2 Estimasi Parameter Model Parametrik Weibull AFT.....	33
3.2.1 Konstruksi Fungsi <i>Likelihood</i> untuk Data <i>Survival</i>	34
3.2.2 Estimasi Parameter Model Parametrik Weibull AFT dengan Metode <i>Maximum Likelihood</i>	36
3.3 Pengecekan Model Weibull AFT.....	43

4. APLIKASI PADA DATA.....	48
4.1 Data Aplikasi.....	48
4.1.1 Pengecekan Asumsi Distribusi Weibull.....	49
4.1.2 Model Parametrik Weibull AFT.....	51
4.1.3 Model Cox PH.....	54
4.1.4 Perbandingan Model Cox PH dan Model Weibull AFT.....	55
4.2 Data Simulasi.....	58
4.2.1 Simulasi dengan Data Memenuhi Asumsi Distribusi Weibull.....	59
4.2.2 Simulasi dengan Data yang Tidak Berdistribusi Weibull.....	64
5. KESIMPULAN DAN SARAN.....	71
5.1. Kesimpulan.....	71
5.2. Saran.....	72
DAFTAR PUSTAKA.....	73
LAMPIRAN.....	74



DAFTAR GAMBAR

Gambar 2.1. Kurva fungsi <i>survival</i> secara teori dan praktik	7
Gambar 2.2. Berbagai bentuk kurva fungsi <i>hazard</i>	11
Gambar 2.3. Himpunan data dengan <i>survival time</i> eksak dan tersensor kanan	14
Gambar 2.4. Fungsi <i>hazard</i> Weibull untuk nilai $\lambda = 1$ dan $p = 1$ (\cdots), $p = 1.5$ ($---$), $p = 0.5$ ($---$).....	18
Gambar 3.1. Ilustrasi grafik hubungan $\ln[-\ln S(x)]$ dengan $\ln x$	26
Gambar 3.2. Garis waktu periode pengamatan	35
Gambar 3.3. Ilustrasi grafik hubungan $-\ln S_R(r_c)$ dengan r_c	45
Gambar 3.4. Bagan alur pengecekan Cox-Snell residual	47
Gambar 4.1. Grafik hubungan $\ln[-\ln S(x)]$ dengan $\ln x$ dari data leukemia	50
Gambar 4.2. Cox-Snell residual dari model Weibull AFT dengan data leukemia	57
Gambar 4.3. Cox-Snell residual dari model Cox PH dengan data leukemia	58
Gambar 4.4. Grafik hubungan $\ln[-\ln S(x)]$ dengan $\ln x$ dari data Weibull	59
Gambar 4.5. Cox-Snell residual dari model parametrik Weibull AFT dengan data Weibull	63
Gambar 4.6. Cox-Snell residual dari model Cox PH dengan data Weibull	64
Gambar 4.7. Grafik hubungan $\ln[-\ln S(x)]$ dengan $\ln x$ dari data lognormal.....	65
Gambar 4.8. Cox-Snell residual dari model parametrik Weibull AFT dengan data lognormal	69
Gambar 4.9. Cox-Snell residual dari model parametrik Cox PH dengan data lognormal	69

DAFTAR TABEL

Tabel 2.1.	Contoh penyajian data tersensor kanan	13
Tabel 3.1.	Bentuk Model Parametrik Weibull PH.....	28
Tabel 3.2.	Perbandingan Bentuk Model Weibull PH dan Model Weibull AFT .	33
Tabel 4.1.	Hasil estimasi <i>survival</i> Kaplan-Meier dari data leukemia.....	49
Tabel 4.2.	Hasil penaksiran parameter model Weibull AFT dari data leukemia.....	51
Tabel 4.3.	Hasil estimasi berdasarkan model parametrik Weibull AFT dari data leukemia.....	54
Tabel 4.4.	Hasil penaksiran parameter model Cox PH dari data leukemia.....	55
Tabel 4.5.	Estimasi Cox-Snell residual dari model Weibull AFT pada data leukemia.....	56
Tabel 4.6.	Estimasi Cox-Snell residual dari model Cox PH pada data leukemia.....	57
Tabel 4.7.	Hasil penaksiran parameter model Weibull AFT dari data Weibull .	60
Tabel 4.8.	Hasil estimasi berdasarkan model parametrik Weibull AFT dari data Weibull	61
Tabel 4.9.	Hasil penaksiran parameter model Cox PH dari data Weibull	62
Tabel 4.10.	Hasil penaksiran parameter model Weibull AFT dari data lognormal.....	66
Tabel 4.11.	Hasil estimasi berdasarkan model parametrik Weibull AFT dari data lognormal.....	66
Tabel 4.12.	Hasil penaksiran parameter model Cox PH dari data lognormal.....	67

DAFTAR LAMPIRAN

Lampiran 1. Data pasien leukemia	74
Lampiran 2. Estimasi Kaplan-Meier pada data pasien leukemia	75
Lampiran 3. Estimasi model Weibull AFT pada data pasien leukemia.....	76
Lampiran 4. Estimasi model Cox PH pada data pasien leukemia.....	77
Lampiran 5. Estimasi fungsi <i>survival</i> dari Cox-Snell residual untuk model Weibull AFT pada data pasien leukemia	78
Lampiran 6. Estimasi <i>survival</i> dari model Cox PH pada data pasien leukemia..	79
Lampiran 7. Estimasi fungsi <i>survival</i> dari Cox-Snell residual untuk model Cox PH pada data pasien leukemia	80
Lampiran 8. Data berdistribusi Weibull dengan parameter $\lambda = 6$ dan $p = 2$	81
Lampiran 9. Data berdistribusi lognormal dengan parameter $\mu = 2.5$ dan $\sigma = 1$	83
Lampiran 10. Estimasi Kaplan-Meier pada data Weibull.....	85
Lampiran 11. Estimasi model Weibull AFT pada data Weibull	87
Lampiran 12. Estimasi model Cox PH pada data Weibull	88
Lampiran 13. Estimasi fungsi <i>survival</i> dari Cox-Snell residual untuk model Weibull AFT pada data Weibull	89
Lampiran 14. Estimasi <i>survival</i> dari model Cox PH pada data Weibull	91
Lampiran 15. Estimasi fungsi <i>survival</i> dari Cox-Snell residual untuk model Cox PH pada data Weibull	93
Lampiran 16. Estimasi Kaplan-Meier pada data lognormal	95
Lampiran 17. Estimasi model parametrik Weibull AFT pada data lognormal	97
Lampiran 18. Estimasi model Cox PH pada data lognormal	98
Lampiran 19. Estimasi fungsi <i>survival</i> dari Cox-Snell residual untuk model Weibull AFT pada data lognormal	99
Lampiran 20. Estimasi <i>survival</i> dari model Cox PH pada data lognormal.....	101
Lampiran 21. Estimasi fungsi <i>survival</i> dari Cox-Snell residual untuk model Cox PH pada data lognormal	103

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Analisis *survival* adalah sekumpulan prosedur statistika untuk menganalisis data di mana variabel yang diperhatikan adalah waktu sampai terjadinya suatu *event* (Kleinbaum & Klein, 2005). Analisis pada data *survival* melibatkan data *time-to-event*, di mana waktu sampai terjadinya suatu *event* yang diinginkan disebut *survival time* atau *failure time* (Zhao, 2008).

Event yang diamati dapat berupa kematian, munculnya suatu penyakit, kambuhnya suatu penyakit setelah dilakukan *treatment*, atau hal-hal lainnya yang dapat terjadi dan diperhatikan pada suatu individu. *Event* pada analisis *survival* biasanya menyatakan suatu kegagalan (*failure*) karena *event* yang diamati biasanya berupa kejadian negatif, seperti kematian atau munculnya suatu penyakit. Akan tetapi, waktu *survival* juga dapat berupa kejadian positif, seperti waktu pemulihan pasien yang terkena demam berdarah.

Waktu yang diukur pada analisis *survival* disebut sebagai waktu *survival* karena variabel tersebut menunjukkan waktu individu dapat bertahan (*survive*) selama dilakukan pengamatan. Skala waktu harus dinyatakan dengan jelas, misal minggu, hari, atau jam. Awal dan akhir dari pengamatan juga harus didefinisikan dan terukur dengan jelas.

Pada analisis *survival* terdapat beberapa skema data, yaitu penyensoran dan pemancungan. Suatu data dikatakan tersensor jika *event* yang diamati pada suatu individu hanya diketahui terjadi setelah atau sebelum titik waktu tertentu. Jika waktu pengamatan diketahui secara pasti, maka data dikatakan eksak. Adapun pemancungan merupakan suatu mekanisme seleksi apakah suatu individu dapat dimasukkan ke dalam pengamatan atau tidak.

Pada saat melakukan pengamatan, terkadang waktu *survival* tidak berdiri sendiri, ada beberapa faktor yang dapat mempengaruhi waktu terjadinya suatu *event*. Faktor-faktor tersebut dilihat pada setiap individu pada saat pengamatan, misalnya umur, jenis kelamin, *treatment* yang diberikan, dan lainnya. Informasi

tersebut dapat digunakan dalam perhitungan secara bersamaan sebagai variabel prediktor atau disebut *covariate*.

Salah satu tujuan analisis *survival* adalah untuk mengetahui hubungan antara waktu kejadian (*time to failure*) dan variabel prediktor (*covariate*) yang diukur pada saat dilakukan penelitian. Metode yang seringkali digunakan adalah dengan menggunakan model *Cox proportional hazard* (Cox PH). Model Cox PH menggunakan asumsi *proportional hazard* (PH) di mana *hazard ratio* diasumsikan konstan sepanjang waktu.

Alasan utama mengapa model Cox PH lebih sering digunakan adalah karena model Cox PH tidak bergantung pada asumsi distribusi dari waktu kejadiannya (Qi, 2009). Model Cox PH merupakan model semiparametrik karena walaupun parameter regresinya diketahui, distribusi dari waktu kejadiannya tetap tidak diketahui. Fungsi *baseline hazard*, yaitu fungsi *hazard* yang melibatkan waktu tetapi tidak melibatkan kovariat, tidak ditentukan pada model Cox PH.

Meskipun model Cox PH sering digunakan secara luas, model ini juga memiliki kelemahan. Karena fungsi *baseline hazard* tidak ditentukan pada model, maka, dari hasil penaksiran parameter efek kovariat yang diperoleh, hanya dapat dilihat *hazard ratio*, yaitu perbandingan *hazard* antara dua individu dengan kovariat yang berbeda. Dengan kata lain, model Cox PH hanya memberikan informasi pengaruh dari suatu kovariat terhadap *hazard*. Kuantitas lainnya seperti fungsi *survival*, fungsi *hazard*, dan prediksi waktu *survival* tidak dapat dihitung dengan menggunakan model Cox PH.

Pada tugas akhir ini, akan dibahas model *survival* lainnya, yaitu model parametrik. Pada model parametrik, waktu *survival* diasumsikan mengikuti suatu distribusi tertentu. Fungsi *baseline hazard* pada model parametrik dapat ditentukan berdasarkan distribusi waktu kejadiannya, sehingga kuantitas seperti fungsi *survival*, fungsi *hazard*, dan prediksi waktu *survival* dapat dihitung dengan menggunakan model parametrik.

Model parametrik dapat berupa model parametrik PH dan model *accelerated failure time* (AFT). Model parametrik PH adalah bentuk parametrik dari model Cox PH, di mana fungsi *baseline hazard* mengikuti suatu distribusi tertentu. Walaupun model parametrik PH mudah diaplikasikan untuk analisis data *survival*,

hanya ada sedikit distribusi yang memungkinkan untuk digunakan pada model ini. Sehingga, model AFT dapat menjadi alternatif untuk melakukan analisis *survival* secara parametrik.

Pada model AFT, waktu *survival* dimodelkan sebagai fungsi dari variabel prediktor. Oleh karena itu, model AFT dapat mengukur secara langsung efek dari variabel prediktor terhadap waktu *survival*, bukan terhadap *hazard* seperti pada model PH. Hal tersebut akan lebih memudahkan pada saat melakukan interpretasi parameternya.

Dari beberapa distribusi yang dapat digunakan, pada tugas akhir ini akan dibahas secara khusus untuk penggunaan model parametrik dengan waktu *survival* yang berdistribusi Weibull. Model Weibull merupakan model parametrik yang dapat memenuhi asumsi PH dan AFT. Namun, hanya model Weibull dengan asumsi AFT yang akan dibahas pada tugas akhir ini. Selain itu juga akan dibandingkan antara hasil regresi parametrik (Weibull AFT) dan semiparametrik (Cox PH) yang diaplikasikan pada suatu data. Data yang digunakan adalah data tersensor kanan untuk menyesuaikan penggunaan pada model Cox PH.

1.2 Perumusan Masalah dan Ruang Lingkup Masalah

Perumusan masalah yang diajukan pada tugas akhir ini adalah:

1. Bagaimana pemodelan pada regresi parametrik Weibull AFT?
2. Bagaimana mencari taksiran parameter pada model Weibull AFT?
3. Bagaimana penerapan model Weibull AFT pada suatu data dan perbandingannya dengan model Cox PH?

Ruang lingkup pembahasan masalah pada tugas akhir ini meliputi:

1. Data yang digunakan untuk simulasi adalah data tersensor kanan.
2. Variabel prediktor yang digunakan adalah variabel yang tidak bergantung pada waktu (*time independent*).

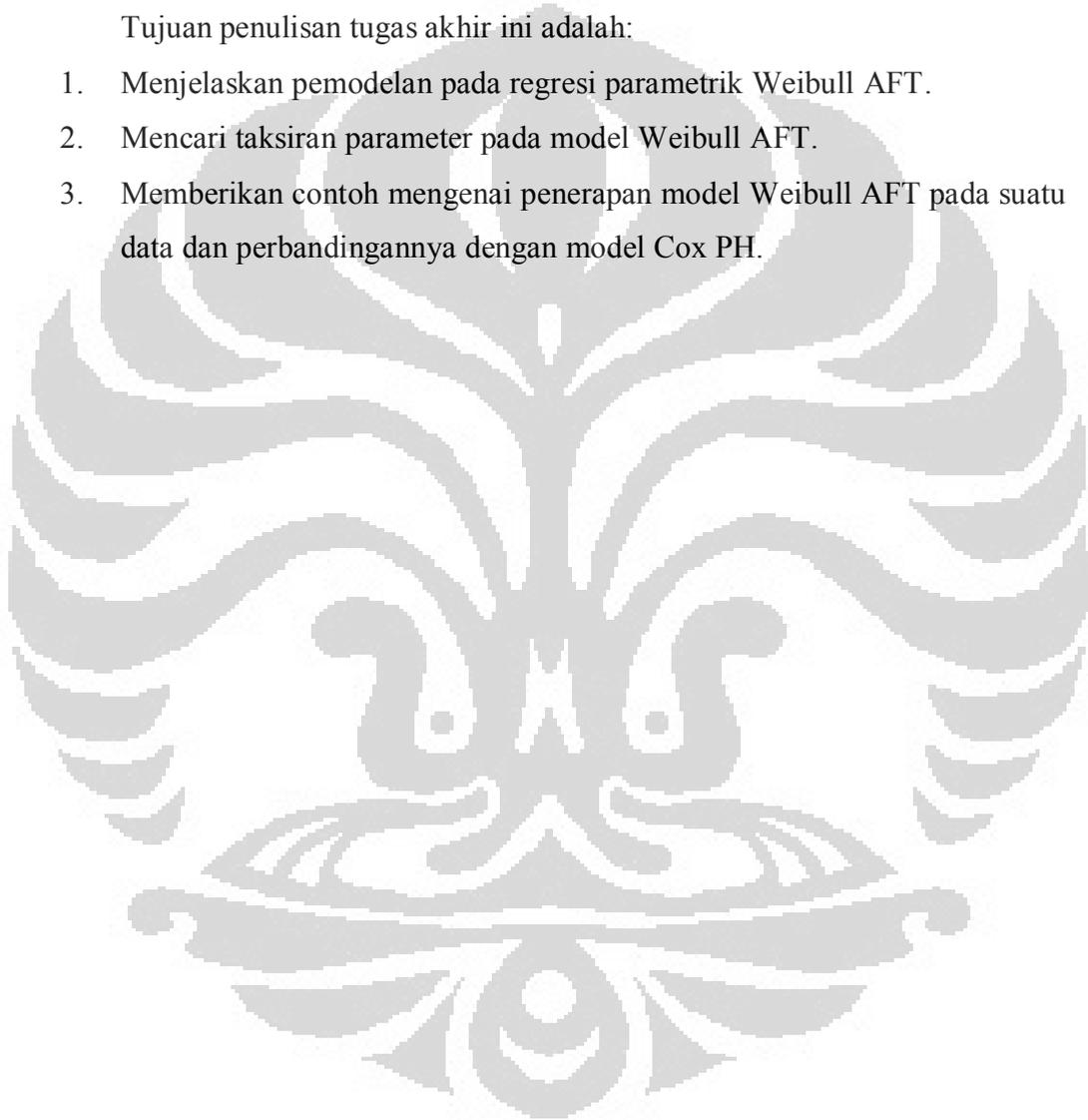
1.3 Jenis Penelitian dan Metode yang Digunakan

Jenis penelitian yang digunakan dalam pembuatan tugas akhir ini adalah studi literatur. Metode yang digunakan untuk menaksir parameter pada model Weibull AFT adalah *maximum likelihood*.

1.4 Tujuan Penulisan

Tujuan penulisan tugas akhir ini adalah:

1. Menjelaskan pemodelan pada regresi parametrik Weibull AFT.
2. Mencari taksiran parameter pada model Weibull AFT.
3. Memberikan contoh mengenai penerapan model Weibull AFT pada suatu data dan perbandingannya dengan model Cox PH.



BAB 2

LANDASAN TEORI

Pada bab ini akan diberikan beberapa konsep dasar yang digunakan pada analisis *survival*, seperti *survival time*, kuantitas dasar analisis *survival*, dan data tersensor. Selain itu juga akan dibahas mengenai model *Cox proportional hazard* (PH), estimasi Kaplan-Meier, distribusi Weibull, model *Accelerated Failure Time* (AFT), serta estimasi *maximum likelihood*.

2.1 *Survival Time*

Survival time didefinisikan sebagai variabel yang mengukur waktu dari suatu titik awal (misalkan waktu di mana suatu *treatment* dimulai) sampai titik akhir yang ditentukan (misalkan waktu sampai munculnya tumor) (Zhao, 2008).

Terdapat tiga hal yang harus diperhatikan untuk menentukan waktu *survival* secara tepat, yaitu sebagai berikut :

1. Waktu awal pengamatan yang terdefinisi dengan baik, yaitu waktu ketika suatu individu masuk dalam pengamatan.
2. Skala waktu sebagai satuan pengukuran harus jelas. Skala yang digunakan dapat berupa tahun, bulan, minggu, atau hari.
3. Waktu akhir pengamatan yang terdefinisi dengan baik, yaitu waktu ketika individu mengalami *event* atau batas waktu dilakukannya pengamatan terhadap individu tersebut.

Data *survival* seringkali dikumpulkan dalam periode waktu tertentu yang lamanya dibatasi. Selain waktu *survival* dari individu yang diamati, data *survival* ini juga dapat mengandung nilai dari kovariat (yaitu, faktor-faktor yang dapat mempengaruhi waktu *survival*) yang juga diperhatikan pada saat pengamatan.

2.2 Kuantitas Dasar pada Analisis *Survival*

Misalkan X adalah waktu sampai suatu *event* terjadi. Dalam hal ini, X adalah sebuah variabel random yang nonnegatif, $X \geq 0$, yang berasal dari populasi yang homogen. Ada empat kuantitas dasar yang biasa digunakan dalam analisis

survival untuk menggambarkan keadaan dari data *survival*, yaitu *probability density function*, fungsi *survival*, fungsi *hazard (hazard rate)*, dan *mean residual life*. Pada subbab ini hanya akan dibahas dua kuantitas dasar, yaitu fungsi *survival* dan fungsi *hazard*.

2.2.1 Fungsi *Survival*

Kuantitas dasar yang digunakan untuk menggambarkan kejadian *time to event* adalah fungsi *survival*, yaitu probabilitas seorang individu dapat bertahan sampai waktu x (baru mengalami *event* setelah waktu x). Secara matematis, fungsi *survival* dinyatakan sebagai:

$$S(x) = \Pr(X > x). \quad (2.1)$$

Secara teori, X adalah variabel random kontinu. Fungsi *survival* merupakan komplemen dari fungsi distribusi kumulatif, yaitu probabilitas bahwa variabel random X kurang dari atau sama dengan waktu x atau secara matematis dinyatakan dengan

$$F(x) = \Pr(X \leq x),$$

sehingga fungsi *survival* dapat dinyatakan dengan

$$S(x) = \Pr(X > x) = 1 - \Pr(X \leq x) = 1 - F(x). \quad (2.2)$$

Fungsi *survival* juga dapat dinyatakan dalam bentuk *pdf*, $f(x)$, yaitu

$$S(x) = \Pr(X > x) = \int_x^{\infty} f(t) dt. \quad (2.3)$$

Untuk X berupa variabel random diskrit atau hasil diskritisasi dari variabel random kontinu (yang biasanya merupakan pengukuran pada data *real*), maka diperlukan teknik yang berbeda. Misalkan nilai-nilai dari X yang teramati adalah $x_j, j = 1, 2, \dots$ dengan *probability mass function (p.m.f)*

$$p(x_j) = \Pr(X = x_j), \quad j = 1, 2, \dots$$

di mana $x_1 < x_2 < \dots$.

Fungsi *survival* untuk variabel random diskrit X diberikan oleh

$$S(x) = \Pr(X > x) = \sum_{x_j > x} p(x_j). \quad (2.4)$$

(Klein and Moeschberger, 1997 : 26)

Fungsi *survival* dapat ditunjukkan dalam berbagai bentuk kurva. Namun, semua fungsi *survival* memiliki sifat-sifat dasar seperti berikut:

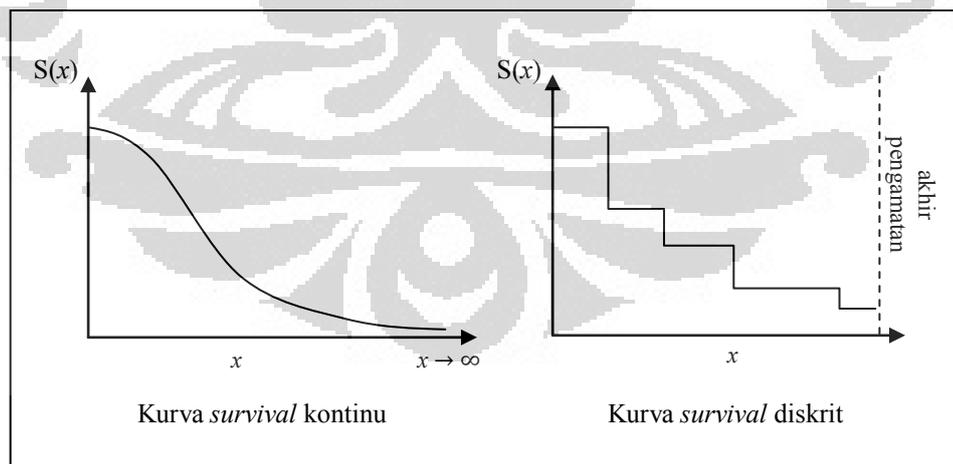
1. Fungsi *survival* adalah fungsi monoton tak naik.
2. Fungsi *survival* bernilai satu pada saat $x = 0$, atau $S(0) = 1$.

Hal ini karena semua individu masih bertahan pada awal pengamatan sehingga probabilitas *survival* pada saat itu adalah 1.

3. Fungsi *survival* bernilai nol pada saat x menuju tak hingga.

Hal ini karena pada akhirnya tidak ada individu yang akan bertahan jika waktu pengamatan berlangsung tanpa batas, sehingga probabilitas *survival* akan bernilai 0.

Penyajian fungsi *survival* secara grafis lebih dapat menggambarkan sifat-sifat di atas. Jika waktu *survival* adalah variabel random kontinu, maka grafik *survival* secara teoritis akan berbentuk kurva mulus. Namun, jika waktu *survival* adalah variabel random diskrit, maka kurva *survival* yang diperoleh merupakan fungsi tangga (seperti pada gambar 2.1). Selain itu, karena lamanya periode pengamatan tidak mungkin sampai menuju tak berhingga, bisa saja terdapat individu yang tidak mengalami *event* sampai akhir periode pengamatan, sehingga taksiran fungsi *survival* yang diperoleh tidak sama dengan nol pada akhir masa studi.



Gambar 2.1. Kurva fungsi *survival* kontinu dan diskrit

2.2.2 Fungsi Hazard

Secara matematis, fungsi *hazard* dapat dinyatakan sebagai:

$$h(x) = \lim_{\Delta x \rightarrow 0} \frac{\Pr[x < X \leq x + \Delta x \mid X > x]}{\Delta x} \quad (2.5)$$

Berdasarkan perumusan tersebut, fungsi *hazard* menyatakan *rate* bahwa seorang individu akan mengalami *event*, tepat sesaat sesudah titik waktu x .

Jika X adalah variabel random kontinu dengan *pdf* $f(x)$, maka dari persamaan (2.5) diperoleh:

$$\begin{aligned} h(x) &= \lim_{\Delta x \rightarrow 0} \frac{\Pr(x < X \leq x + \Delta x \mid X > x)}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{\Pr[(x < X \leq x + \Delta x) \cap (X > x)]}{\Pr(X > x) \cdot \Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{P(x < X \leq x + \Delta x)}{\Pr(X > x) \cdot \Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{P(x < X \leq x + \Delta x)}{S(x) \cdot \Delta x} \\ &= \frac{1}{S(x)} \cdot \lim_{\Delta x \rightarrow 0} \frac{P(x < X \leq x + \Delta x)}{\Delta x} \\ &= \frac{f(x)}{S(x)} \end{aligned} \quad (2.6)$$

Dari persamaan (2.2) diperoleh hubungan $S(x) = 1 - F(x)$, dan

$$f(x) = F'(x) = \frac{d(F(x))}{dx} = \frac{d(1 - S(x))}{dx} = -\frac{d(S(x))}{dx} = -S'(x), \quad (2.7)$$

maka

$$\begin{aligned} h(x) &= \frac{f(x)}{S(x)} = -\frac{S'(x)}{S(x)} = -S'(x) \cdot \frac{d \ln S(x)}{dS(x)} = -\frac{dS(x)}{dx} \cdot \frac{d \ln S(x)}{dS(x)} \\ &= -\frac{d \ln S(x)}{dx} \end{aligned} \quad (2.8)$$

Dari persamaan tersebut diperoleh,

$$\begin{aligned} \int_0^x h(u) du &= -\int_0^x \frac{d \ln S(u)}{du} du \\ -\int_0^x h(u) du &= \int_0^x \frac{d}{du} \ln S(u) du \end{aligned}$$

$$-\int_0^x h(u)du = \ln S(u) \Big|_0^x$$

$$-\int_0^x h(u)du = \ln S(x) - \ln S(0)$$

Karena $S(0) = 1$, maka $\ln S(0) = 0$, sehingga persamaan di atas menjadi

$$-\int_0^x h(u)du = \ln S(x)$$

dan diperoleh

$$S(x) = \exp \left[-\int_0^x h(u)du \right]. \quad (2.9)$$

Berdasarkan penjelasan di atas, diperoleh hubungan antara $S(x)$, $h(x)$, dan $f(x)$ sebagai berikut:

- $f(x) = -S'(x)$
- $h(x) = \frac{f(x)}{S(x)}$
- $S(x) = \exp \left[-\int_0^x h(u)du \right]$

Sehingga dapat disimpulkan bahwa jika salah satu kuantitas dasar diketahui, maka kuantitas dasar yang lain akan dapat ditentukan juga.

Adapun untuk variabel random diskrit, fungsi *hazard*-nya adalah sebagai berikut:

$$h(x_j) = \Pr(X = x_j | X \geq x_j), \quad j = 1, 2, \dots, n$$

$$= \frac{\Pr(X = x_j \cap X \geq x_j)}{\Pr(X \geq x_j)}$$

$$= \frac{\Pr(X = x_j)}{\Pr(X \geq x_j)} = \frac{\Pr(X = x_j)}{\Pr(X > x_{j-1})}$$

$$= \frac{p(x_j)}{S(x_{j-1})} \quad (2.10)$$

di mana $p(x_j) = \Pr(X = x_j)$ dan $S(x_0) = 1$.

Berikut akan ditunjukkan hubungan fungsi *hazard* dan fungsi *survival* untuk titik waktu yang diskrit. Dari persamaan (2.1) diketahui

$$S(x_j) = \Pr(X > x_j)$$

dan

$$\begin{aligned} S(x_{j-1}) &= \Pr(X > x_{j-1}) = \Pr(X = x_j) + \Pr(X > x_j) \\ S(x_{j-1}) &= p(x_j) + S(x_j) \\ p(x_j) &= S(x_{j-1}) - S(x_j) \end{aligned} \quad (2.11)$$

Dari persamaan (2.10) dan (2.11) di atas, diperoleh

$$h(x_j) = \frac{p(x_j)}{S(x_{j-1})} = \frac{S(x_{j-1}) - S(x_j)}{S(x_{j-1})} = 1 - \frac{S(x_j)}{S(x_{j-1})}, \quad j = 1, 2, \dots \quad (2.12)$$

Fungsi *survival* dapat juga dinyatakan sebagai perkalian dari *survival* bersyarat,

$$S(x) = \prod_{x_j \leq x} \frac{S(x_j)}{S(x_{j-1})} \quad (2.13)$$

(Klein & Moeschberger, 1997 : 31)

Dengan demikian, berdasarkan persamaan (2.12) dan (2.13), dapat diperoleh hubungan fungsi *survival* dengan fungsi *hazard*, yaitu

$$S(x) = \prod_{x_j \leq x} [1 - h(x_j)]. \quad (2.14)$$

Kuantitas lainnya yang berhubungan dengan fungsi *hazard* adalah *cumulative hazard function* $H(x)$, yang didefinisikan sebagai

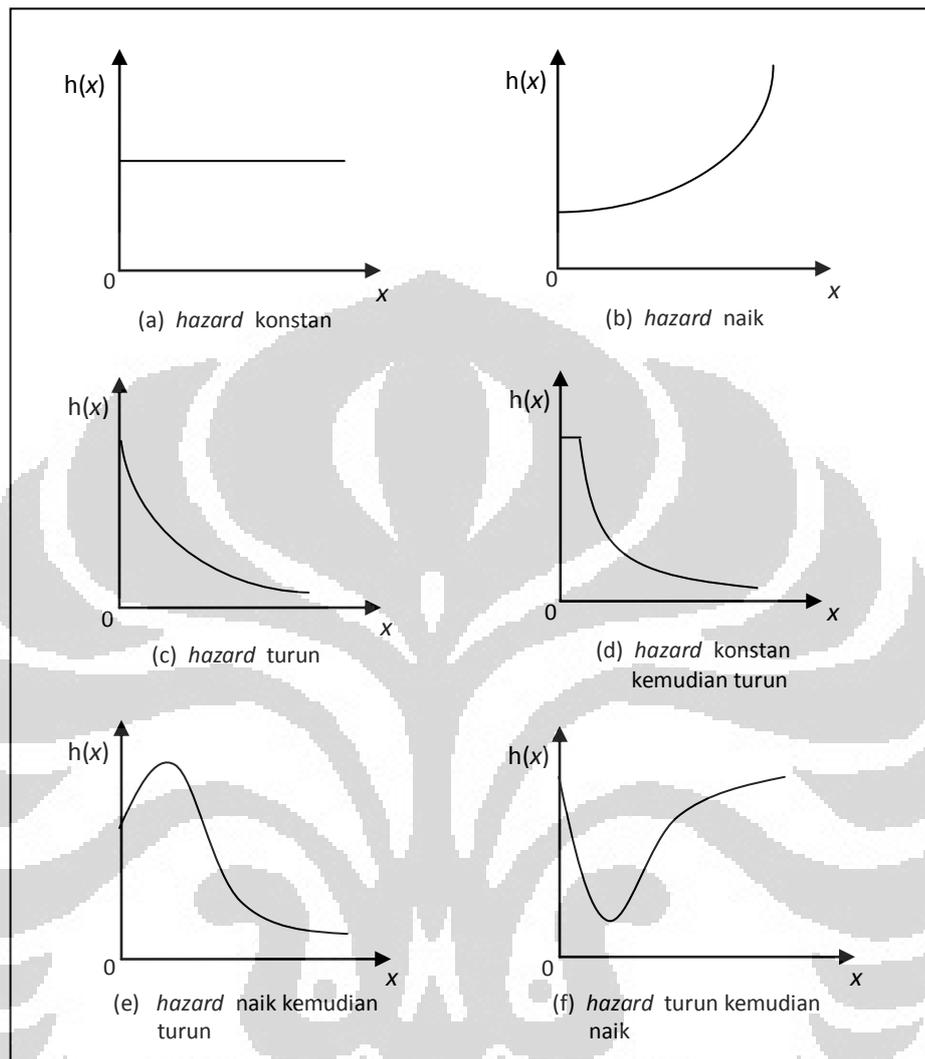
$$H(x) = \int_0^x h(u) du = -\ln[S(x)] \quad (2.15)$$

$$S(x) = \exp[-H(x)] \quad (2.16)$$

Penyajian fungsi *hazard* secara grafis dengan plot nilai-nilai *hazard rate* pada setiap titik waktu dapat mempermudah pemahaman terhadap waktu *survival* yang diamati. Berdasarkan persamaan (2.5), dapat ditentukan sifat dari fungsi *hazard*, yaitu:

- Selalu bernilai nonnegatif, $h(x) \geq 0$.
- Tidak memiliki batas atas.

Berikut adalah beberapa contoh grafik fungsi *hazard*.



Gambar 2.2. Berbagai bentuk kurva fungsi *hazard*

2.3 Data Tersensor

Pada analisis *survival*, data yang digunakan adalah data *time-to-event*. Data tersebut dapat berupa data eksak, data tersensor, ataupun data terpancung. Data eksak diperoleh apabila waktu tepatnya suatu *event* yang diinginkan terjadi diketahui. Suatu data dikatakan tersensor jika hanya diketahui sebagian informasi mengenai waktu sampai terjadinya *event* pada individu yang bersangkutan tetapi tidak diketahui waktu kejadiannya secara pasti. Sedangkan pemancungan pada suatu data merupakan suatu mekanisme seleksi seorang individu akan masuk sebagai sampel atau tidak.

Pada penulisan tugas akhir ini, data yang digunakan hanya data eksak dan data tersensor. Data tersensor terdiri dari tersensor kiri, tersensor kanan, dan tersensor interval. Berikut ini akan dijelaskan ketiga jenis data tersensor tersebut.

2.3.1 Data Tersensor Kiri

Perhatikan skenario berikut. Akan dilakukan penelitian mengenai usia saat seorang anak pertama kali bisa membaca. Pengamatan dilakukan selama satu tahun pada suatu taman kanak-kanak, yang diukur dalam tahun. Misalkan terdapat seorang anak yang ternyata sudah bisa membaca sebelum terdaftar sebagai siswa. Jika tidak dapat diketahui pasti usia anak tersebut saat pertama kali bisa membaca, maka titik waktu di mana anak tersebut terdaftar sebagai siswa di taman kanak-kanak tersebut dapat dianggap sebagai waktu penyensoran kiri. Skema pengamatan tersebut dapat dituliskan sebagai berikut:

- Waktu asal (*time origin*) : saat anak baru lahir,
- Skala waktu pengukuran : usia anak (dalam tahun),
- *Event* yang diamati : pertama kali bisa membaca,
- Waktu akhir (*end of study*) : setelah satu tahun pengamatan,
- Waktu penyensoran kiri : usia anak saat terdaftar sebagai siswa.

Dari cerita di atas, suatu data dikatakan tersensor kiri apabila *event* yang diamati pada individu sudah terjadi sebelum individu tersebut masuk dalam pengamatan.

2.3.2 Data Tersensor Kanan

Pada data tersensor kanan hanya diketahui bahwa waktu *survival* melebihi suatu nilai tertentu. Data tersensor kanan dapat terjadi karena beberapa hal sebagai berikut:

- Seorang individu belum mengalami *event* setelah periode pengamatan berakhir;
- Seorang individu keluar dari pengamatan di tengah-tengah masa studi;
- Seorang individu mengalami *event* tetapi bukan karena penyebab yang menjadi fokus penelitian.

Berikut ini diberikan contoh kasus data yang tersensor kanan. Misalkan akan dilakukan pengamatan masa pakai laptop merk A sebelum rusak. Misalkan waktu pengamatan dibatasi selama lima tahun, maka dalam periode pengamatan tersebut dapat terjadi hal-hal seperti berikut:

- Setelah satu tahun pemakaian terdapat laptop yang rusak, sehingga waktu ini disebut eksak.
- Mengalami kerusakan pada dua tahun awal pemakaian, tapi karena kecerobohan pengguna (terjatuh atau terkena air). Data seperti ini disebut tersensor kanan.
- Sesudah akhir masa pengamatan terdapat laptop yang masih dalam kondisi bagus, sehingga data ini juga tersensor kanan.

Pengukuran data tersensor kanan mencakup dua variabel, sebut (x, δ) , di mana x menunjukkan waktu *survival* seorang individu dalam pengamatan dan δ adalah variabel indikator yang menunjukkan waktu tersebut adalah waktu *survival* yang eksak atau tersensor kanan. δ akan bernilai 1 jika waktu *survival* diketahui secara pasti dan bernilai 0 jika waktu *survival* tersebut tersensor kanan.

Sebagai ilustrasi, berikut ini diberikan himpunan data yang terdiri dari lima individu yang diamati selama sepuluh tahun periode pengamatan, yaitu sebagai berikut:

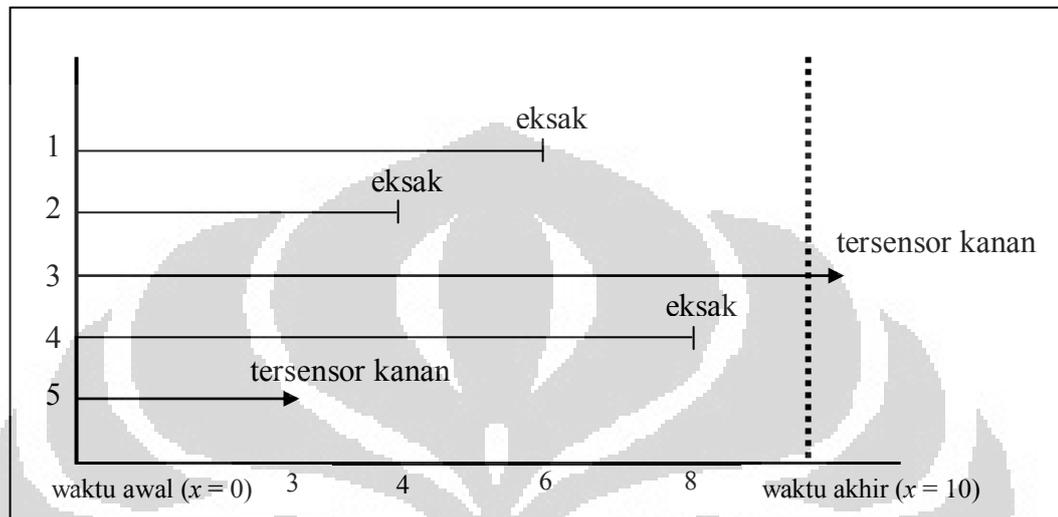
$$6, 4, 10^+, 8, 3^+.$$

Maka data di atas dapat disajikan sebagai berikut:

Tabel 2.1. Contoh penyajian data tersensor kanan

Individu	(x, δ)	Keterangan
1	(6, 1)	Eksak
2	(4, 1)	Eksak
3	(10, 0)	Sensor kanan
4	(8, 1)	Eksak
5	(3, 0)	Sensor kanan

Gambar 2.3 berikut ini menunjukkan gambaran garis waktu *survival* dari contoh sebelumnya. Waktu *survival* yang tersensor kanan ditunjukkan dengan anak panah di akhir garis waktu *survival*.



Gambar 2.3. Himpunan data dengan *survival time* eksak dan tersensor kanan

Dari gambar di atas, *survival time* individu 1, 2, dan 4 diketahui secara pasti, sedangkan *survival time* untuk individu 3 dan 5 tersensor kanan. Individu 3 belum mengalami *event* sampai akhir periode pengamatan, sedangkan individu 5 keluar dari pengamatan sebelum masa studi berakhir, sehingga keduanya memiliki *survival time* yang tersensor kanan.

2.3.3 Data Tersensor Interval

Jika pada data tersensor kanan yang diketahui hanyalah bahwa *event* terjadi sesudah suatu titik waktu tertentu, dan sebaliknya pada sensor kiri, maka gabungan dari kedua kondisi ini merupakan sensor interval. Dengan perkataan lain, suatu *event* hanya diketahui terjadi di antara suatu interval waktu tertentu, namun tidak ada informasi tepatnya titik waktu *survival* tersebut.

Sebagai contoh, misalkan ingin diamati usia berapa waktu pertama kali munculnya *angina pectoris* pada pasien penderita jantung koroner. Hal ini hanya dapat diketahui terjadi di antara dua waktu pemeriksaan secara klinis, sehingga data observasi yang diperoleh adalah data tersensor interval.

Data *survival* yang diperoleh bisa saja terdiri dari lebih dari satu skema penyensoran. Data tersensor kanan terdiri dari dua macam skema *censoring*, yaitu tipe I dan tipe II. Pada data tersensor kanan tipe I, *event* yang diamati hanya *event* yang terjadi selama waktu yang telah ditentukan. Pada data tersensor kanan tipe II, jumlah *event* yang akan diamati telah ditentukan terlebih dahulu dan pengamatan akan diakhiri jika jumlah *event* tersebut telah terpenuhi. Pada tugas akhir ini, data *survival* yang digunakan berupa data tersensor kanan tipe I.

2.4 Model Cox *Proportional Hazard*

Model *Cox Proportional Hazard* (Cox PH) adalah model regresi pada analisis *survival* yang dapat digunakan untuk melihat pengaruh kovariat terhadap waktu *survival* dari individu yang diamati.

Model Cox PH menyatakan *hazard rate* dari satu individu pada waktu x dengan diketahui kovariat \mathbf{Z} dengan model:

$$h(x, \mathbf{Z}) = h_0(x) \exp \left[\sum_{k=1}^m \beta_k Z_k \right] = h_0(x) \exp(\boldsymbol{\beta}^t \mathbf{Z}) \quad (2.17)$$

di mana

x = waktu *survival*

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

Z_k = kovariat ke- k , $k = 1, 2, \dots, m$

β_k = koefisien regresi yang bersesuaian dengan kovariat Z_k

Dari bentuk persamaan model Cox PH di atas, terlihat bahwa model ini mengandung dua komponen, yaitu:

- Efek waktu, yang diakomodir melalui *baseline hazard* $h_0(x)$
- Efek kovariat, yang diakomodir melalui $\exp[\boldsymbol{\beta}^t \mathbf{Z}]$, di mana \mathbf{Z} merupakan kovariat yang tidak bergantung pada waktu (*time independent covariate*). Jika \mathbf{Z} bergantung pada waktu (*time dependent covariate*), maka model Cox PH perlu dimodifikasi melalui bentuk *extended cox model*. Pada tugas akhir ini, yang akan digunakan adalah model Cox PH dengan *time independent covariate*, yang nilainya konstan sepanjang waktu.

Pada efek waktu, bentuk fungsional $h_0(x)$ bergantung pada distribusi dari waktu *survival* X . Namun, karena distribusi dari waktu *survival* tidak diketahui dan $h_0(x)$ tidak dapat ditentukan bentuknya, maka bagian ini adalah nonparametrik. Akan tetapi, pada efek kovariat, di bawah asumsi tertentu, taksiran parsial maksimum *likelihood* dari β secara asimtotik berdistribusi normal:

$$\hat{\beta} \sim N(\beta, [-l''(\beta)]^{-1}), \text{ di mana } l''(\beta) = \left[\frac{\partial^2 l}{\partial \beta_i \partial \beta_j} \right]$$

untuk $i = 1, 2, \dots, m$ dan $j = 1, 2, \dots, m$.

(Galbraits, 2004)

Oleh karena itu, model Cox PH termasuk dalam metode semiparametrik.

Mengingat bahwa distribusi dari X tidak perlu diketahui, model Cox PH umumnya sering dijadikan metode pilihan (alternatif) dari metode parametrik. Untuk mengatasi keterbatasan tidak diketahuinya $h_0(x)$, maka penggunaan Cox PH pun dibatasi pada *hazard ratio*, yaitu

$$HR = \frac{h_A(x, \mathbf{Z})}{h_B(x, \mathbf{Z}^*)} \quad (2.18)$$

di mana $h_A(x, \mathbf{Z})$ adalah *hazard rate* dari individu A dengan nilai kovariat

$$\mathbf{Z} = (Z_1, Z_2, \dots, Z_m)$$

dan $h_B(x, \mathbf{Z}^*)$ adalah *hazard rate* dari individu B dengan nilai kovariat \mathbf{Z} yang berbeda dari individu A, yaitu

$$\mathbf{Z}^* = (Z_1^*, Z_2^*, \dots, Z_m^*)$$

Berikut ditunjukkan bahwa HR dapat dihitung

$$\begin{aligned} HR &= \frac{h_A(x, \mathbf{Z})}{h_B(x, \mathbf{Z}^*)} = \frac{h_0(x) \exp[\sum_{k=1}^m \beta_k Z_k]}{h_0(x) \exp[\sum_{k=1}^m \beta_k Z_k^*]} \\ &= \exp\left(\sum_{k=1}^m \beta_k Z_k - \sum_{k=1}^m \beta_k Z_k^*\right) \\ &= \exp\left[\sum_{k=1}^m \beta_k (Z_k - Z_k^*)\right]. \end{aligned}$$

Apabila *hazard ratio* (HR) yang diperoleh tidak bergantung waktu, atau dengan kata lain, selalu bernilai sama di setiap titik waktu x , maka dapat dikatakan bahwa Z_1, Z_2, \dots, Z_m memenuhi asumsi *proportional hazard* (PH).

2.5 Estimasi Kaplan-Meier

Estimasi Kaplan-Meier adalah salah satu metode nonparametrik yang dapat digunakan untuk menaksir kuantitas dasar pada analisis *survival*. Misalkan *event* yang diamati terjadi pada D titik waktu yang berbeda, yaitu $x_1 < x_2 < \dots < x_D$, dan pada saat x_i terdapat d_i *event* yang terjadi. Misalkan Y_i adalah jumlah individu yang beresiko mengalami *event* pada waktu x_i . Atau dengan kata lain, Y_i adalah jumlah individu yang masih ada pada waktu x_i atau dapat mengalami *event* pada saat x_i . Estimasi fungsi *survival* yang dikemukakan oleh Kaplan dan Meier (1958) didefinisikan sebagai berikut untuk setiap interval nilai x di mana terdapat *event*:

$$\hat{S}(x) = \begin{cases} 1 & x < x_1 \\ \prod_{x_i \leq x} \left[1 - \frac{d_i}{Y_i} \right], & x_1 \leq x \end{cases} \quad (2.19)$$

(Klein and Moeschberger, 1997 : 84)

Estimasi Kaplan-Meier tersebut merupakan fungsi tangga dengan loncatan pada titik waktu *event* yang terobservasi. Besarnya loncatan tersebut tidak hanya tergantung pada banyaknya jumlah *event* yang terobservasi pada setiap waktu x_i , tetapi juga observasi yang tersensor sesaat sebelum x_i .

Estimasi Kaplan-Meier juga dapat digunakan untuk mengestimasi fungsi kumulatif *hazard* $H(x) = -\ln[S(x)]$. Sehingga estimasi yang diperoleh adalah

$$\hat{H}(x) = -\ln[\hat{S}(x)]$$

2.6 Distribusi Weibull

Distribusi Weibull merupakan salah satu distribusi yang sering digunakan dalam analisis data *survival*. Misalkan variabel random kontinu dari waktu kejadian X berdistribusi Weibull dengan parameter $\lambda > 0$ dan $p > 0$, atau disingkat $X \sim \text{Wei}(\lambda, p)$. Bentuk *pdf* dari distribusi Weibull adalah

$$f(x) = \lambda p x^{p-1} \exp(-\lambda x^p); \quad x \geq 0. \quad (2.20)$$

Fungsi *survival* dari distribusi Weibull diberikan oleh

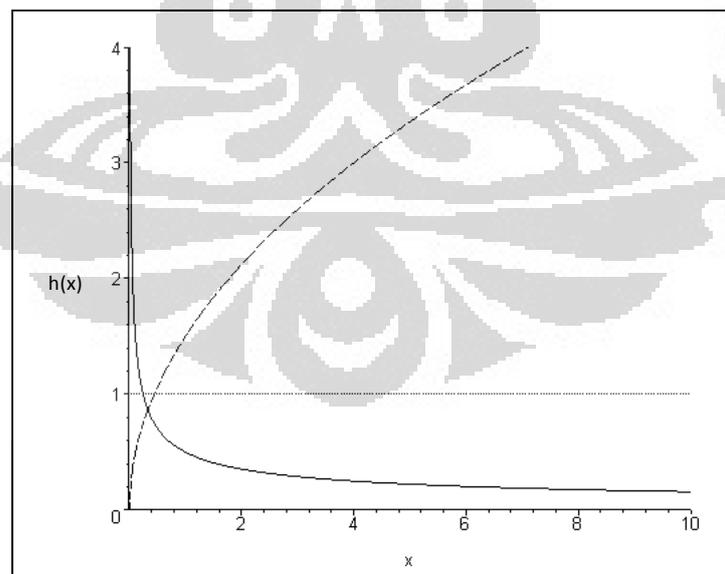
$$S(x) = \Pr(X > x) = \int_x^{\infty} f(t) dt$$

$$\begin{aligned}
&= \int_x^{\infty} \lambda p t^{p-1} \exp(-\lambda t^p) dt \\
&= \exp(-\lambda t^p) \Big|_x^{\infty} \\
S(x) &= \exp(-\lambda x^p); \quad x \geq 0
\end{aligned} \tag{2.21}$$

dan fungsi *hazard* dari distribusi Weibull adalah

$$\begin{aligned}
h(x) &= \frac{-d}{dx} \ln S(x) \\
&= \frac{d}{dx} \lambda x^p \\
&= \lambda p x^{p-1}
\end{aligned} \tag{2.22}$$

Parameter p disebut sebagai *shape parameter* karena parameter p ini menentukan perubahan bentuk dari kurva *hazard* seiring bertambahnya waktu x . Jika $p > 1$, maka kurva *hazard* monoton naik. Jika $p < 1$, maka kurva *hazard* monoton turun. Sedangkan untuk $p = 1$, maka *hazard* konstan dan mengikuti bentuk khusus distribusi eksponensial dengan $h(x) = \lambda$. Adanya *shape parameter* ini membuat distribusi Weibull menjadi fleksibel untuk suatu data *survival*.



Gambar 2.4. Fungsi *hazard* Weibull untuk nilai $\lambda = 1$ dan $p = 1$ (\cdots), $p = 1.5$ ($-\cdot-$), $p = 0.5$ ($—$)

2.7 Model *Accelerated Failure Time* (AFT)

Asumsi dasar untuk model AFT adalah efek dari kovariat konstan dan multiplikatif terhadap waktu *survival*. Dengan kata lain, kovariat mempengaruhi *survival* oleh suatu faktor konstan yang disebut *acceleration factor* (AF).

Model AFT ini menggambarkan penguluran waktu *survival* sebagai fungsi dari variabel prediktor. Sebagai ilustrasi dari asumsi AFT, bandingkan fungsi *survival* seorang perokok, $S_1(x)$ dan yang bukan perokok, $S_2(x)$. Asumsi AFT dapat dinyatakan dengan $S_2(x) = S_1(\gamma x)$ untuk $x \geq 0$, di mana γ adalah suatu konstanta yang disebut sebagai *acceleration factor* (AF) yang membandingkan waktu *survival* seorang perokok dan bukan perokok. Misalnya diketahui bahwa probabilitas seseorang yang “tidak merokok” dapat bertahan hidup pada usia 40 tahun sama dengan probabilitas seorang “perokok” dapat bertahan hidup pada usia 30 tahun, maka didapat nilai dari AF sebesar 0,75. Dengan cara yang sama dapat dinyatakan secara umum bahwa probabilitas seseorang yang tidak merokok dapat bertahan x tahun sama dengan probabilitas seorang perokok dapat bertahan 0,75 kali x tahun, atau secara matematis dapat dinyatakan bahwa $S_2(x) = S_1(0,75x)$.

Dalam melakukan regresi, AF γ dapat diparameterisasi dengan fungsi kovariat $\exp(\boldsymbol{\theta}^t \mathbf{Z})$, di mana \mathbf{Z} adalah kovariat dan $\boldsymbol{\theta}^t = (\theta_1, \dots, \theta_m)$ adalah koefisien regresi.

Definisi 2.1

Misalkan X adalah waktu *survival* dan \mathbf{Z} adalah vektor dari kovariat yang tidak bergantung pada waktu. Model AFT menyatakan bahwa fungsi *survival* dari individu dengan kovariat \mathbf{Z} pada saat x sama dengan fungsi *survival* individu dengan fungsi *baseline survival* pada saat $x \exp(\boldsymbol{\theta}^t \mathbf{Z})$, di mana $\boldsymbol{\theta}^t = (\theta_1, \dots, \theta_m)$ adalah vektor koefisien regresi (Klein and Moeschberger, 1997). Atau secara matematis, model AFT didefinisikan oleh persamaan:

$$S(x|\mathbf{Z}) = S_0[\exp(\boldsymbol{\theta}^t \mathbf{Z})x], \quad \forall x > 0 \quad (2.23)$$

dengan faktor $\exp(\boldsymbol{\theta}^t \mathbf{Z})$ disebut *acceleration factor*, yang mengukur perubahan efek kovariat terhadap waktu *survival*.

Berdasarkan model AFT tersebut, maka fungsi *hazard* untuk individu dengan kovariat \mathbf{Z} adalah

$$\begin{aligned}
h(x|\mathbf{Z}) &= \frac{-d}{dx} \ln S(x|\mathbf{Z}) \\
&= \frac{-d}{dx} \ln S_0[\exp(\boldsymbol{\theta}^t \mathbf{Z}) x] \\
&= \exp(\boldsymbol{\theta}^t \mathbf{Z}) h_0[\exp(\boldsymbol{\theta}^t \mathbf{Z}) x]
\end{aligned} \tag{2.24}$$

Pada model AFT diasumsikan bahwa efek dari kovariat multiplikatif terhadap waktu *survival*. Selain itu, hubungan antara kovariat dan waktu *survival* dapat dinyatakan sebagai hubungan linier antara log waktu dan nilai dari kovariat. Diasumsikan bahwa model linier untuk log waktu yang biasa digunakan adalah

$$Y = \ln X = \mu + \boldsymbol{\alpha}^t \mathbf{Z} + \sigma W \tag{2.25}$$

di mana $\boldsymbol{\alpha}$ adalah vektor parameter regresi, μ adalah *intercept*, σ adalah *scale parameter*, dan W adalah komponen error.

Kedua bentuk model (2.23) dan (2.25) berkaitan erat. Misalkan $S_0(x)$ adalah fungsi *survival* dari $X = e^Y$. Saat $\mathbf{Z} = 0$, $S_0(x)$ adalah fungsi *baseline survival* dari $\exp(\mu + \sigma W)$.

$$\begin{aligned}
S(x|\mathbf{Z}) &= \Pr(X > x | \mathbf{Z}) \\
&= \Pr(Y > \ln x | \mathbf{Z}) \\
&= \Pr(\mu + \boldsymbol{\alpha}^t \mathbf{Z} + \sigma W > \ln x | \mathbf{Z}) \\
&= \Pr(\mu + \sigma W > \ln x - \boldsymbol{\alpha}^t \mathbf{Z} | \mathbf{Z}) \\
&= \Pr(\exp(\mu + \sigma W) > x \exp(-\boldsymbol{\alpha}^t \mathbf{Z}) | \mathbf{Z}) \\
&= S_0(x \exp(-\boldsymbol{\alpha}^t \mathbf{Z}) | \mathbf{Z})
\end{aligned} \tag{2.26}$$

Sehingga model linier log waktu tersebut adalah model AFT dengan $\boldsymbol{\theta} = -\boldsymbol{\alpha}$.

2.8 Metode *Maximum Likelihood*

Metode *maximum likelihood* adalah salah satu metode yang dapat digunakan untuk mencari taksiran nilai parameter distribusi dari data. Misalkan X_1, X_2, \dots, X_n merupakan suatu sampel acak dari suatu distribusi yang mempunyai pdf $f(x; \theta)$: $\theta \in \Omega$. Pdf bersama dari X_1, X_2, \dots, X_n adalah $f(x_1; \theta)f(x_2; \theta) \dots f(x_n; \theta)$. Pdf bersama ini dapat dinyatakan sebagai fungsi dari θ dan disebut sebagai fungsi *likelihood* L dari sampel acak yang dinotasikan dengan:

$$L(\theta; x_1, x_2, \dots, x_n) = f(x_1; \theta)f(x_2; \theta) \dots f(x_n; \theta), \quad \theta \in \Omega \tag{2.27}$$

Misalkan dapat ditemukan suatu fungsi nontrivial dari x_1, x_2, \dots, x_n , yaitu $u(x_1, x_2, \dots, x_n)$, sedemikian sehingga jika θ diganti dengan $u(x_1, x_2, \dots, x_n)$, maka fungsi *likelihood* L akan bernilai maksimum, yaitu $L[u(x_1, x_2, \dots, x_n); x_1, x_2, \dots, x_n]$ sedikitnya sebesar $L(\theta; x_1, x_2, \dots, x_n)$ untuk setiap $\theta \in \Omega$. Statistik $u(X_1, X_2, \dots, X_n)$ disebut penaksir maksimum *likelihood* (m.l.e. : *maximum likelihood estimator*) dari θ dan dinotasikan dengan simbol:

$$\hat{\theta} = u(X_1, X_2, \dots, X_n).$$

(Hogg and Craig, 1995 : 262)

Misalkan terdapat k parameter yang tidak diketahui, maka penaksir maksimum *likelihood* untuk θ_i diperoleh dengan menyelesaikan

$$\frac{\partial L(\theta_1, \theta_2, \dots, \theta_k)}{\partial \theta_i} = 0 \quad (2.28)$$

atau melalui bentuk logaritma natural

$$\frac{\partial \ln L(\theta_1, \theta_2, \dots, \theta_k)}{\partial \theta_i} = 0 \quad (2.29)$$

untuk $i = 1, 2, \dots, k$.

2.9 Metode *Newton-Raphson*

Metode *Newton-Raphson* adalah salah satu metode numerik yang paling sering digunakan untuk menyelesaikan masalah pencarian akar dari suatu fungsi $f(x) = 0$ dengan menggunakan suatu titik awal, kemudian mendekatinya dengan memperhatikan *slope* atau kemiringan dari fungsi $f(x)$ di titik tersebut.

Langkah-langkah dari metode *Newton-Raphson* dimulai dengan menentukan aproksimasi awal p_0 dan membentuk barisan $\{p_n\}_{n=0}^{\infty}$, di mana

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}, \text{ untuk } n \geq 1 \quad (2.30)$$

Langkah-langkah metode *Newton-Raphson* di atas digunakan untuk mencari akar dari satu persamaan dengan satu variabel. Jika $F(\mathbf{x})$ adalah sistem persamaan yang terdiri dari dua atau lebih variabel, yaitu

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{bmatrix},$$

maka langkah-langkah dari metode *Newton-Raphson* untuk mencari solusi \mathbf{x} sehingga $\mathbf{F}(\mathbf{x}) = \mathbf{0}$, dimulai dengan menentukan aproksimasi awal $\mathbf{x}^{(0)}$, kemudian pada setiap $r \geq 1$, hitung

$$\mathbf{x}^{(r)} = \mathbf{x}^{(r-1)} - \{\mathbf{J}(\mathbf{x}^{(r-1)})\}^{-1} \mathbf{F}(\mathbf{x}^{(r-1)}) \quad (2.31)$$

di mana

$\mathbf{x}^T = [x_1 \ \dots \ x_n]$, dengan \mathbf{x}^T adalah transpos dari matriks \mathbf{x}

$$\mathbf{F}^T(\mathbf{x}) = [f_1(\mathbf{x}) \ \dots \ f_n(\mathbf{x})]$$

$$\mathbf{J}(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f_1(\mathbf{x})}{\partial x_j} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_j(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f_j(\mathbf{x})}{\partial x_j} \end{bmatrix}$$

(Burden and Faires, 2001 : 613)

Walaupun metode *Newton-Raphson* sering digunakan, metode ini juga memiliki kelemahan, yaitu bagus atau tidaknya hasil aproksimasi yang diperoleh bergantung pada ketepatan nilai dari taksiran awal yang ditentukan.

BAB 3

MODEL PARAMETRIK WEIBULL AFT

Model parametrik adalah salah satu metode yang dapat digunakan untuk mengetahui pengaruh kovariat terhadap waktu *survival*. Pada model parametrik, waktu *survival* diasumsikan mengikuti suatu distribusi tertentu. Distribusi yang biasa digunakan untuk waktu *survival* antara lain Weibull, eksponensial (bentuk khusus dari Weibull), log-logistik, lognormal, dan *generalized gamma*.

Asumsi *accelerated failure time* (AFT) seringkali digunakan sebagai asumsi dasar model parametrik, yaitu efek kovariat diasumsikan linier terhadap log waktu. Selain itu, asumsi *proportional hazard* (PH) seperti pada model Cox PH juga dapat digunakan pada beberapa model parametrik.

Pada bab ini akan dibahas model parametrik dengan waktu *survival* diasumsikan berdistribusi Weibull. Model parametrik Weibull ini memiliki sifat unik, yaitu satu-satunya model parametrik yang dapat memenuhi asumsi AFT dan juga PH (Klein & Moeschberger, 1997).

3.1 Model Parametrik Weibull

Model Weibull adalah model parametrik *survival* yang paling sering digunakan (Kleinbaum & Klein, 2005). Model Weibull memiliki sifat unik, yaitu jika asumsi AFT terpenuhi maka asumsi PH juga terpenuhi (dan *vice versa*). Sifat tersebut berlaku jika *shape parameter* p tidak berubah untuk setiap nilai dari kovariat.

Berikut ini akan ditunjukkan sifat model Weibull tersebut. Secara matematis, asumsi PH dinyatakan dengan

$$\frac{h_1(x)}{h_2(x)} = c \quad (3.1)$$

di mana $h_1(x)$ adalah *hazard* dari individu pertama, $h_2(x)$ adalah *hazard* dari individu kedua, dan c adalah suatu konstanta.

Asumsi AFT secara matematis dinyatakan dengan

$$S_1(x) = S_2(\gamma x) \quad (3.2)$$

di mana $S_1(x)$ adalah fungsi *survival* dari individu pertama, $S_2(x)$ adalah fungsi *survival* dari individu kedua, dan γ adalah suatu konstanta yang disebut sebagai *accelerated factor*.

Pertama, akan dilihat untuk model Weibull dengan nilai p yang konstan, jika asumsi AFT terpenuhi maka asumsi PH juga terpenuhi. Berdasarkan asumsi AFT dan bentuk fungsi *survival* dari distribusi Weibull pada persamaan (2.21), yaitu

$$S(x) = \exp[-\lambda x^p]$$

dan bentuk fungsi *hazard* dari distribusi Weibull pada persamaan (2.22), yaitu

$$h(x) = \lambda p x^{p-1},$$

maka diperoleh

$$\begin{aligned} S_1(x) &= S_2(\gamma x) \\ \exp[-\lambda_1 x^p] &= \exp[-\lambda_2 (\gamma x)^p] \\ -\lambda_1 x^p &= -\lambda_2 (\gamma x)^p \\ \lambda_1 &= \lambda_2 \gamma^p. \end{aligned}$$

Kemudian dengan menghitung *hazard ratio* dari model Weibull dan mensubstitusi persamaan di atas, diperoleh

$$\begin{aligned} \frac{h_1(x)}{h_2(x)} &= \frac{\lambda_1 p x^{p-1}}{\lambda_2 p x^{p-1}} \\ \frac{h_1(x)}{h_2(x)} &= \frac{\lambda_2 \gamma^p p x^{p-1}}{\lambda_2 p x^{p-1}} \\ \frac{h_1(x)}{h_2(x)} &= \gamma^p = c \end{aligned}$$

Terbukti bahwa untuk model Weibull dengan nilai p yang konstan, jika asumsi AFT terpenuhi maka asumsi PH juga terpenuhi.

Selanjutnya akan ditunjukkan untuk model Weibull dengan nilai p yang konstan, jika asumsi PH terpenuhi maka asumsi AFT juga terpenuhi. Berdasarkan bentuk fungsi *hazard* dari distribusi Weibull pada persamaan (3.3), maka

$$\begin{aligned} \frac{h_1(x)}{h_2(x)} &= c \\ \frac{\lambda_1 p x^{p-1}}{\lambda_2 p x^{p-1}} &= c \\ \lambda_1 &= \lambda_2 c. \end{aligned}$$

Kemudian, substitusi persamaan tersebut ke dalam fungsi *survival* dari model Weibull pada persamaan (2.21), sehingga diperoleh

$$\begin{aligned} S_1(x) &= \exp(-\lambda_1 x^p) \\ &= \exp(-c\lambda_2 x^p) \\ &= \exp\left[-\left(c^{\frac{1}{p}}\right)^p \lambda_2 x^p\right] \\ &= \exp\left[-\lambda_2 \left(c^{\frac{1}{p}}x\right)^p\right] \\ &= \exp[-\lambda_2(\gamma x)^p] \end{aligned}$$

$$S_1(x) = S_2(\gamma x)$$

dengan $\gamma = c^{1/p}$. Terbukti bahwa untuk model Weibull dengan nilai p yang konstan, jika asumsi PH terpenuhi maka asumsi AFT juga terpenuhi. Secara keseluruhan terbukti bahwa sifat pada model Weibull, yaitu jika nilai dari p konstan maka berlaku asumsi PH terpenuhi jika dan hanya jika asumsi AFT terpenuhi.

Sifat khusus pada model Weibull lainnya adalah log negatif log dari fungsi *survival* ($\ln[-\ln S(x)]$) linier terhadap log waktu ($\ln x$). Sifat ini memungkinkan untuk melihat kecocokan penggunaan model Weibull melalui grafik dengan memplot $\ln[-\ln S(x)]$ melalui estimasi fungsi *survival* dari Kaplan-Meier terhadap $\ln x$. Hubungan linier tersebut dapat dilihat sebagai berikut.

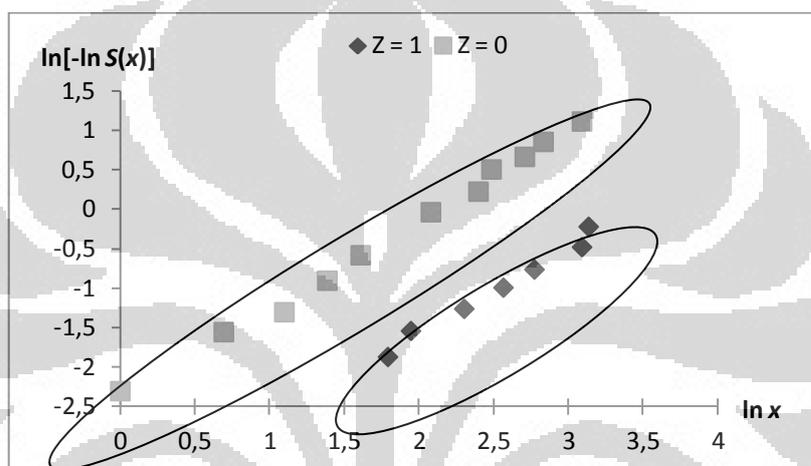
$$\begin{aligned} S(x) &= \exp(-\lambda x^p) \\ \ln S(x) &= -\lambda x^p \\ -\ln S(x) &= \lambda x^p \\ \ln[-\ln S(x)] &= \ln \lambda + p \ln x \end{aligned}$$

Dari persamaan di atas didapat bahwa $\ln[-\ln(S(x))]$ adalah fungsi linier dari $\ln(x)$ dengan *slope* p dan *intercept* $\ln(\lambda)$. Jika *slope* bernilai satu maka x mengikuti distribusi eksponensial.

Berdasarkan hubungan linier tersebut, jika grafik dari $\ln[-\ln S(x)]$ terhadap $\ln x$ menghasilkan garis yang lurus (linier) maka dapat disimpulkan bahwa data waktu *survival* tersebut berdistribusi Weibull.

Misalkan hanya diamati satu kovariat yang terdiri dari dua level (contohnya kovariat jenis kelamin yang terdiri dari laki-laki dan perempuan), maka grafik

akan menghasilkan dua buah garis yang merepresentasikan kedua kelompok individu berdasarkan kovariatnya tersebut, seperti yang terlihat pada gambar 3.1 di bawah ini. Jika grafik yang diperoleh berupa garis-garis yang sejajar maka asumsi PH terpenuhi, karena garis-garis yang sejajar menunjukkan bahwa nilai dari p sama untuk setiap nilai dari kovariat. Jika data tersebut berdistribusi Weibull dan nilai p konstan maka selain asumsi PH, berdasarkan sifat model Weibull, asumsi AFT juga terpenuhi.



Gambar 3.1. Ilustrasi grafik hubungan $\ln[-\ln S(x)]$ dengan $\ln x$

Berikut ini akan dibahas kedua bentuk model parametrik Weibull dengan asumsi PH dan AFT.

3.1.1 Model Weibull PH

Model Weibull PH merupakan bentuk parametrik dari model Cox PH dengan fungsi *baseline hazard* mengikuti bentuk fungsi *hazard* dari distribusi Weibull. Misalkan X adalah waktu terjadinya suatu *event* dan Z adalah vektor kovariat yang mempengaruhi waktu *survival*. Jika X berdistribusi Weibull, maka, dari persamaan (2.22), fungsi *hazard*-nya adalah

$$h(x) = \lambda p x^{p-1}.$$

Berdasarkan model Cox PH,

$$h(x|Z) = h_0(x) \exp(\beta^t Z),$$

maka dengan mensubstitusi fungsi *baseline hazard* $h_0(x)$ pada model Cox PH dengan fungsi *hazard* dari distribusi Weibull diperoleh model Weibull PH

$$h(x|\mathbf{Z}) = \lambda px^{p-1} \exp(\boldsymbol{\beta}^t \mathbf{Z}) \quad (3.3)$$

di mana $\boldsymbol{\beta}^t = (\beta_1, \dots, \beta_m)$ adalah vektor koefisien regresi yang mengukur besarnya pengaruh kovariat terhadap *hazard*.

Bentuk lain dari model Weibull PH juga dapat didefinisikan dengan memparameterisasi ulang λ sebagai fungsi eksponensial dari kovariat, yaitu

$$\lambda = \exp(\beta_0 + \beta_1 Z_1 + \dots + \beta_m Z_m),$$

sehingga diperoleh model Weibull PH

$$\begin{aligned} h(x|\mathbf{Z}) &= \lambda px^{p-1} \\ &= px^{p-1} \exp(\beta_0 + \beta_1 Z_1 + \dots + \beta_m Z_m) \\ &= \exp(\beta_0) px^{p-1} \exp(\boldsymbol{\beta}^t \mathbf{Z}). \end{aligned} \quad (3.4)$$

Bentuk model Weibull PH tersebut sama dengan model Weibull PH pada persamaan sebelumnya (3.3), dengan fungsi *baseline hazard* didefinisikan secara parametrik oleh

$$h_0^*(x) = \exp(\beta_0) px^{p-1}$$

Dari bentuk fungsi *hazard* model Weibull PH tersebut, dapat dilihat bahwa model tersebut memenuhi asumsi PH, yaitu *hazard ratio* konstan sepanjang waktu, seperti berikut. Misalkan individu A memiliki *hazard rate* $h_A(x|\mathbf{Z})$ di mana

$$\mathbf{Z} = (Z_1, Z_2, \dots, Z_m)$$

dan individu B memiliki *hazard rate* $h_B(x|\mathbf{Z}^*)$ di mana

$$\mathbf{Z}^* = (Z_1^*, Z_2^*, \dots, Z_m^*)$$

maka dengan asumsi bahwa p konstan untuk setiap nilai dari kovariat, *hazard ratio*-nya adalah

$$\begin{aligned} \text{HR} &= \frac{h_A(x|\mathbf{Z})}{h_B(x|\mathbf{Z}^*)} = \frac{\exp(\beta_0) px^{p-1} \exp[\sum_{k=1}^m \beta_k Z_k]}{\exp(\beta_0) px^{p-1} \exp[\sum_{k=1}^m \beta_k Z_k^*]} \\ &= \exp\left(\sum_{k=1}^m \beta_k Z_k - \sum_{k=1}^m \beta_k Z_k^*\right) \\ &= \exp\left[\sum_{k=1}^m \beta_k (Z_k - Z_k^*)\right] \end{aligned} \quad (3.5)$$

Hazard ratio tersebut tidak bergantung pada variabel waktu x sehingga nilainya konstan sepanjang waktu.

Model Weibull PH juga dapat dinyatakan dalam bentuk fungsi *survival*, yaitu

$$\begin{aligned}
 S(x|\mathbf{Z}) &= \exp\left[-\int_0^x h(t|\mathbf{Z}) dt\right] \\
 &= \exp\left[-\int_0^x \lambda p t^{p-1} \exp(\boldsymbol{\beta}^t \mathbf{Z}) dt\right] \\
 &= \exp\left[-\lambda t^p \exp(\boldsymbol{\beta}^t \mathbf{Z}) \Big|_0^x\right] \\
 &= \exp[-\lambda x^p \exp(\boldsymbol{\beta}^t \mathbf{Z})]
 \end{aligned} \tag{3.6}$$

Tabel 3.1. Bentuk Model Parametrik Weibull PH

<i>Response</i>	Model Weibull PH
Fungsi <i>hazard</i>	$h(x \mathbf{Z}) = \lambda p x^{p-1} \exp(\boldsymbol{\beta}^t \mathbf{Z})$
Fungsi <i>survival</i>	$S(x \mathbf{Z}) = \exp[-\lambda x^p \exp(\boldsymbol{\beta}^t \mathbf{Z})]$

Distribusi Weibull adalah salah satu dari sedikit distribusi yang dapat digunakan dalam model parametrik PH. Pada distribusi waktu yang lain, seringkali model parametriknya tidak memenuhi asumsi PH, sehingga tidak banyak distribusi yang memungkinkan untuk digunakan pada model parametrik PH (Qi, 2009). Contohnya pada distribusi log-logistik yang memiliki fungsi *hazard*

$$h_0(x) = \frac{\lambda p x^{p-1}}{1 + \lambda x^p} \tag{3.7}$$

Dengan parametrisasi ulang $\lambda = \exp(\beta_0 + \beta_1 Z_1 + \dots + \beta_m Z_m)$, maka *hazard ratio* untuk individu A dengan *hazard rate* $h_A(x|\mathbf{Z})$ dan individu B dengan *hazard rate* $h_B(x|\mathbf{Z}^*)$ adalah

$$\begin{aligned}
 \text{HR} &= \frac{h_A(x|\mathbf{Z})}{h_B(x|\mathbf{Z}^*)} = \frac{\frac{\exp(\beta_0) p x^{p-1} \exp[\sum_{k=1}^m \beta_k Z_k]}{1 + \exp(\beta_0) x^p \exp[\sum_{k=1}^m \beta_k Z_k]}}{\frac{\exp(\beta_0) p x^{p-1} \exp[\sum_{k=1}^m \beta_k Z_k^*]}{1 + \exp(\beta_0) x^p \exp[\sum_{k=1}^m \beta_k Z_k^*]}} \\
 &= \exp\left[\sum_{k=1}^m \beta_k (Z_k - Z_k^*)\right] \frac{1 + \exp(\beta_0) x^p \exp[\sum_{k=1}^m \beta_k Z_k^*]}{1 + \exp(\beta_0) x^p \exp[\sum_{k=1}^m \beta_k Z_k]}
 \end{aligned}$$

Universitas Indonesia

Hazard ratio yang dihasilkan nilainya tergantung pada waktu x sehingga model parametrik dengan waktu *survival* yang berdistribusi log-logistik tidak memenuhi asumsi PH.

3.1.2 Model Weibull AFT

Misalkan variabel random waktu *survival* X berdistribusi Weibull dengan parameter $\lambda > 0$ dan $p > 0$. Dari persamaan (2.21), fungsi *survival* untuk distribusi Weibull diberikan oleh

$$S_X(x) = \exp(-\lambda x^p); \quad x \geq 0.$$

Seperti yang sudah dijelaskan pada subbab 2.7, pada model AFT diasumsikan bahwa model linier terhadap log waktu. Model linier untuk log waktu yang biasa digunakan adalah

$$Y = \ln X = \mu + \alpha^t \mathbf{Z} + \sigma W$$

di mana α adalah vektor parameter regresi, μ adalah *intercept*, σ adalah *scale parameter*, dan W adalah distribusi error.

Berikut ini akan dibahas terlebih dahulu model parametrik Weibull AFT tanpa memasukkan efek kovariat. Dengan transformasi log waktu, fungsi *survival* untuk $Y = \ln X$ adalah

$$S_Y(y) = \exp(-\lambda e^{yp}). \quad (3.8)$$

Persamaan tersebut kemudian ditransformasi sehingga didapat persamaan yang linier dari log waktu sebagai berikut.

$$S_Y(y) = \exp(-\lambda e^{yp})$$

$$\ln S_Y(y) = -\lambda e^{yp}$$

$$-\ln S_Y(y) = \lambda e^{yp}$$

$$\ln[-\ln S_Y(y)] = \ln \lambda + yp$$

misalkan $w = \ln[-\ln S_Y(y)]$, maka

$$w = \ln \lambda + yp$$

$$y = -\frac{\ln \lambda}{p} + \frac{w}{p}$$

Misalkan didefinisikan parameter baru yaitu σ dan μ , di mana $\sigma = 1/p$ dan $\mu = -(\ln \lambda)/p$, sehingga diperoleh

$$y = \mu + \sigma w.$$

Y mengikuti bentuk log linier model dengan

$$Y = \ln X = \mu + \sigma W, \quad (3.9)$$

di mana W adalah *error* yang memiliki bentuk fungsi *survival*

$$\begin{aligned} S_W(w) &= \Pr(W > w) \\ &= \Pr\left(\frac{Y - \mu}{\sigma} > w\right) \\ &= \Pr(Y > \mu + \sigma w) \\ &= S_Y(\mu + \sigma w) \\ &= \exp[-\lambda \exp((\mu + \sigma w)p)] \\ &= \exp\left[-\exp(-\mu/\sigma) \exp\left((\mu + \sigma w)\frac{1}{\sigma}\right)\right] \\ &= \exp(-e^w) \end{aligned} \quad (3.10)$$

dan bentuk *pdf*

$$\begin{aligned} f_W(w) &= \frac{-d}{dw} S_W(w) \\ &= \frac{-d}{dw} [\exp(-e^w)] \\ &= e^w \exp(-e^w) \\ &= \exp(w - e^w) \end{aligned} \quad (3.11)$$

Dari bentuk *pdf* variabel error W tersebut, maka W berdistribusi Gumbel (yaitu salah satu bentuk dari distribusi *extreme value*).

Untuk menggabungkan efek kovariat ke dalam model Weibull AFT, *scale parameter* λ diparameterisasi ulang sebagai fungsi dari kovariat, yaitu

$$\lambda^* = \exp(b_0 + b_1 Z_1 + \dots + b_m Z_m)$$

di mana Z_k adalah kovariat dan b_k adalah koefisien regresi, untuk $k = 1, 2, \dots, m$.

Dengan menganggap λ sebagai *baseline* dari *scale* parameter yang tidak dipengaruhi oleh efek kovariat, maka konstanta $\exp(b_0)$ dapat dinyatakan sebagai nilai dari λ , sehingga λ^* dapat dinyatakan ulang dalam bentuk

$$\begin{aligned} \lambda^* &= \exp(b_0) \exp(b_1 Z_1 + \dots + b_m Z_m) \\ &= \lambda \exp(b_1 Z_1 + \dots + b_m Z_m) \\ &= \lambda \exp(\mathbf{b}^t \mathbf{Z}) \end{aligned}$$

Oleh karena itu, berdasarkan persamaan (3.8), fungsi *survival* dari Y adalah

$$S_Y(y) = \exp(-\lambda \exp(\mathbf{b}^t \mathbf{Z}) e^{yp})$$

Persamaan tersebut kemudian ditransformasi sehingga didapat persamaan yang linier dari log waktu sebagai berikut:

$$S_Y(y) = \exp[-\lambda \exp(\mathbf{b}^t \mathbf{Z}) e^{yp}]$$

$$\ln S_Y(y) = -\lambda \exp(\mathbf{b}^t \mathbf{Z}) e^{yp}$$

$$-\ln S_Y(y) = \lambda \exp(\mathbf{b}^t \mathbf{Z}) e^{yp}$$

$$\ln[-\ln S_Y(y)] = \ln \lambda + \mathbf{b}^t \mathbf{Z} + yp$$

misalkan $w = \ln[-\ln S_Y(y)]$, maka

$$w = \ln \lambda + \mathbf{b}^t \mathbf{Z} + yp$$

$$y = -\frac{\ln \lambda}{p} - \left(\frac{b_1}{p} Z_1 + \dots + \frac{b_m}{p} Z_m \right) + \frac{w}{p}$$

Misalkan didefinisikan parameter baru yaitu σ dan μ , di mana $\sigma = 1/p$, $\mu = -(\ln \lambda)/p$, dan $\alpha_k = -b_k/p$, $k = 1, 2, \dots, m$, sehingga diperoleh

$$y = \mu + \boldsymbol{\alpha}^t \mathbf{Z} + \sigma w$$

Y mengikuti bentuk log linier model dengan

$$Y = \ln X = \mu + \boldsymbol{\alpha}^t \mathbf{Z} + \sigma W, \quad (3.12)$$

di mana W adalah *error* yang berdistribusi Gumbel dan memiliki bentuk fungsi *survival* dan *pdf* seperti pada persamaan (3.10) dan (3.11).

Fungsi *survival* untuk Y adalah

$$\begin{aligned} S_Y(y) &= \Pr(Y > y) \\ &= \Pr(\mu + \boldsymbol{\alpha}^t \mathbf{Z} + \sigma W > y) \\ &= \Pr\left(W > \frac{y - \mu - \boldsymbol{\alpha}^t \mathbf{Z}}{\sigma}\right) \\ &= S_W\left(\frac{y - \mu - \boldsymbol{\alpha}^t \mathbf{Z}}{\sigma}\right) \end{aligned} \quad (3.13)$$

Berdasarkan fungsi *survival* dari W pada persamaan (3.9), maka diperoleh

$$S_Y(y) = \exp(-e^{[(y - \mu - \boldsymbol{\alpha}^t \mathbf{Z})/\sigma]}). \quad (3.14)$$

Bentuk *pdf* dari Y adalah

$$\begin{aligned} f_Y(y) &= \frac{-d}{dy} S_Y(y) \\ &= \frac{-d}{dy} [\exp(-e^{[(y - \mu - \boldsymbol{\alpha}^t \mathbf{Z})/\sigma])}] \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\sigma} e^{[(y-\mu-\alpha^t \mathbf{Z})/\sigma]} \exp(-e^{[(y-\mu-\alpha^t \mathbf{Z})/\sigma]}) \\
&= (1/\sigma) \exp([(y-\mu-\alpha^t \mathbf{Z})/\sigma] - e^{[(y-\mu-\alpha^t \mathbf{Z})/\sigma]}) \quad (3.15)
\end{aligned}$$

Berikut ini akan dibandingkan bentuk model Weibull PH dan AFT yang menggunakan fungsi *survival* dan fungsi *hazard*. Hal ini dilakukan untuk menunjukkan kesamaan bentuk model parametrik Weibull dengan asumsi PH maupun asumsi AFT.

❖ Perbandingan model Weibull PH dan AFT yang menggunakan fungsi *survival*

Dari persamaan (3.14), fungsi *survival* untuk model Weibull AFT dapat ditulis ulang dalam bentuk

$$\begin{aligned}
S_Y(y) &= \exp \left[- \exp \left(\frac{y - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \\
S_X(x) &= \exp \left[- \exp \left(\frac{\ln x - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \\
&= \exp \left[- \exp \left(\frac{-\mu - \alpha^t \mathbf{Z}}{\sigma} \right) x^{1/\sigma} \right] \quad (3.16)
\end{aligned}$$

Perhatikan kembali persamaan fungsi *survival* dari model Weibull PH dari persamaan (3.6)

$$S(x|\mathbf{Z}) = \exp[-\lambda x^p \exp(\boldsymbol{\beta}^t \mathbf{Z})]$$

Bandingkan kedua persamaan di atas, terlihat bahwa model PH dapat dinyatakan dengan parameter μ , σ , dan α_k pada model AFT dengan

$$\begin{aligned}
\lambda &= \exp(-\mu/\sigma) \\
p &= 1/\sigma \\
\beta_k &= -\alpha_k/\sigma \quad (3.17)
\end{aligned}$$

di mana α_k dan β_k adalah koefisien regresi, dengan $k = 1, 2, \dots, m$.

❖ Perbandingan model Weibull PH dan AFT yang menggunakan fungsi *hazard*

Dari persamaan (2.24), fungsi *hazard* dari model AFT adalah

$$h(x|\mathbf{Z}) = \exp(\boldsymbol{\theta}^t \mathbf{Z}) h_0[\exp(\boldsymbol{\theta}^t \mathbf{Z}) x]$$

dengan *baseline hazard* mengikuti distribusi Weibull, maka $h_0(x) = \lambda p x^{p-1}$, sehingga diperoleh fungsi *hazard* dari model Weibull AFT adalah

$$\begin{aligned}
 h(x|\mathbf{Z}) &= \exp(\boldsymbol{\theta}^t \mathbf{Z}) \lambda p (\exp(\boldsymbol{\theta}^t \mathbf{Z}) x)^{p-1} \\
 &= \lambda p x^{p-1} \exp[p(\boldsymbol{\theta}^t \mathbf{Z})].
 \end{aligned}
 \tag{3.18}$$

Dari persamaan (3.3), fungsi *hazard* dari model Weibull PH adalah

$$h(x|\mathbf{Z}) = \lambda p x^{p-1} \exp(\boldsymbol{\beta}^t \mathbf{Z})$$

Model PH tersebut sama dengan model AFT dengan $\boldsymbol{\beta} = p\boldsymbol{\theta}$ dan dari persamaan (3.17) diperoleh

$$\boldsymbol{\theta} = \frac{\boldsymbol{\beta}}{p} = \frac{\boldsymbol{\beta}}{1/\sigma} = \boldsymbol{\beta}\sigma = -\boldsymbol{\alpha}. \tag{3.19}$$

Tabel 3.2. Perbandingan Bentuk Model Weibull PH dan Model Weibull AFT

<i>Response</i>	Model Weibull PH	Model Weibull AFT
Fungsi <i>survival</i>	$\exp[\lambda x^p \exp(\boldsymbol{\beta}^t \mathbf{Z})]$	$\exp\left[-\exp\left(\frac{-\mu - \boldsymbol{\alpha}^t \mathbf{Z}}{\sigma}\right) x^{1/\sigma}\right]$
Parameter	$\lambda = \exp(-\mu/\sigma),$ $p = 1/\sigma, \beta_k = -\alpha_k/\sigma$	μ, σ, α_k
Fungsi <i>hazard</i>	$\lambda p x^{p-1} \exp(\boldsymbol{\beta}^t \mathbf{Z})$	$\lambda p x^{p-1} \exp[p(\boldsymbol{\theta}^t \mathbf{Z})]$
Parameter	$\lambda, p, \beta_k = p\theta_k$	λ, p, θ_k

Dari pembahasan tersebut diperoleh bahwa model parametrik Weibull PH memiliki bentuk yang sama dengan model parametrik Weibull AFT. Parameter dari model Weibull PH dapat dinyatakan sebagai fungsi dari parameter pada model Weibull AFT. Oleh karena itu, dengan memperoleh hasil penaksiran parameter pada model Weibull AFT, maka taksiran dari model Weibull PH juga dapat diperoleh.

3.2 Estimasi Parameter Model Parametrik Weibull AFT

Untuk mengestimasi parameter pada model parametrik Weibull AFT digunakan metode *maximum likelihood*. Sebelum membahas hal tersebut, akan dibahas terlebih dahulu bagaimana mengkonstruksi fungsi *likelihood* pada data *survival*.

3.2.1 Konstruksi Fungsi *Likelihood* untuk Data *Survival*

Dalam analisis *survival* dengan data yang tersensor, konstruksi fungsi *likelihood* berbeda dari fungsi *likelihood* yang biasa. Asumsi utama yang dibutuhkan untuk mengkonstruksi fungsi *likelihood* untuk data *survival* yang tersensor adalah waktu hidup dan waktu sensor saling bebas. Jika asumsi tersebut tidak terpenuhi, maka dibutuhkan teknik lainnya.

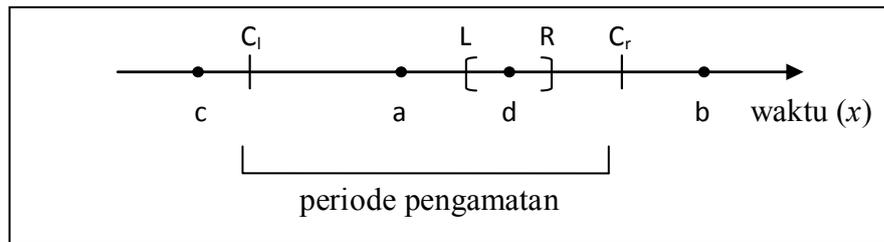
Misalkan X adalah waktu kejadian dari suatu *event*. Dalam mengkonstruksi fungsi *likelihood* untuk data tersensor, harus diperhatikan informasi yang diberikan dari setiap observasi. Pengamatan dengan waktu kejadian yang diketahui pasti dapat memberikan informasi mengenai probabilitas bahwa *event* tersebut terjadi pada waktu tersebut. Oleh karena itu, *probability density function (pdf)* dari X pada waktu x tertentu, $f(x)$, memberikan kontribusi pada fungsi *likelihood* untuk data waktu yang eksak.

Untuk data observasi yang tersensor kanan, diketahui bahwa waktu terjadinya suatu *event* lebih besar dari suatu waktu tertentu x , $X > x$, artinya *survival time* lebih besar dari x . Untuk itu, fungsi *survival* berkontribusi pada fungsi *likelihood* untuk data tersensor kanan. Secara matematis, informasi yang diberikan oleh observasi tersebut adalah $P(X > x) = S(x)$.

Untuk data observasi yang tersensor kiri, diketahui bahwa *event* tersebut sudah terjadi sebelum waktu pengamatan, sehingga kontribusi untuk fungsi *likelihood* adalah fungsi distribusi kumulatif. Secara matematis, informasi yang diberikan oleh observasi tersebut adalah $P(X < x) = F(x)$ atau dapat juga dinyatakan sebagai fungsi *survival* yaitu $1 - S(x)$.

Untuk data yang tersensor interval, hanya diketahui bahwa suatu *event* terjadi pada selang waktu tertentu. Oleh karena itu, maka informasi yang diberikan adalah probabilitas bahwa waktu kejadian berada dalam interval waktu tersebut. Misalkan waktu *survival* tersebut berada di dalam suatu interval waktu $[L, R]$. Informasi yang diberikan oleh observasi tersebut dapat dinyatakan dengan $\Pr(L < X < R) = S(L) - S(R)$.

Untuk memperjelas uraian di atas, pada gambar 3.2 berikut ini diberikan gambaran titik-titik waktu *survival* yang diperoleh dari suatu pengamatan.



Gambar 3.2. Garis waktu periode pengamatan

Periode pengamatan dimulai dari waktu C_l sampai dengan C_r , dengan a adalah waktu pengamatan yang eksak, b adalah waktu pengamatan yang tersensor kanan, c adalah waktu pengamatan yang tersensor kiri, dan d adalah waktu pengamatan yang tersensor interval pada selang waktu L dan R . Berdasarkan penjelasan sebelumnya, maka kontribusi *likelihood* dari keempat jenis observasi tersebut dapat diringkas sebagai berikut:

- (a). Waktu kejadian eksak : $\Pr(X = x) = f(x)$
- (b). Observasi tersensor kanan : $\Pr(X > C_r) = S(C_r)$
- (c). Observasi tersensor kiri : $\Pr(X < C_l) = F(C_l) = 1 - S(C_l)$
- (d). Observasi tersensor interval : $\Pr(L < X < R) = S(L) - S(R)$

Misalkan terdapat n buah pengamatan dengan n_1 pengamatan yang terobservasi, n_2 pengamatan yang tersensor kanan, n_3 pengamatan yang tersensor kiri, dan n_4 pengamatan yang tersensor interval, dengan $n_1 + n_2 + n_3 + n_4 = n$. Dengan asumsi bahwa setiap pengamatan saling bebas, maka fungsi *likelihood* dapat dikonstruksi dengan cara menggabungkan setiap komponen observasi menjadi

$$L = \prod_{i \in D} f(x_i) \prod_{i \in R} S(C_{r_i}) \prod_{i \in L} [1 - S(C_{l_i})] \prod_{i \in I} [S(L_i) - S(R_i)], \quad (3.20)$$

di mana D adalah himpunan waktu kejadian yang eksak, R adalah himpunan observasi yang tersensor kanan, L adalah himpunan observasi yang tersensor kiri, dan I adalah himpunan observasi yang tersensor interval.

Berikut ini akan lebih dijabarkan konstruksi fungsi *likelihood* khusus untuk data *survival* yang hanya mengandung observasi tersensor kanan. Misalkan

terdapat n buah pengamatan dengan n_1 pengamatan yang terobservasi dan n_2 pengamatan yang tersensor kanan, di mana $n_1 + n_2 = n$. Data dari suatu *event* yang terdapat observasi tersensor kanan dapat dinyatakan dalam variabel random berpasangan (T, δ) , di mana δ mengindikasikan apakah waktu kejadian X terobservasi ($\delta = 1$) atau tidak ($\delta = 0$). Sedangkan T sama dengan X jika waktu kejadian terobservasi dan sama dengan C_r jika tersensor kanan, atau dengan kata lain $T = \min(X, C_r)$.

Detail dari konstruksi fungsi *likelihood* untuk data yang hanya terdapat sensor kanan adalah sebagai berikut. Untuk $\delta = 0$, dapat dilihat bahwa

$$\begin{aligned}\Pr[T, \delta = 0] &= \Pr[T = X \mid \delta = 0] \Pr[\delta = 0] + \Pr[T = C_r \mid \delta = 0] \Pr[\delta = 0] \\ &= \Pr[T = C_r \mid \delta = 0] \Pr[\delta = 0] \\ &= \Pr[\delta = 0] \\ &= \Pr(X > C_r) = S(C_r)\end{aligned}$$

dan juga untuk $\delta = 1$,

$$\begin{aligned}\Pr[T, \delta = 1] &= \Pr[T = X \mid \delta = 1] \Pr[\delta = 1] + \Pr[T = C_r \mid \delta = 1] \Pr[\delta = 1] \\ &= \Pr[T = X \mid \delta = 1] \Pr[\delta = 1] \\ &= \Pr[X = T \mid X \leq C_r] \Pr[X \leq C_r] \\ &= \left[\frac{f(t)}{1 - S(C_r)} \right] [1 - S(C_r)] = f(t)\end{aligned}$$

Kedua persamaan tersebut dapat digabungkan menjadi sebuah persamaan

$$\Pr(t, \delta) = [f(t)]^\delta [S(t)]^{1-\delta}.$$

Sehingga untuk sampel acak dengan $(T_i, \delta_i), i = 1, \dots, n$, fungsi *likelihood* adalah

$$L = \prod_{i=1}^n \Pr[t_i, \delta_i] = \prod_{i=1}^n [f(t_i)]^{\delta_i} [S(t_i)]^{1-\delta_i} \quad (3.21)$$

3.2.2 Estimasi Parameter Model Parametrik Weibull AFT dengan Metode *Maximum Likelihood*

Untuk mengestimasi parameter pada model parametrik Weibull AFT, digunakan metode *maximum likelihood*. Perhatikan kembali model linier log waktu

$$Y = \mu + \alpha^t \mathbf{Z} + \sigma W$$

di mana Y adalah log dari waktu atau $Y = \ln X$.

Dari persamaan (3.14) dan (3.15), Y memiliki fungsi *survival*

$$S_Y(y) = \exp(-e^{[(y-\mu-\alpha^t \mathbf{Z})/\sigma]})$$

dan bentuk *pdf*

$$f_Y(y) = (1/\sigma) \exp([(y-\mu-\alpha^t \mathbf{Z})/\sigma] - e^{[(y-\mu-\alpha^t \mathbf{Z})/\sigma]}).$$

Dari persamaan (3.10) dan (3.11), W memiliki fungsi *survival*

$$S_W(w) = \exp(-e^w)$$

dan bentuk *pdf*

$$f_W(w) = \exp(w - e^w).$$

Oleh karena itu, bentuk fungsi *survival* dan *pdf* dari Y dapat dinyatakan dalam fungsi *survival* dan *pdf* dari W , yaitu

$$S_Y(y) = S_W\left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma}\right)$$

dan

$$f_Y(y) = \frac{1}{\sigma} f_W\left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma}\right)$$

Berdasarkan persamaan (3.20), konstruksi fungsi *likelihood* untuk data *survival* adalah

$$L = \prod_{i \in D} f(x_i) \prod_{i \in R} S(C_{r_i}) \prod_{i \in L} [1 - S(C_{l_i})] \prod_{i \in I} [S(L_i) - S(R_i)]$$

Oleh karena itu, untuk menaksir parameter pada model tersebut, fungsi *likelihood* dikonstruksi sebagai berikut:

$$\begin{aligned} L &= \prod_{i \in D} f(y_i) \prod_{i \in R} S(y_i) \prod_{i \in L} [1 - S(y_i)] \prod_{i \in I} [S(y_{L_i}) - S(y_{R_i})] \\ &= \prod_{i \in D} \left[\frac{1}{\sigma} f_W\left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma}\right) \right] \prod_{i \in R} \left[S_W\left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma}\right) \right] \\ &\quad \prod_{i \in L} \left[1 - \left[S_W\left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma}\right) \right] \right] \\ &\quad \prod_{i \in I} \left[S_W\left(\frac{y_{L_i} - \mu - \alpha^t \mathbf{Z}}{\sigma}\right) - S_W\left(\frac{y_{R_i} - \mu - \alpha^t \mathbf{Z}}{\sigma}\right) \right] \end{aligned} \quad (3.22)$$

Dengan mensubstitusi fungsi *survival* dari W pada fungsi *likelihood* di atas, maka diperoleh fungsi *likelihood*

$$\begin{aligned}
 L = & \prod_{i \in D} \left(\frac{1}{\sigma} \right) \exp \left[\left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) - \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \\
 & \prod_{i \in R} \exp \left[- \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \prod_{i \in L} \left\{ 1 - \exp \left[- \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \right\} \\
 & \prod_{i \in I} \left\{ \exp \left[- \exp \left(\frac{y_{Li} - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] - \exp \left[- \exp \left(\frac{y_{Ri} - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \right\}
 \end{aligned} \tag{3.23}$$

Bentuk logaritma natural dari fungsi *likelihood* tersebut adalah

$$\begin{aligned}
 \ln L = & -n_1 \ln \sigma + \sum_{i \in D} \left[\left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) - \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \\
 & - \sum_{i \in R} \left[\exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] + \sum_{i \in L} \ln \left\{ 1 - \exp \left[- \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \right\} \\
 & + \sum_{i \in I} \ln \left\{ \exp \left[- \exp \left(\frac{y_{Li} - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] - \exp \left[- \exp \left(\frac{y_{Ri} - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \right\}
 \end{aligned} \tag{3.24}$$

Taksiran parameter μ diperoleh dengan cara menyelesaikan persamaan

$$\frac{\partial \ln L}{\partial \mu} = 0$$

yaitu

$$\begin{aligned}
 & \sum_{i \in D} \left[-\frac{1}{\sigma} + \frac{1}{\sigma} \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] + \sum_{i \in R} \left[\frac{1}{\sigma} \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \\
 & + \sum_{i \in L} \left\{ \frac{-\frac{1}{\sigma} \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \exp \left[- \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right]}{1 - \exp \left[- \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right]} \right\} \\
 & + \sum_{i \in I} \left\{ \frac{\frac{1}{\sigma} e^{(y_{Li} - \mu - \alpha^t \mathbf{Z})/\sigma} \exp \left[-e^{(y_{Li} - \mu - \alpha^t \mathbf{Z})/\sigma} \right] - \frac{1}{\sigma} e^{(y_{Ri} - \mu - \alpha^t \mathbf{Z})/\sigma} \exp \left[-e^{(y_{Ri} - \mu - \alpha^t \mathbf{Z})/\sigma} \right]}{\exp \left[- \exp \left(\frac{y_{Li} - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] - \exp \left[- \exp \left(\frac{y_{Ri} - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right]} \right\} = 0
 \end{aligned} \tag{3.25}$$

Taksiran parameter σ diperoleh dengan cara menyelesaikan persamaan

$$\frac{\partial \ln L}{\partial \sigma} = 0$$

yaitu

$$\begin{aligned} & -\frac{n_1}{\sigma} + \sum_{i \in D} \left[-\left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma^2} \right) + \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma^2} \right) \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \\ & + \sum_{i \in R} \left[\left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma^2} \right) \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \\ & + \sum_{i \in L} \left\{ \frac{-\left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma^2} \right) \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \exp \left[-\exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right]}{1 - \exp \left[-\exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right]} \right\} \\ & + \sum_{i \in I} \left\{ \frac{\left(\frac{y_{L_i} - \mu - \alpha^t \mathbf{Z}}{\sigma^2} \right) \exp \left(\frac{y_{L_i} - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \exp \left[-\exp \left(\frac{y_{L_i} - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right]}{\exp \left[-\exp \left(\frac{y_{L_i} - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] - \exp \left[-\exp \left(\frac{y_{R_i} - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right]} \right\} \\ & - \sum_{i \in I} \left\{ \frac{\left(\frac{y_{R_i} - \mu - \alpha^t \mathbf{Z}}{\sigma^2} \right) \exp \left(\frac{y_{R_i} - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \exp \left[-\exp \left(\frac{y_{R_i} - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right]}{\exp \left[-\exp \left(\frac{y_{L_i} - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] - \exp \left[-\exp \left(\frac{y_{R_i} - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right]} \right\} = 0 \end{aligned} \quad (3.26)$$

Taksiran parameter α_k untuk $k = 1, 2, \dots, m$ diperoleh dengan cara menyelesaikan persamaan

$$\frac{\partial \ln L}{\partial \alpha_k} = 0$$

untuk $k = 1, 2, \dots, m$, yaitu

$$\begin{aligned} & \sum_{i \in D} \left[-\frac{Z_k}{\sigma} + \frac{Z_k}{\sigma} \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] + \sum_{i \in R} \left[\frac{Z_k}{\sigma} \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \\ & + \sum_{i \in L} \left\{ \frac{-\frac{Z_k}{\sigma} \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \exp \left[-\exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right]}{1 - \exp \left[-\exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right]} \right\} \\ & + \sum_{i \in I} \left\{ \frac{\frac{Z_k}{\sigma} e^{(y_{L_i} - \mu - \alpha^t \mathbf{Z})/\sigma} \exp \left[-e^{(y_{L_i} - \mu - \alpha^t \mathbf{Z})/\sigma} \right] - \frac{Z_k}{\sigma} e^{(y_{R_i} - \mu - \alpha^t \mathbf{Z})/\sigma} \exp \left[-e^{(y_{R_i} - \mu - \alpha^t \mathbf{Z})/\sigma} \right]}{\exp \left[-\exp \left(\frac{y_{L_i} - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] - \exp \left[-\exp \left(\frac{y_{R_i} - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right]} \right\} = 0 \end{aligned} \quad (3.27)$$

Dari bentuk persamaan (3.25) sampai (3.27) terlihat bahwa estimasi μ , σ , dan α_k tidak dapat diselesaikan secara eksak, sehingga diperlukan penyelesaian secara numerik.

Pada tugas akhir ini, model parametrik Weibull digunakan pada data yang hanya mengandung skema tersensor kanan, untuk menyesuaikan penggunaan pada model Cox PH.

Untuk data yang tersensor kanan, dari persamaan (3.21) fungsi *likelihood*-nya adalah

$$L = \prod_{i=1}^n \Pr[t_i, \delta_i] = \prod_{i=1}^n [f(t_i)]^{\delta_i} [S(t_i)]^{1-\delta_i}$$

sehingga fungsi *likelihood* untuk menaksir model Weibull AFT untuk data yang tersensor kanan adalah

$$\begin{aligned} L &= \prod_{i=1}^n [f_Y(y_i)]^{\delta_i} [S_Y(y_i)]^{1-\delta_i} \\ &= \prod_{i=1}^n \left[\frac{1}{\sigma} f_W \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right]^{\delta_i} \left[S_W \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right]^{1-\delta_i} \end{aligned} \quad (3.28)$$

Dengan mensubstitusi fungsi *survival* dari W pada fungsi *likelihood* di atas, maka diperoleh fungsi *likelihood*

$$\begin{aligned} L &= \prod_{i=1}^n \left\{ \left(\frac{1}{\sigma} \right) \exp \left[\left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) - \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \right\}^{\delta_i} \\ &\quad \left\{ \exp \left[- \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \right\}^{1-\delta_i} \end{aligned} \quad (3.29)$$

Bentuk logaritma natural dari fungsi *likelihood* tersebut adalah

$$\begin{aligned} \ln L &= \sum_{i=1}^n \left[\delta_i \ln \left\{ \frac{1}{\sigma} \exp \left[\left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) - \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \right\} \right. \\ &\quad \left. + (1 - \delta_i) \ln \left\{ \exp \left[- \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \right\} \right] \\ \ln L &= - \sum_{i=1}^n \delta_i \ln \sigma + \sum_{i=1}^n \delta_i \left[\left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) - \exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \\ &\quad - \sum_{i=1}^n (1 - \delta_i) \left[\exp \left(\frac{y_i - \mu - \alpha^t \mathbf{Z}}{\sigma} \right) \right] \end{aligned} \quad (3.30)$$

Taksiran parameter μ diperoleh dengan cara menyelesaikan persamaan

$$\frac{\partial \ln L}{\partial \mu} = 0$$

yaitu

$$\sum_{i=1}^n \delta_i \left[-\frac{1}{\sigma} + \frac{1}{\sigma} \exp\left(\frac{y_i - \mu - \boldsymbol{\alpha}^t \mathbf{Z}}{\sigma}\right) \right] + \sum_{i=1}^n (1 - \delta_i) \left[\frac{1}{\sigma} \exp\left(\frac{y_i - \mu - \boldsymbol{\alpha}^t \mathbf{Z}}{\sigma}\right) \right] = 0 \quad (3.31)$$

Taksiran parameter σ diperoleh dengan cara menyelesaikan persamaan

$$\frac{\partial \ln L}{\partial \sigma} = 0$$

yaitu

$$\begin{aligned} & -\sum_{i=1}^n \frac{\delta_i}{\sigma} + \sum_{i=1}^n \delta_i \left[-\left(\frac{y_i - \mu - \boldsymbol{\alpha}^t \mathbf{Z}}{\sigma^2}\right) + \left(\frac{y_i - \mu - \boldsymbol{\alpha}^t \mathbf{Z}}{\sigma^2}\right) \exp\left(\frac{y_i - \mu - \boldsymbol{\alpha}^t \mathbf{Z}}{\sigma}\right) \right] \\ & + \sum_{i=1}^n (1 - \delta_i) \left[\left(\frac{y_i - \mu - \boldsymbol{\alpha}^t \mathbf{Z}}{\sigma^2}\right) \exp\left(\frac{y_i - \mu - \boldsymbol{\alpha}^t \mathbf{Z}}{\sigma}\right) \right] = 0 \end{aligned} \quad (3.32)$$

Taksiran parameter α_k untuk $k = 1, 2, \dots, m$ diperoleh dengan cara menyelesaikan persamaan

$$\frac{\partial \ln L}{\partial \alpha_k} = 0$$

untuk $k = 1, 2, \dots, m$, yaitu

$$\sum_{i=1}^n \delta_i \left[-\frac{Z_k}{\sigma} + \frac{Z_k}{\sigma} \exp\left(\frac{y_i - \mu - \boldsymbol{\alpha}^t \mathbf{Z}}{\sigma}\right) \right] + \sum_{i=1}^n (1 - \delta_i) \left[\frac{Z_k}{\sigma} \exp\left(\frac{y_i - \mu - \boldsymbol{\alpha}^t \mathbf{Z}}{\sigma}\right) \right] = 0 \quad (3.33)$$

Dari bentuk persamaan (3.31) sampai (3.33) terlihat bahwa estimasi μ , σ , dan α_k tidak dapat diselesaikan secara eksak, sehingga diperlukan penyelesaian secara numerik.

Berikut ini akan dijelaskan metode numerik dengan menggunakan *Newton-Rhapson* untuk menaksir parameter pada model Weibull AFT dengan data tersensor kanan. Misalkan matriks $\hat{\mathbf{a}}$ adalah matriks kolom ukuran $m + 2$ yang isinya taksiran parameter μ , σ , α_1 , \dots , α_m , maka

$$\hat{\mathbf{a}} = \begin{bmatrix} \hat{\mu} \\ \hat{\sigma} \\ \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_m \end{bmatrix}. \quad (3.34)$$

Matriks $\mathbf{G}(\mu, \sigma, \alpha_1, \dots, \alpha_m)$ adalah matriks kolom ukuran $m + 2$ yang isinya turunan pertama fungsi *log-likelihood* terhadap parameter $\mu, \sigma, \alpha_1, \dots, \alpha_m$, maka matriks $\mathbf{G}(\mu, \sigma, \alpha_1, \dots, \alpha_m) = \mathbf{0}$ bersesuaian dengan persamaan (3.31) sampai (3.33). Jika $\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m$ adalah taksiran parameter $\mu, \sigma, \alpha_1, \dots, \alpha_m$ yang memaksimumkan fungsi *likelihood*, maka matriks $\mathbf{G}(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m) = \mathbf{0}$ adalah

$$\mathbf{G}(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m) = \begin{bmatrix} G_1(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m) \\ G_2(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m) \\ G_3(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m) \\ \vdots \\ G_{m+2}(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m) \end{bmatrix} = \begin{bmatrix} \frac{\partial \ln L(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m)}{\partial \hat{\mu}} \\ \frac{\partial \ln L(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m)}{\partial \hat{\sigma}} \\ \frac{\partial \ln L(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m)}{\partial \hat{\alpha}_1} \\ \vdots \\ \frac{\partial \ln L(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m)}{\partial \hat{\alpha}_m} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (3.35)$$

Definisikan matriks $\mathbf{J}(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m)$, yaitu matriks *Jacobian* berukuran $([m + 2] \times [m + 2])$ yang isinya adalah turunan fungsi pada matriks $\mathbf{G}(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m)$ terhadap taksiran $\mu, \sigma, \alpha_1, \dots, \alpha_m$.

$$\mathbf{J}(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m) = \begin{bmatrix} \frac{\partial G_1}{\partial \hat{\mu}} & \dots & \frac{\partial G_1}{\partial \hat{\alpha}_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial G_{m+2}}{\partial \hat{\mu}} & \dots & \frac{\partial G_{m+2}}{\partial \hat{\alpha}_m} \end{bmatrix} \quad (3.36)$$

Prosedur *Newton-Raphson* untuk mencari taksiran $\hat{\mathbf{a}}$ sehingga $\mathbf{G}(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m) = \mathbf{0}$ adalah sebagai berikut:

1. Input:
 - Batas toleransi (*tol*)
 - Taksiran awal ($\hat{\mathbf{a}}^{(0)}$)
2. Pada setiap iterasi ke- r
 - Hitung taksiran baru untuk $\hat{\mathbf{a}}$

$$\hat{\mathbf{a}}^{(r)} = \hat{\mathbf{a}}^{(r-1)} - \{\mathbf{J}^{(r-1)}(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m)\}^{-1} \mathbf{G}^{(r-1)}(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m)$$

- Hitung error taksiran dibandingkan dengan taksiran sebelumnya

$$error^{(r)} = \hat{\mathbf{a}}^{(r)} - \hat{\mathbf{a}}^{(r-1)}$$

- Periksa kondisi berikut

$$\|error^{(r)}\|_{\infty} > tol, \text{ lanjut ke iterasi berikutnya.}$$

$$\|error^{(r)}\|_{\infty} \leq tol, \text{ iterasi selesai.}$$

3. Output:

- Taksiran untuk $\hat{\mathbf{a}}$ adalah $\hat{\mathbf{a}}^{(r)}$
- Error dari taksiran yang diperoleh adalah $\|error^{(r)}\|_{\infty}$
- Iterasi yang dilakukan adalah sebanyak r

$\{J^{(r-1)}(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m)\}^{-1}$ adalah invers dari matriks *Jacobian* dengan menggunakan nilai $\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m$ yang diperoleh pada iterasi ke $r-1$.

$G^{(r-1)}(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m)$ adalah matriks seperti pada persamaan (3.38) dengan nilai $\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \dots, \hat{\alpha}_m$ yang diperoleh pada iterasi ke $r-1$.

Dengan metode *Newton-Raphson* di atas maka taksiran untuk $\hat{\mathbf{a}}$ pada persamaan (3.34) adalah $\hat{\mathbf{a}}^{(r)}$.

Misalkan estimasi yang diperoleh adalah $\hat{\mu}, \hat{\sigma}$, dan $\hat{\alpha}_k$ untuk $k=1, 2, \dots, m$. Berdasarkan hubungan pada persamaan (3.17), maka estimasi *maximum likelihood* untuk λ, p , dan β_k diberikan oleh

$$\begin{aligned}\hat{\lambda} &= \exp(-\hat{\mu}/\hat{\sigma}) \\ \hat{p} &= 1/\hat{\sigma} \\ \hat{\beta}_k &= -\hat{\alpha}_k/\hat{\sigma}\end{aligned}\tag{3.37}$$

3.3 Pengecekan Model Weibull AFT

Plot dari residual dapat digunakan untuk menguji kecocokan model. Prosedur yang berdasarkan residual dari model AFT ini relevan dengan yang digunakan pada model Cox PH. Salah satu plot yang dapat digunakan adalah dengan membandingkan distribusi dari Cox-Snell residual dengan distribusi unit eksponensial. Berikut ini akan dijelaskan terlebih dahulu mengenai Cox-Snell residual.

Cox-Snell residual didefinisikan sebagai

$$r_c = \hat{H}(x|\mathbf{Z}) = -\ln[\hat{S}(x|\mathbf{Z})], \quad (3.38)$$

di mana

- x adalah waktu *survival* dari individu
- \mathbf{Z} adalah vektor kovariat
- $\hat{H}(x)$ dan $\hat{S}(x)$ adalah estimasi dari fungsi kumulatif *hazard* dan fungsi *survival* dari model.

Misalkan waktu *survival* X memiliki distribusi *survival* yang kontinu, $S(x)$, dengan fungsi kumulatif *hazard*

$$H(x) = -\ln S(x)$$

atau

$$S_X(x) = \exp[-H(x)].$$

Misalkan $Q = H(X)$ adalah fungsi kumulatif *hazard* dari X . Maka fungsi *survival* untuk Q adalah

$$\begin{aligned} S_Q(q) &= \Pr(Q > q) = \Pr(H(x) > q) \\ &= \Pr(X > H_X^{-1}(q)) = S_X(H_X^{-1}(q)) \\ &= \exp[-H_X(H_X^{-1}(q))] = \exp(-q) \end{aligned}$$

Fungsi *survival* tersebut adalah bentuk *survival* dari distribusi eksponensial, yaitu

$$S(t) = \exp(-\lambda t)$$

dengan nilai parameter $\lambda = 1$. Oleh karena itu, tanpa memperhatikan distribusi dari X , variabel baru $Q = H(X)$ memiliki distribusi unit eksponensial.

Jika model yang digunakan cocok, maka nilai dari $\hat{S}(x)$ akan memiliki sifat yang sama seperti $S(x)$. Dengan demikian, $r_c = -\ln \hat{S}(x)$ juga memiliki distribusi unit eksponensial dan fungsi distribusinya adalah $f_R(r) = \exp(-r)$. Misalkan $S_R(r)$ menyatakan fungsi *survival* dari Cox-Snell residual r_c , maka

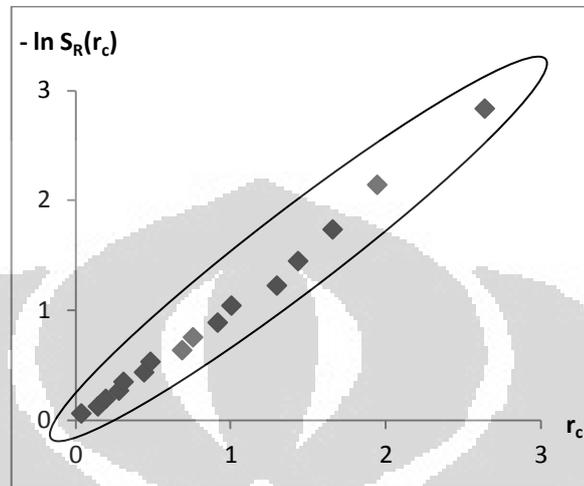
$$S_R(r) = \int_r^\infty f_R(t) dt = \int_r^\infty \exp(-t) dt = \exp(-r) \quad (3.39)$$

dan

$$H_R(r) = -\ln S_R(r) = -\ln[\exp(-r)] = r \quad (3.40)$$

Oleh karena itu, plot $H_R(r)$ terhadap r_c atau $-\ln S_R(r_c)$ terhadap r_c dapat digunakan untuk memeriksa kecocokan model untuk sembarang distribusi dari X .

Plot tersebut akan memberikan garis lurus dengan nilai *slope* sebesar satu dan *intercept* bernilai nol jika model yang digunakan tepat.



Gambar 3.3. Ilustrasi grafik hubungan $-\ln S_R(r_c)$ dengan r_c

Pada model Weibull AFT, berdasarkan persamaan (3.13) estimasi fungsi *survival* diberikan oleh

$$\hat{S}_Y(y) = S_W\left(\frac{y - \hat{\mu} - \hat{\alpha}^t \mathbf{Z}}{\hat{\sigma}}\right)$$

di mana $\hat{\mu}$, $\hat{\alpha}$, dan $\hat{\sigma}$ berturut-turut adalah estimasi *maximum likelihood* dari μ , α , dan σ , serta $S_W(w)$ adalah fungsi *survival* dari W pada model AFT.

Misalkan

$$\frac{y - \hat{\mu} - \hat{\alpha}^t \mathbf{Z}}{\hat{\sigma}} = r_s \quad (3.41)$$

Pada model Weibull AFT, fungsi *survival* untuk W dari persamaan (3.10) adalah

$$S_W(w) = \exp(-e^w).$$

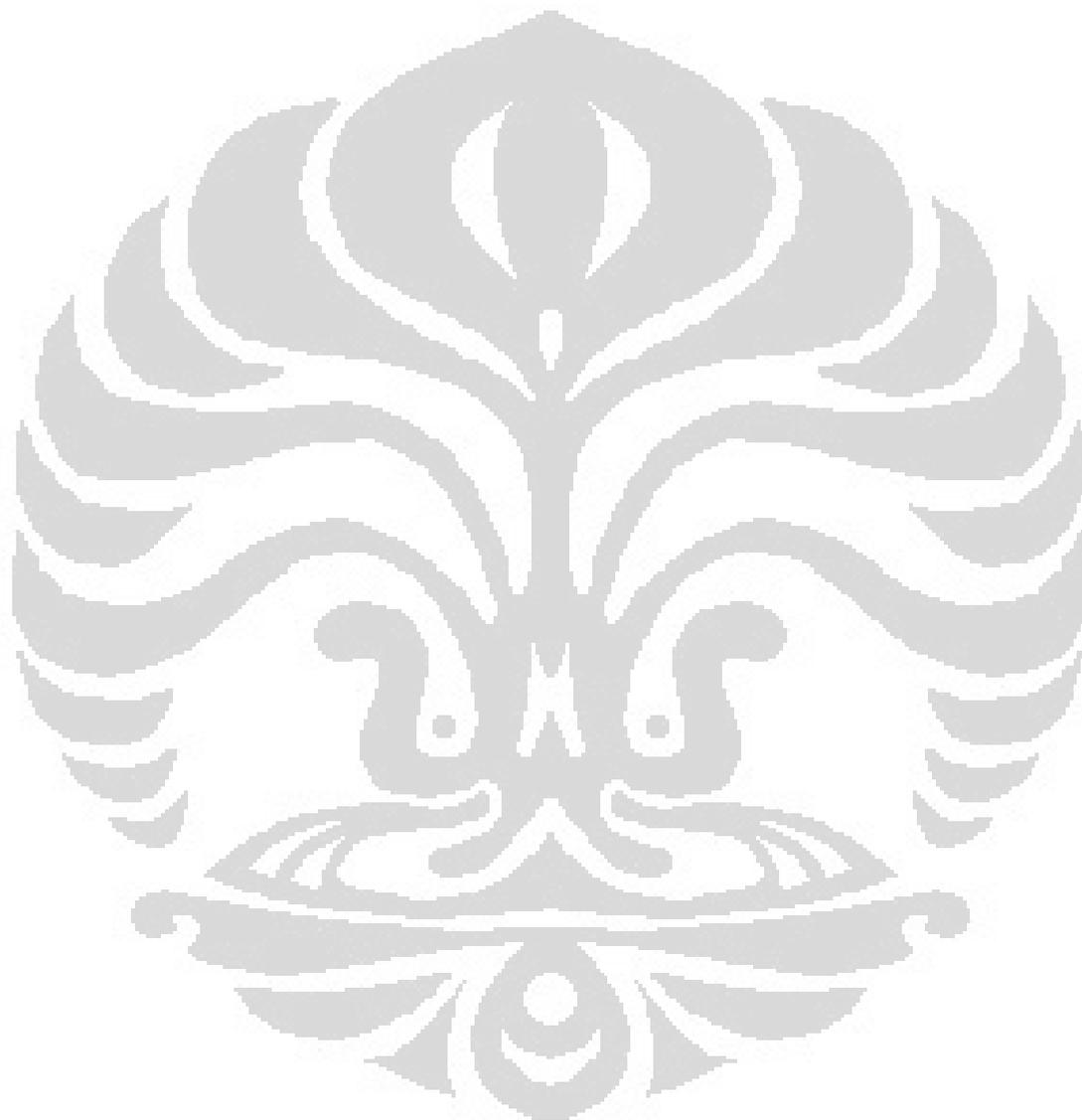
Dengan menggunakan fungsi *survival* dari W , maka Cox-Snell residualnya adalah

$$\begin{aligned} r_c &= -\ln[\hat{S}(x)] = -\ln[S_W(r_s)] \\ &= -\ln[\exp(-e^{r_s})] \\ r_c &= \exp(r_s) \end{aligned} \quad (3.42)$$

Jika model Weibull AFT yang digunakan cocok, maka plot dari $-\ln S(r_c)$ terhadap r_c adalah garis lurus dengan *slope* sebesar satu dan melalui titik asal

(seperti yang terlihat pada gambar 3.3 sebelumnya), dengan nilai r_c berdasarkan persamaan (3.41) dan (3.42) yaitu

$$r_c = \exp\left(\frac{y - \hat{\mu} - \hat{\alpha}^t \mathbf{Z}}{\hat{\sigma}}\right). \quad (3.44)$$





Gambar 3.4. Bagan alur pengecekan Cox-Snell residual

BAB 4

APLIKASI PADA DATA

Pada bab ini akan dibahas penerapan model parametrik Weibull AFT pada data. Kemudian akan dijelaskan interpretasi dari taksiran yang telah diperoleh. Selain itu, akan dibentuk juga model Cox PH sebagai perbandingannya. Walaupun model parametrik Weibull AFT dapat diterapkan untuk data dengan skema penyensoran kanan, kiri, ataupun interval, pada contoh penerapan ini hanya digunakan data tersensor kanan untuk menyesuaikan penggunaannya dengan model Cox PH.

Dalam metode parametrik diasumsikan bahwa distribusi dari waktu *survival* diketahui. Akan tetapi, pada saat diberikan suatu data *survival*, tidak selalu dapat ditentukan distribusi yang tepat. Oleh karena itu, pada bab ini juga akan dilakukan simulasi dengan men-*generate* data dari distribusi Weibull (asumsi terpenuhi) dan dari distribusi selain Weibull (untuk contoh asumsi tidak terpenuhi).

Kemudian akan dilihat nilai Cox-Snell residual, baik untuk model parametrik Weibull AFT maupun model Cox PH. Hal ini dilakukan untuk membandingkan model mana yang lebih baik. Suatu model dikatakan baik jika plot dari $-\ln S_R(r_c)$ terhadap r_c mendekati garis lurus dengan *slope* sebesar satu dan melalui titik asal, seperti yang dijelaskan pada subbab 3.3.

4.1 Data Aplikasi

Pada subbab ini akan dibahas penerapan model Weibull AFT serta perbandingannya dengan model Cox PH dengan menggunakan data yang diambil dari Freireich et al. 1963. Data diperoleh dengan mengamati waktu meninggal (dalam minggu) pasien yang menderita leukemia.

- Awal masa studi : pertama kali dilakukan pengobatan
- Akhir masa studi : meninggal atau akhir masa studi
- *Event* : meninggal
- Skala waktu : minggu

Pasien leukemia tersebut dibagi menjadi dua grup dengan jenis *treatment* yang berbeda, yaitu dengan *treatment* tertentu dan pemberian *placebo*. Perbedaan jenis *treatment* (TRT) yang diberikan kepada pasien leukemia tersebut menjadi kovariat yang dinyatakan dalam variabel Z , di mana

$$Z = \begin{cases} 1, & \text{treatment} \\ 0, & \text{placebo} \end{cases}$$

Data ini terdiri dari 42 observasi, di mana

- terdapat 21 pasien yang menerima *treatment* tertentu dan 21 pasien lainnya yang menerima *placebo*;
 - terdapat 30 pasien yang mengalami *event* dan 12 pasien yang tersensor kanan.
- Untuk lebih jelasnya, data dapat dilihat pada lampiran 1.

4.1.1 Pengecekan Asumsi Distribusi Weibull

Untuk dapat menggunakan model parametrik Weibull AFT, pertama akan diperiksa terlebih dahulu apakah data *survival* tersebut memenuhi asumsi distribusi Weibull. Salah satu cara yang dapat digunakan adalah dengan menggunakan sifat khusus dari distribusi Weibull yang sudah dijabarkan pada bab 3, yaitu log dari negatif log fungsi *survival* ($\ln[-\ln S(x)]$) linier terhadap log waktu ($\ln x$). Nilai dari $S(x)$ diperoleh melalui estimasi *survival* Kaplan-Meier. Dengan menggunakan suatu perangkat lunak statistika, estimasi dari *survival* $S(x)$ yang diperoleh adalah sebagai berikut:

Tabel 4.1. Hasil estimasi *survival* Kaplan-Meier dari data leukemia

➤ $Z = 1, \text{treatment}$

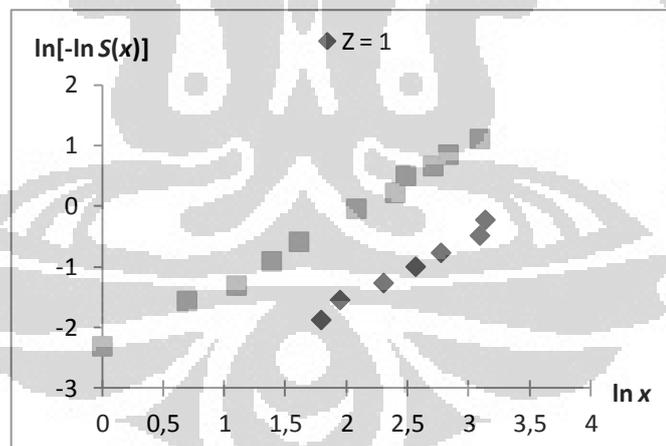
Waktu	Jumlah resiko	Jumlah event	Survival
6	21	3	0.857
7	17	1	0.807
10	15	1	0.753
13	12	1	0.690
16	11	1	0.627
22	7	1	0.538
23	6	1	0.448

➤ $Z = 0, placebo$

Waktu	Jumlah resiko	Jumlah event	Survival
1	21	2	0.9048
2	19	2	0.8095
3	17	1	0.7619
4	16	2	0.6667
5	14	2	0.5714
8	12	4	0.3810
11	8	2	0.2857
12	6	2	0.1905
15	4	1	0.1429
17	3	1	0.0952
22	2	1	0.0476
23	1	1	0.0000

Jumlah resiko adalah jumlah pasien yang beresiko mengalami *event* pada waktu x dan jumlah *event* adalah banyaknya pasien yang meninggal.

Kemudian dari nilai *survival* yang diperoleh untuk setiap kelompok pasien tersebut, dibuat plot antara $\ln[-\ln S(x)]$ dengan $\ln x$, sehingga diperoleh hasil sebagai berikut:



Gambar 4.1. Grafik hubungan $\ln[-\ln S(x)]$ dengan $\ln x$ dari data leukemia

Dari plot di atas terlihat bahwa kedua garis yang terbentuk dari kedua kelompok pasien tersebut adalah garis-garis yang linier, sehingga asumsi data tersebut berdistribusi Weibull terpenuhi dan model parametrik Weibull dapat digunakan. Kedua garis tersebut juga merupakan garis-garis yang sejajar sehingga

asumsi PH terpenuhi dan begitu juga dengan asumsi AFT terpenuhi berdasarkan sifat model Weibull yang sudah dijelaskan sebelumnya pada bab 3.

Berikut akan dibahas penerapan data leukemia tersebut dengan menggunakan model parametrik Weibull AFT, serta perbandingannya dengan model Cox PH.

4.1.2 Model Parametrik Weibull AFT

Berdasarkan persamaan (3.12), model Weibull AFT yang dibentuk untuk data leukemia tersebut adalah

$$Y = \mu + \alpha Z + \sigma W$$

di mana

- $Y = \ln X$ adalah log dari waktu *survival*
- Z adalah kovariat *treatment* (TRT) yang diberikan, dengan

$$Z = \begin{cases} 1, & \text{treatment} \\ 0, & \text{placebo} \end{cases}$$

- W adalah komponen error.
- μ adalah *intercept*
- α adalah koefisien regresi
- σ adalah *scale* parameter

Dengan menggunakan perangkat lunak statistika, estimasi parameter untuk model parametrik Weibull AFT yang diperoleh adalah sebagai berikut:

Tabel 4.2. Hasil penaksiran parameter model Weibull AFT dari data leukemia

	Value	Std. error	z	p
<i>intercept</i> (μ)	2.248	0.166	13.55	8.30e-042
TRT (α)	1.267	0.311	4.08	4.51e-005
log (<i>scale</i>)	-0.312	0.147	-2.12	3.43e-002
<i>scale</i> (σ)	0.732			

Berdasarkan output di atas, maka estimasi model yang didapat adalah

$$\hat{Y} = \ln \hat{X} = 2.248 + 1.267Z.$$

Estimasi waktu *survival* untuk pasien yang diberi *treatment* ($Z = 1$) adalah

$$\hat{Y} = \ln \hat{X} = 2.248 + 1.267 = 3.515$$

$$\hat{X}_t = e^{3.515} = 33.616 \text{ minggu.}$$

Atau dengan kata lain, rata-rata waktu hidup pasien leukemia yang diberikan *treatment* adalah 33,616 minggu sejak pertama kali diberikan pengobatan tersebut.

Estimasi waktu *survival* untuk pasien yang diberikan *placebo* ($Z=0$) adalah

$$\hat{Y} = \ln \hat{X} = 2.248$$

$$\hat{X}_p = e^{2.248} = 9.4688 \text{ minggu.}$$

Atau dengan kata lain, rata-rata waktu hidup pasien leukemia yang diberi *placebo* adalah 9,4688 minggu sejak pertama kali diberikan pengobatan tersebut.

Acceleration factor γ dapat dihitung dengan

$$\hat{\gamma} = \frac{\hat{X}_t}{\hat{X}_p} = \frac{33.616}{9.4688} = 3.55$$

Berdasarkan model parametrik Weibull AFT, waktu *survival* dari pasien leukemia yang diberikan *treatment* tertentu adalah 3,55 kali dari waktu *survival* pasien leukemia yang diberikan *placebo*. Atau dengan kata lain, pasien yang diberikan *treatment*, secara rata-rata, akan mengalami *event* (meninggal) lebih lama dibandingkan pasien yang diberi *placebo*.

Selain waktu *survival* di atas, fungsi *survival* dan fungsi *hazard* dari pasien leukemia tersebut juga dapat dihitung dengan menggunakan model parametrik Weibull AFT. Untuk menghitung fungsi *survival*, digunakan fungsi *survival* model Weibull AFT pada persamaan (3.16), yaitu

$$S_X(x|Z) = \exp \left[- \exp \left(\frac{-\mu - \alpha Z}{\sigma} \right) x^{1/\sigma} \right]$$

Berdasarkan hasil estimasi pada tabel 4.2, estimasi fungsi *survival* untuk pasien yang diberikan *treatment* tertentu adalah

$$\hat{S}_X(x|Z = 1) = \exp[-(0.008)x^{1.366}].$$

Estimasi fungsi *survival* untuk pasien yang diberikan *placebo* adalah

$$\hat{S}_X(x|Z = 0) = \exp[-(0.046)x^{1.366}].$$

Untuk melihat perbandingan nilai dari fungsi *survival* untuk kedua kelompok pasien leukemia tersebut, misalkan waktu pengamatannya diambil pada saat $x = 10$, maka diperoleh nilai *survival* untuk pasien yang diberikan *treatment* adalah 0,8304, sedangkan nilai *survival* untuk pasien yang diberikan *placebo* adalah 0,3435.

Dari hasil estimasi fungsi *survival* kedua kelompok pasien leukemia tersebut, didapat bahwa fungsi *survival* atau probabilitas untuk bertahan hidup dari pasien yang diberikan *treatment* tertentu lebih besar dibandingkan dengan pasien yang diberikan *placebo*.

Untuk mengestimasi fungsi *hazard* dari pasien leukemia tersebut, digunakan bentuk fungsi *hazard* model Weibull AFT pada persamaan (3.18) yaitu

$$h(x|Z) = \lambda p x^{p-1} e^{p\theta Z}.$$

Dengan menggunakan hubungan antar parameter pada persamaan (3.17) dan (3.19), yaitu

$$\lambda = \exp(-\mu/\sigma), p = 1/\sigma, \theta = -\alpha$$

maka fungsi *hazard* model Weibull AFT tersebut dapat ditulis ulang sebagai

$$h(x|Z) = \frac{1}{\sigma} \left[\exp\left(\frac{-\mu - \alpha Z}{\sigma}\right) x^{\frac{1}{\sigma}-1} \right]$$

Berdasarkan hasil estimasi pada tabel 4.3, estimasi fungsi *hazard* untuk pasien yang diberikan *treatment* tertentu adalah

$$\hat{h}(x|Z = 1) = \frac{1}{0.732} \left[\exp\left(\frac{-2.248 - 1.267}{0.732}\right) x^{0.732-1} \right]$$

$$\hat{h}(x|Z = 1) = (0.011)x^{0.366}.$$

Estimasi fungsi *hazard* untuk pasien yang diberikan *placebo* adalah

$$\hat{h}(x|Z = 0) = \frac{1}{0.732} \left[\exp\left(\frac{-2.248}{0.732}\right) x^{0.732-1} \right]$$

$$\hat{h}(x|Z = 0) = (0.063)x^{0.366}.$$

Untuk melihat perbandingan nilai dari fungsi *hazard* untuk kedua kelompok pasien leukemia tersebut, kembali dimisalkan waktu pengamatannya diambil pada saat $x = 10$, maka diperoleh nilai *hazard* untuk pasien yang diberikan *treatment* adalah 0,0255, sedangkan nilai *hazard* untuk pasien yang diberikan *placebo* adalah 0,1463.

Dari hasil estimasi fungsi *hazard* kedua kelompok pasien leukemia tersebut, didapat bahwa fungsi *hazard* atau resiko mengalami *event* dari pasien yang diberikan *treatment* tertentu lebih kecil dibandingkan dengan pasien yang diberikan *placebo*.

Berdasarkan hasil ketiga jenis pengukuran, baik waktu *survival*, fungsi *survival*, maupun fungsi *hazard*, seluruhnya menyatakan bahwa pengobatan dengan cara memberikan suatu *treatment* tertentu pada pasien leukemia lebih baik dibandingkan pengobatan dengan *placebo*.

Berikut ini rangkuman hasil estimasi dari waktu *survival*, fungsi *survival*, dan fungsi *hazard* yang diperoleh dengan menggunakan model parametrik Weibull AFT.

Tabel 4.3. Hasil estimasi berdasarkan model parametrik Weibull AFT dari data leukemia

	Pasien dengan <i>treatment</i>	Pasien dengan <i>placebo</i>
Waktu <i>survival</i>	33.616 minggu	9.4688 minggu
Fungsi <i>survival</i>	$\hat{S}_x(x) = \exp[-(0.008)x^{1.366}]$	$\hat{S}_x(x) = \exp[-(0.046)x^{1.366}]$
Fungsi <i>hazard</i>	$\hat{h}(x) = (0.011)x^{0.366}$	$\hat{h}(x) = (0.063)x^{0.366}$

4.1.3 Model Cox PH

Berdasarkan persamaan (2.17), model Cox PH yang dibentuk untuk data leukemia tersebut adalah

$$h(x|Z) = h_0(x) \exp(\beta Z)$$

di mana

- x adalah waktu *survival* pasien (dalam minggu)
- $h_0(x)$ adalah fungsi *baseline hazard*
- Z adalah *treatment* (TRT) yang diberikan, dengan

$$Z = \begin{cases} 1, & \text{treatment} \\ 0, & \text{placebo} \end{cases}$$

Dengan menggunakan suatu perangkat lunak statistika, akan didapatkan taksiran parameter pada model Cox PH seperti pada tabel di bawah ini.

Tabel 4.4. Hasil penaksiran parameter model Cox PH dari data leukemia

	coef	exp(coef)	se(coef)	z	p
TRT	- 1.57	0.208	0.412	- 3.81	0.00014

Estimasi model Cox PH yang diperoleh adalah

$$h(x|Z) = h_0(x) \exp(-1.57Z).$$

Berdasarkan output di atas, dapat dihitung *hazard ratio* dari kelompok pasien yang diberikan *treatment* tertentu dan pasien yang diberikan *placebo*.

Hazard dari pasien yang mendapatkan *treatment* tertentu adalah

$$h_t = h(x|Z = 1) = h_0(x)e^{-1.57},$$

sedangkan *hazard* dari pasien yang diberikan *placebo* adalah

$$h_p = h(x|Z = 0) = h_0(x).$$

Hazard ratio yang diperoleh adalah

$$\frac{h_t}{h_p} = \frac{h_0(x)e^{-1.57}}{h_0(x)} = e^{-1.57} = 0.208$$

$$h_t = 0.208 h_p$$

Berdasarkan model Cox PH dapat disimpulkan bahwa *hazard* atau resiko kematian pasien yang diberikan *treatment* tertentu adalah 0.208 kali *hazard* pasien yang diberikan *placebo*. Atau dengan kata lain, pasien leukemia yang diberikan *treatment* lebih *survive* dibandingkan pasien yang diberi *placebo*. Kesimpulan yang diambil dari model Cox PH ini juga sama seperti yang diperoleh pada model parametrik Weibull AFT.

4.1.4 Perbandingan Model Cox PH dan Model Weibull AFT

Setelah diperoleh hasil estimasi dengan menggunakan model parametrik Weibull AFT dan model Cox PH, berikut ini akan dijelaskan perbandingan kedua model tersebut dengan menggunakan Cox-Snell residual. Suatu model dikatakan baik jika plot dari $-\ln S_R(r_c)$ terhadap r_c memiliki *slope* sebesar satu dan melalui titik asal, seperti yang dijelaskan pada subbab 3.3.

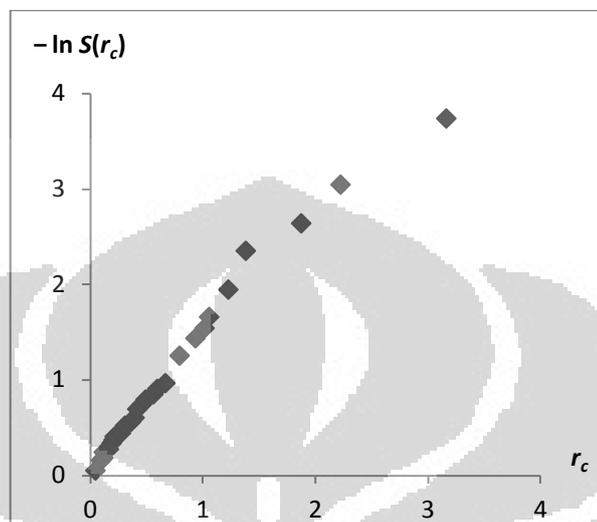
➤ Model parametrik Weibull AFT

Nilai Cox-Snell residual (r_c) pada model AFT dihitung dengan menggunakan persamaan (3.40). Kemudian nilai *survival* dari Cox-Snell residual tersebut dihitung dengan menggunakan estimasi Kaplan-Meier, sehingga diperoleh hasil sebagai berikut:

Tabel 4.5. Estimasi Cox-Snell residual dari model Weibull AFT pada data leukemia

r_c	$S_R(r_c)$	$-\ln S(r_c)$
0.0464	0.9524	0.0488
0.095	0.8571	0.1542
0.1172	0.8333	0.1824
0.1195	0.7857	0.2412
0.1653	0.7619	0.2719
0.1908	0.7143	0.3365
0.208	0.6905	0.3703
0.2174	0.6667	0.4054
0.2731	0.6429	0.4418
0.3081	0.5952	0.5189
0.3627	0.5714	0.5597
0.394	0.5476	0.6022
0.418	0.5	0.6931
0.4587	0.4762	0.7419
0.4919	0.4524	0.7932
0.5604	0.4286	0.8472
0.5954	0.4048	0.9044
0.6673	0.381	0.9650
0.7943	0.2857	1.2528
0.9349	0.2381	1.4351
1.0156	0.2143	1.5404
1.0567	0.1905	1.6581
1.2272	0.1429	1.9456
1.3822	0.0952	2.3518
1.8748	0.0714	2.6395
2.2244	0.0476	3.0449
3.163	0.0238	3.7382

Berdasarkan data pada tabel tersebut, maka plot dari $-\ln S_R(r_c)$ terhadap r_c yang diperoleh adalah sebagai berikut:



Gambar 4.2. Cox-Snell residual dari model Weibull AFT dengan data leukemia

Berdasarkan gambar di atas, terlihat bahwa plot dari $-\ln S_R(r_c)$ terhadap r_c memiliki nilai *slope* mendekati satu dan melalui titik asal, sehingga penggunaan model Weibull AFT untuk data pasien leukemia tersebut cukup tepat.

➤ Model Cox PH

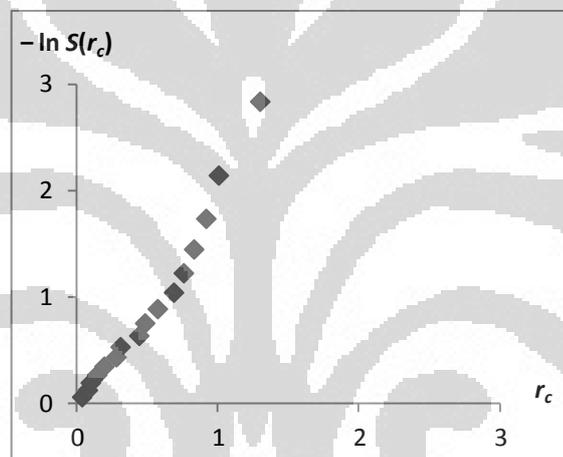
Nilai Cox-Snell residual (r_c) pada model PH, dihitung dengan menggunakan suatu perangkat lunak statistika. Kemudian nilai *survival* dari Cox-Snell residual tersebut dihitung dengan menggunakan estimasi Kaplan-Meier, sehingga diperoleh hasil sebagai berikut:

Tabel 4.6. Estimasi Cox-Snell residual dari model Cox PH pada data leukemia

r_c	$S_R(r_c)$	$-\ln S(r_c)$
0.0367	0.9412	0.0606
0.0769	0.8824	0.1251
0.0976	0.8235	0.1942
0.1439	0.7647	0.2683
0.1948	0.7059	0.3483
0.2797	0.6471	0.4353
0.3092	0.5882	0.5307

0.4416	0.5294	0.6360
0.4829	0.4706	0.7537
0.5727	0.4118	0.8872
0.6872	0.3529	1.0416
0.7572	0.2941	1.2238
0.8301	0.2353	1.4469
0.9163	0.1765	1.7344
1.0051	0.1176	2.1405
1.2983	0.0588	2.8336

Berdasarkan data pada tabel di atas, maka plot dari $-\ln S_R(r_c)$ terhadap r_c yang diperoleh adalah sebagai berikut:



Gambar 4.3. Cox-Snell residual dari model Cox PH dengan data leukemia

Berdasarkan gambar di atas, terlihat bahwa plot dari $-\ln S_R(r_c)$ terhadap r_c memiliki nilai *slope* mendekati dua dan melalui titik asal, sehingga penggunaan model Cox PH untuk data pasien leukemia tersebut kurang tepat dibandingkan dengan model parametrik Weibull AFT.

4.2 Data Simulasi

Simulasi dengan data yang memenuhi asumsi distribusi Weibull dilakukan dengan cara men-*generate* suatu data yang berasal dari distribusi Weibull dengan parameter $\lambda = 6$ dan $p = 2$. Selain itu, simulasi dengan data yang tidak memenuhi asumsi distribusi Weibull dilakukan dengan cara men-*generate* suatu data yang

berasal dari suatu distribusi non-Weibull, yaitu lognormal, dengan parameter $\mu = 2,5$ dan $\sigma = 1$.

Misalkan X menyatakan waktu *survival* suatu invidu dalam satuan waktu minggu. Misalkan terdapat suatu kovariat Z yang menyatakan suatu *treatment* (TRT) yang diberikan kepada individu tersebut, dengan

$$Z = \begin{cases} 1, & \text{treatment A} \\ 0, & \text{treatment B} \end{cases}$$

Data yang di-*generate* ada sebanyak 100 data, di mana

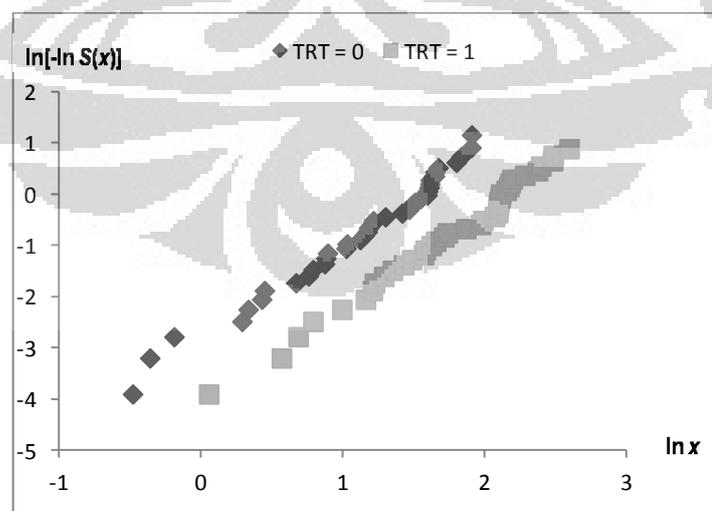
- terdapat dua grup data, yaitu 50 data dengan *treatment* A dan 50 data lainnya dengan *treatment* B;
- masing-masing grup data tersebut mengalami sensor kanan sebanyak 20%.

Untuk lebih jelasnya, data dapat dilihat pada lampiran 8 dan 9.

4.2.1 Simulasi dengan Data Memenuhi Asumsi Distribusi Weibull

Pada subbab ini akan ditunjukkan perbandingan hasil yang diperoleh dari model parametrik Weibull AFT dan model Cox PH jika data yang digunakan memenuhi asumsi distribusi Weibull.

Pertama akan ditunjukkan terlebih dahulu bahwa data yang telah di-*generate* dapat digunakan untuk model Weibull AFT. Cara yang digunakan sama seperti subbab sebelumnya, yaitu dengan cara memplot nilai dari $\ln[-\ln S(x)]$ terhadap $\ln x$. Hasil yang diperoleh adalah sebagai berikut:



Gambar 4.4. Grafik hubungan $\ln[-\ln S(x)]$ dengan $\ln x$ dari data Weibull

Dari gambar 4.4 terlihat bahwa garis-garis yang dihasilkan adalah garis-garis yang linier dan sejajar, sehingga asumsi distribusi Weibull, PH, dan AFT terpenuhi.

Berikut ini akan dijelaskan penggunaan model parametrik Weibull AFT, Cox PH, serta perbandingan kedua model tersebut pada data Weibull yang telah di-generate.

4.2.1.1 Model Parametrik Weibull AFT dengan Data Weibull

Berdasarkan persamaan (3.12), model Weibull AFT yang dibentuk untuk data Weibull tersebut adalah

$$Y = \mu + \alpha Z + \sigma W$$

di mana

- $Y = \ln X$ adalah log dari waktu *survival*
- Z adalah kovariat *treatment* (TRT) yang diberikan, dengan

$$Z = \begin{cases} 1, & \text{treatment A} \\ 0, & \text{treatment B} \end{cases}$$

- W adalah komponen error.
- μ adalah *intercept*
- α adalah koefisien regresi
- σ adalah *scale* parameter

Dengan menggunakan perangkat lunak statistika, estimasi parameter untuk model parametrik Weibull AFT yang diperoleh adalah sebagai berikut:

Tabel 4.7. Hasil penaksiran parameter model Weibull AFT dari data Weibull

	Value	Std. Error	z	p
<i>intercept</i> (μ)	1.512	0.0773	19.56	3.62e-085
TRT (α)	0.686	0.1090	6.30	3.05e-010
log (<i>scale</i>)	-0.719	0.0868	-8.29	1.15e-016
<i>scale</i> (σ)	0.487			

Berdasarkan output di atas, maka estimasi model yang didapat adalah

$$\hat{Y} = \ln \hat{X} = 1.512 + 0.686Z.$$

Berikut ini adalah hasil estimasi dari waktu *survival*, fungsi *survival*, dan fungsi *hazard* yang diperoleh dengan menggunakan model parametrik Weibull AFT.

Tabel 4.8. Hasil estimasi berdasarkan model parametrik Weibull AFT dari data Weibull

	<i>Treatment A</i>	<i>Treatment B</i>
Waktu <i>survival</i>	9.007 minggu	4.536 minggu
Fungsi <i>survival</i>	$\hat{S}_x(x) = \exp[-(0.011)x^{2.053}]$	$\hat{S}_x(x) = \exp[-(0.045)x^{2.053}]$
Fungsi <i>hazard</i>	$\hat{h}(x) = (0.023)x^{1.053}$	$\hat{h}(x) = (0.094)x^{1.053}$

Berdasarkan estimasi waktu *survival* yang diperoleh, *acceleration factor* $\hat{\gamma}$ dapat dihitung dengan

$$\hat{\gamma} = \frac{\hat{X}_A}{\hat{X}_B} = \frac{9.007}{4.536} = 1.9858 \approx 2$$

Berdasarkan model parametrik Weibull AFT, waktu *survival* dari individu yang diberi *treatment A* adalah hampir 2 kali dari waktu *survival* individu yang diberi *treatment B*. Atau dengan kata lain, individu yang diberikan *treatment A*, secara rata-rata, akan mengalami *event* lebih lama dibandingkan individu yang diberi *treatment B*.

Hazard ratio yang diperoleh dari kedua kelompok individu tersebut adalah

$$\frac{h_A}{h_B} = \frac{(0.023)x^{1.053}}{(0.094)x^{1.053}} = 0.245$$

$$h_A = 0.245 h_B.$$

Dari *hazard ratio* tersebut diperoleh bahwa *hazard* atau resiko mengalami *event* untuk individu yang diberikan *treatment A* adalah 0.245 kali *hazard* individu yang diberikan *treatment B*. Atau dengan kata lain, individu yang diberi *treatment B* lebih beresiko mengalami *event* dibandingkan individu yang diberi *treatment A*.

Berdasarkan hasil ketiga jenis pengukuran, baik waktu *survival*, fungsi *survival*, maupun fungsi *hazard*, seluruhnya menyatakan bahwa pemberian *treatment A* lebih baik dibandingkan *treatment B*.

4.2.1.2 Model Cox PH dengan Data Weibull

Berdasarkan persamaan (2.17), model Cox PH yang dibentuk adalah

$$h(x|Z) = h_0(x) \exp(\beta Z)$$

di mana

- x adalah waktu *survival*
- $h_0(x)$ adalah fungsi *baseline hazard*
- Z adalah *treatment* (TRT) yang diberikan, dengan

$$Z = \begin{cases} 1, & \text{treatment A} \\ 0, & \text{treatment B} \end{cases}$$

Dengan menggunakan suatu perangkat lunak statistika, akan didapatkan taksiran parameter pada model Cox PH seperti pada tabel di bawah ini.

Tabel 4.9. Hasil penaksiran parameter model Cox PH dari data Weibull

	coef	exp(coef)	se(coef)	z	p
TRT	- 1.45	0.235	0.281	- 5.51	2.6e-007

Estimasi model Cox PH yang diperoleh adalah

$$h(x|Z) = h_0(x) \exp(-1.45Z).$$

Berdasarkan output di atas, dapat dihitung *hazard ratio* dari kelompok individu yang diberikan *treatment* A dan individu yang diberikan *treatment* B, yaitu

$$\frac{h_A}{h_B} = \frac{h_0(x)e^{-1.45}}{h_0(x)} = e^{-1.45} = 0.235$$

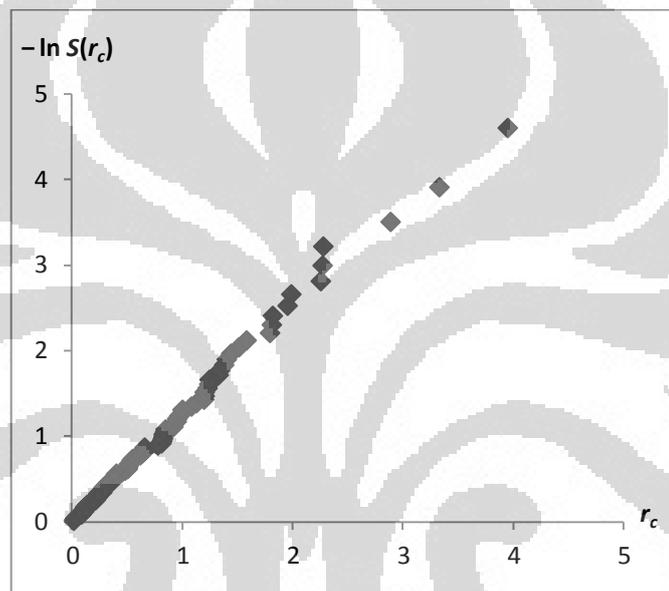
$$h_A = 0.235 h_B$$

Berdasarkan model Cox PH dapat disimpulkan bahwa *hazard* atau resiko mengalami *event* untuk individu yang diberikan *treatment* A adalah 0.235 kali *hazard* individu yang diberikan *treatment* B. Atau dengan kata lain, individu yang diberikan *treatment* A lebih *survive* dibandingkan individu yang diberikan *treatment* B. Kesimpulan yang diambil dari model Cox PH ini juga sama seperti yang diperoleh pada model parametrik Weibull AFT dengan nilai *hazard ratio* yang tidak jauh berbeda.

4.2.1.3 Perbandingan Model Parametrik Weibull AFT dengan Model Cox PH pada Data Berdistribusi Weibull

Untuk membandingkan model parametrik Weibull AFT dengan model Cox PH yang telah dibentuk digunakan metode Cox-Snell residual seperti yang dilakukan pada subbab sebelumnya.

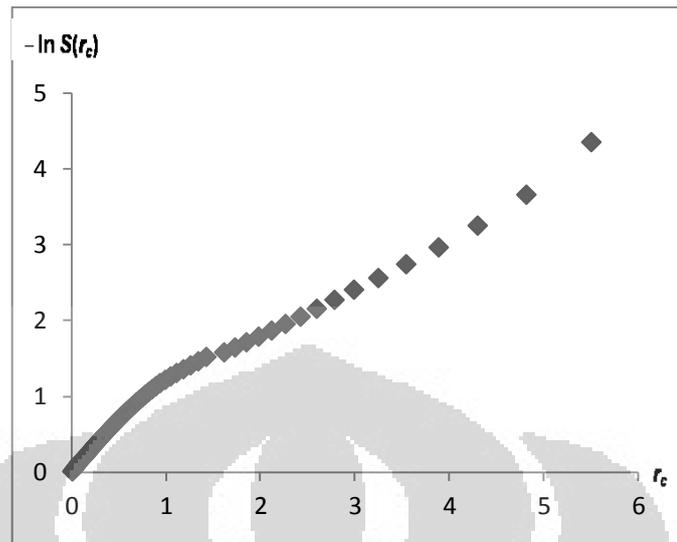
Dengan cara yang sama seperti pada subbab 4.1.4, untuk model parametrik Weibull AFT, plot dari $-\ln S(r_c)$ terhadap r_c yang dihasilkan adalah



Gambar 4.5. Cox-Snell residual dari model parametrik Weibull AFT dengan data Weibull

Berdasarkan gambar di atas, terlihat bahwa plot dari $-\ln S_R(r_c)$ terhadap r_c memiliki nilai *slope* mendekati satu dan melalui titik asal, sehingga penggunaan model Weibull AFT untuk data Weibull tersebut cukup tepat.

Plot $-\ln S(r_c)$ terhadap r_c untuk model Cox PH adalah



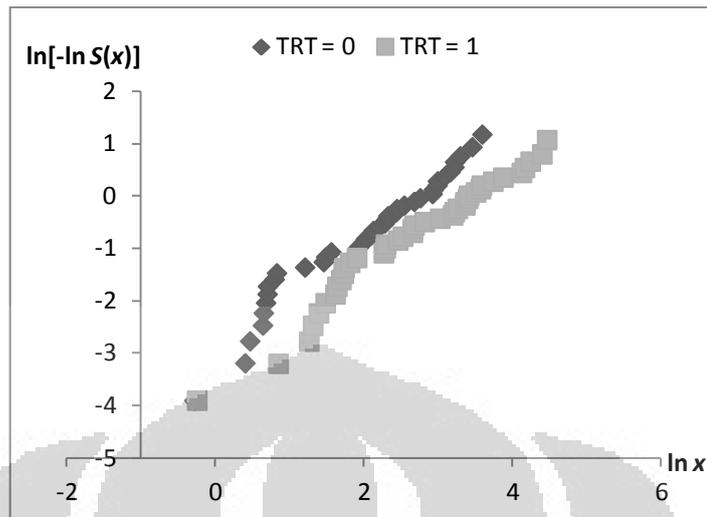
Gambar 4.6. Cox-Snell residual dari model Cox PH dengan data Weibull

Berdasarkan gambar di atas, terlihat bahwa plot dari $-\ln S_R(r_c)$ terhadap r_c melalui titik asal, tetapi nilai *slope* kurang mendekati satu, sehingga penggunaan model Weibull AFT lebih tepat dibandingkan dengan model Cox PH pada data yang di-generate dari distribusi Weibull.

4.2.2 Simulasi dengan Data yang Tidak Berdistribusi Weibull

Pada subbab ini akan ditunjukkan perbandingan hasil yang diperoleh dari model parametrik Weibull AFT dan model Cox PH jika data yang digunakan tidak memenuhi asumsi distribusi Weibull. Data yang di-generate berasal dari distribusi lognormal.

Pertama akan diperiksa terlebih dahulu apakah data yang telah di-generate dapat digunakan untuk model Weibull AFT. Cara yang digunakan sama seperti subbab sebelumnya, yaitu dengan cara memplot nilai dari $\ln[-\ln S(x)]$ terhadap $\ln x$. Hasil yang diperoleh adalah sebagai berikut:



Gambar 4.7. Grafik hubungan $\ln[-\ln S(x)]$ dengan $\ln x$ dari data lognormal

Dari plot di atas terlihat bahwa kedua garis yang terbentuk dari kedua kelompok individu tersebut kurang membentuk garis-garis yang linier dan sejajar. Namun, hal tersebut akan ditoleransi dan data lognormal tersebut tetap akan dianalisis dengan model parametrik Weibull AFT.

Berikut akan dibahas penerapan data leukemia tersebut dengan menggunakan model parametrik Weibull AFT, serta perbandingannya dengan model Cox PH.

4.2.2.1 Model Parametrik Weibull AFT dengan Data Berdistribusi Lognormal

Berdasarkan persamaan (3.12), model Weibull AFT yang dibentuk untuk data lognormal tersebut adalah

$$Y = \mu + \alpha Z + \sigma W$$

di mana

- $Y = \ln X$ adalah log dari waktu *survival*
- Z adalah kovariat *treatment* (TRT) yang diberikan, dengan

$$Z = \begin{cases} 1, & \text{treatment A} \\ 0, & \text{treatment B} \end{cases}$$

- W adalah komponen error.
- μ adalah *intercept*

- α adalah koefisien regresi
- σ adalah *scale* parameter

Dengan menggunakan perangkat lunak statistika, estimasi parameter untuk model parametrik Weibull AFT yang diperoleh adalah sebagai berikut:

Tabel 4.10. Hasil penaksiran parameter model Weibull AFT dari data lognormal

	Value	Std. error	z	p
intercept (μ)	2.7092	0.1521	17.811	5.80e-071
TRT (α)	0.8613	0.2146	4.014	5.97e-005
log (scale)	-0.0418	0.0884	-0.473	6.36e-001
scale (σ)	0.959			

Berdasarkan output pada tabel 4.10, estimasi model yang didapat adalah

$$\hat{Y} = \ln \hat{X} = 2.7092 + 0.8613Z.$$

Berikut ini hasil estimasi dari waktu *survival*, fungsi *survival*, dan fungsi *hazard* yang diperoleh dengan menggunakan model parametrik Weibull AFT.

Tabel 4.11. Hasil estimasi berdasarkan model parametrik Weibull AFT dari data lognormal

	Treatment A	Treatment B
Waktu survival	33.534 minggu	15.017 minggu
Fungsi survival	$\hat{S}_x(x) = \exp[-(0.024)x^{1.043}]$	$\hat{S}_x(x) = \exp[-(0.059)x^{1.043}]$
Fungsi hazard	$\hat{h}(x) = (0.025)x^{0.043}$	$\hat{h}(x) = (0.0615)x^{0.043}$

Acceleration factor γ dapat dihitung dengan

$$\hat{\gamma} = \frac{\hat{X}_A}{\hat{X}_B} = \frac{35.534}{15.017} = 2.366$$

Sehingga berdasarkan model parametrik Weibull AFT, waktu *survival* dari individu yang diberi *treatment* A adalah 2,366 kali dari waktu *survival* individu yang diberi *treatment* B. Atau dengan kata lain, individu yang diberikan *treatment*

A, secara rata-rata, akan mengalami *event* lebih lama dibandingkan individu yang diberi *treatment* B.

Hazard ratio yang diperoleh dari kedua kelompok individu tersebut adalah

$$\frac{h_A}{h_B} = \frac{(0.025)x^{0.043}}{(0.0615)x^{0.043}} = 0.4065$$

$$h_A = 0.4065 h_B.$$

Dari *hazard ratio* tersebut diperoleh bahwa *hazard* atau resiko mengalami *event* untuk individu yang diberikan *treatment* A adalah 0.4065 kali *hazard* individu yang diberikan *treatment* B. Atau dengan kata lain, individu yang diberi *treatment* B lebih beresiko mengalami *event* dibandingkan individu yang diberi *treatment* A.

Berdasarkan hasil ketiga jenis pengukuran, baik waktu *survival*, fungsi *survival*, maupun fungsi *hazard*, seluruhnya menyatakan bahwa pemberian *treatment* A lebih baik dibandingkan *treatment* B.

4.2.2.2 Model Cox PH dengan Data Berdistribusi Lognormal

Berdasarkan persamaan (2.17), model Cox PH yang dibentuk adalah

$$h(x|Z) = h_0(x) \exp(\beta Z)$$

di mana

- x adalah waktu *survival*
- $h_0(x)$ adalah fungsi *baseline hazard*
- Z adalah *treatment* (TRT) yang diberikan, dengan

$$Z = \begin{cases} 1, & \text{treatment A} \\ 0, & \text{treatment B} \end{cases}$$

Dengan menggunakan suatu perangkat lunak statistika, akan didapatkan taksiran parameter pada model Cox PH seperti pada tabel di bawah ini.

Tabel 4.12. Hasil penaksiran parameter model Cox PH dari data lognormal

	coef	exp(coef)	se(coef)	z	p
TRT	- 0.798	0.45	0.244	- 3.28	0.0011

Estimasi model Cox PH yang diperoleh adalah

$$h(x|Z) = h_0(x) \exp(-0.798Z).$$

Berdasarkan estimasi model tersebut, dapat dihitung *hazard ratio* dari kelompok individu yang diberikan *treatment A* dan individu yang diberikan *treatment B*, yaitu

$$\frac{h_A}{h_B} = \frac{h_0(x)e^{-1.45}}{h_0(x)} = e^{-0.798} = 0.45$$

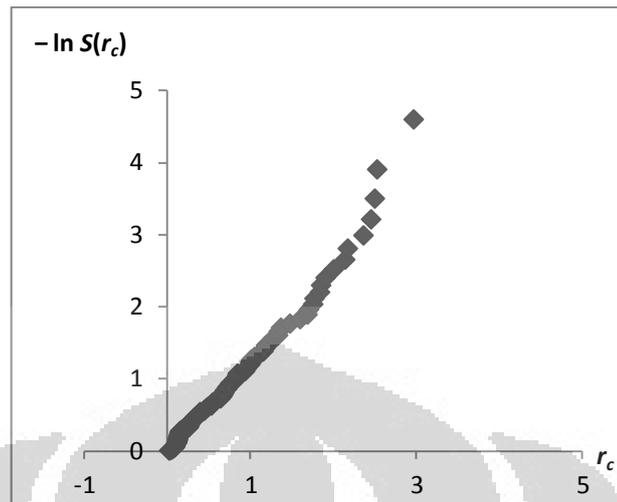
$$h_A = 0.45 h_B$$

Maka berdasarkan model Cox PH dapat disimpulkan bahwa *hazard* atau resiko mengalami *event* untuk individu yang diberikan *treatment A* adalah 0,45 kali *hazard* individu yang diberikan *treatment B*. Atau dengan kata lain, individu yang diberikan *treatment A* lebih *survive* dibandingkan individu yang diberikan *treatment B*. Sehingga kesimpulan yang diambil dari model Cox PH ini juga sama seperti yang diperoleh pada model parametrik Weibull AFT.

4.2.2.3 Perbandingan Model Parametrik Weibull AFT dengan Model Cox PH pada Data Berdistribusi Lognormal

Untuk membandingkan model parametrik Weibull AFT dengan model Cox PH yang telah dibentuk digunakan metode Cox-Snell residual seperti yang dilakukan pada subbab sebelumnya.

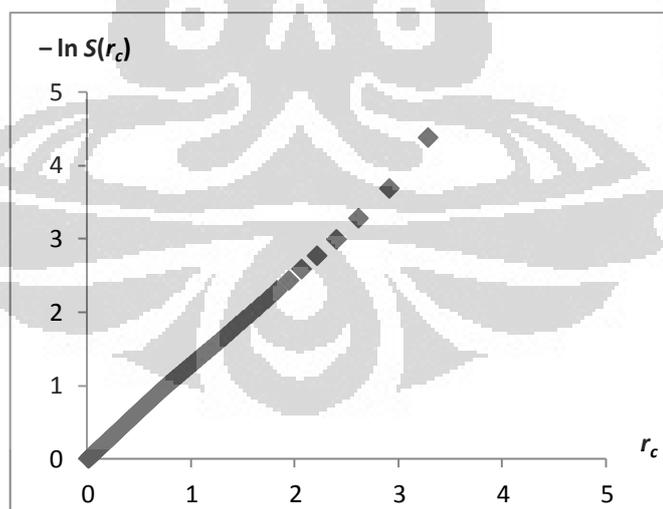
Dengan cara yang sama seperti pada subbab 4.1.4, untuk model parametrik Weibull AFT, plot dari $\ln[-\ln S(r_c)]$ terhadap $\ln r_c$ yang dihasilkan dari data lognormal yang digunakan adalah



Gambar 4.8. Cox-Snell residual dari model parametrik Weibull AFT dengan data lognormal

Berdasarkan gambar di atas, terlihat bahwa plot dari $-\ln S_R(r_c)$ terhadap r_c memiliki nilai *slope* cukup mendekati satu dan melalui titik asal, sehingga penggunaan model Weibull AFT tersebut masih cukup tepat untuk data lognormal tersebut.

Plot $-\ln[-\ln S(r_c)]$ terhadap $\ln r_c$ untuk model Cox PH dengan data lognormal tersebut adalah



Gambar 4.9. Cox-Snell residual dari model parametrik Cox PH dengan data lognormal

Berdasarkan gambar 4.9, terlihat bahwa plot dari $-\ln S_R(r_c)$ terhadap r_c melalui titik asal dan nilai *slope* mendekati satu, sehingga penggunaan model Cox PH cukup tepat untuk model lognormal. Jika dibandingkan dengan plot dari $-\ln S_R(r_c)$ terhadap r_c dari model Weibull AFT, maka terlihat bahwa plot dari model Cox PH menghasilkan garis yang lebih baik, sehingga model Cox PH lebih baik digunakan jika asumsi distribusi pada data tidak terpenuhi pada model parametrik.



BAB 5

KESIMPULAN DAN SARAN

5.1. Kesimpulan

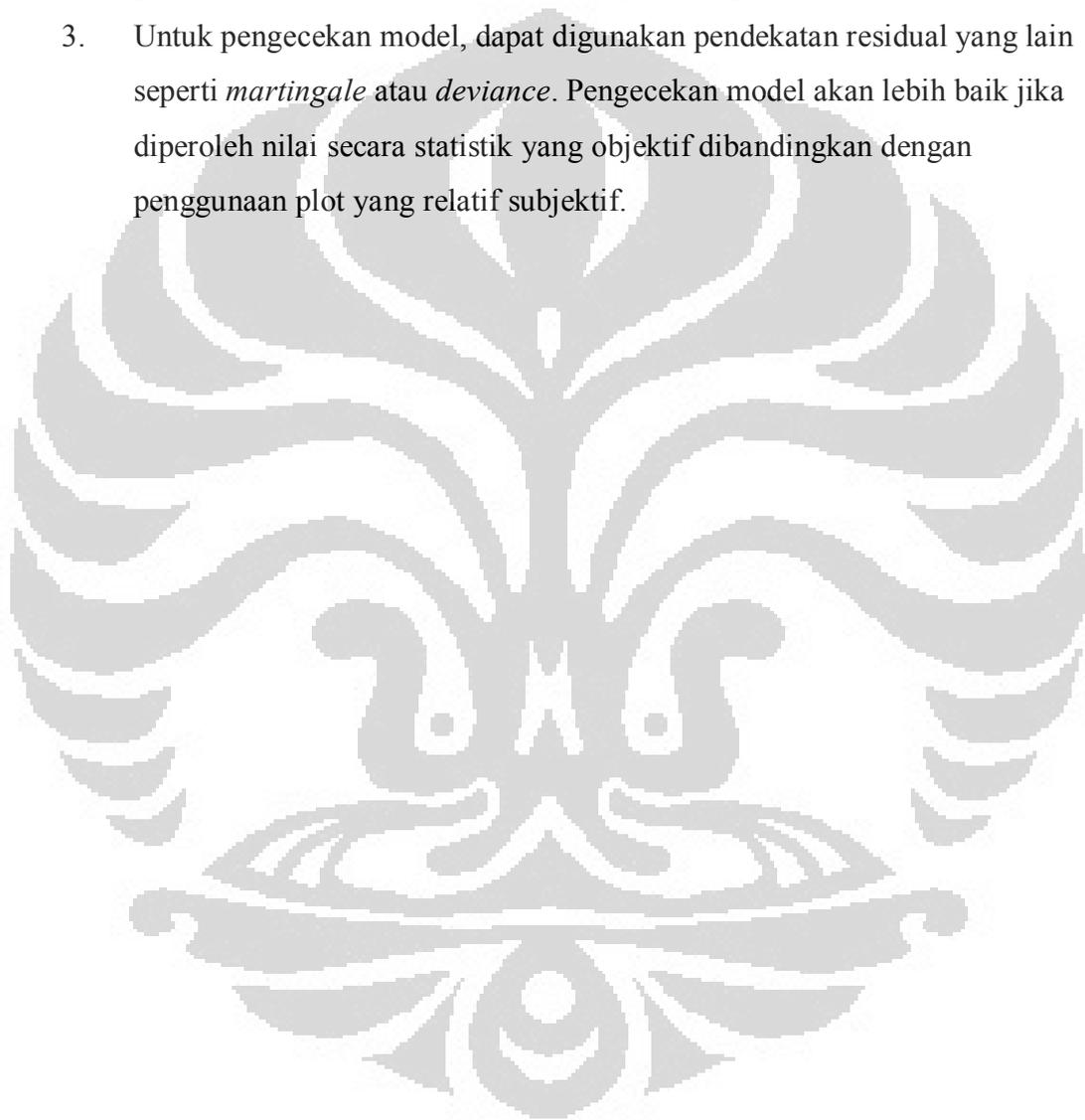
Berdasarkan pembahasan pada bab-bab sebelumnya, kesimpulan yang dapat diperoleh dari penulisan tugas akhir ini adalah:

1. Jika distribusi waktu *survival* diketahui, maka regresi parametrik dapat diterapkan pada data. Untuk waktu *survival* berdistribusi Weibull, dapat dimodelkan secara parametrik dengan dua cara, yaitu Weibull PH dan Weibull AFT. Model Weibull PH diperoleh dengan mensubstitusi fungsi *baseline hazard* pada model Cox PH dengan fungsi *hazard* dari distribusi Weibull. Model Weibull AFT diperoleh dengan melakukan transformasi variabel log waktu sehingga diperoleh model linier log waktu, di mana efek kovariat linier terhadap variabel respon log waktu.
2. Penaksiran parameter pada model parametrik Weibull AFT dilakukan dengan metode *maximum likelihood*. Karena distribusi dari waktu *survival* dapat diketahui, maka setiap objek yang terobservasi maupun yang tersensor dapat berkontribusi langsung terhadap *likelihood*. Parameter pada model Weibull PH memiliki keterkaitan dengan parameter pada model Weibull AFT. Oleh karena itu, jika estimasi parameter pada model Weibull AFT diperoleh, maka estimasi parameter pada model Weibull PH juga dapat diperoleh.
3. Dari contoh penerapan pada data, diperoleh bahwa model parametrik Weibull AFT akan memberikan hasil yang lebih baik jika asumsi distribusi dari waktu *survival* terpenuhi. Jika asumsi distribusi dari waktu *survival* tidak terpenuhi atau distribusi dari waktu *survival* tidak dapat diketahui, maka lebih baik digunakan model Cox PH.

5.2. Saran

Saran untuk pengembangan skripsi ini adalah:

1. Dapat dijelaskan mengenai model parametrik untuk kovariat yang bergantung pada waktu.
2. Dapat digunakan tipe data dengan sensor kiri ataupun interval pada model parametrik, untuk kemudian dibandingkan dengan model *extended* Cox PH.
3. Untuk pengecekan model, dapat digunakan pendekatan residual yang lain seperti *martingale* atau *deviance*. Pengecekan model akan lebih baik jika diperoleh nilai secara statistik yang objektif dibandingkan dengan penggunaan plot yang relatif subjektif.



DAFTAR PUSTAKA

- Burden, R. L., & Faires, J. D. (2001). *Numerical Analysis*. Pacific Grove: Brooks/Cole.
- Galbraits, S. (2004). *Lecture Notes on Medical Statistics*. New South Wales: Department of Statistics, University of New South Wales.
- Hogg, R. V., & Craig, A. T. (1995). *Introduction to Mathematical Statistics*. New Jersey: Prentice-Hall.
- Klein, J. P., & Moeschberger, M. L. (1997). *Survival Analysis - Techniques for Censored and Truncated Data*. New York: Springer-Verlag.
- Kleinbaum, D. G., & Klein, M. (2005). *Survival Analysis - A Self-Learning Text*. New York: Springer.
- Qi, J. (2009). *Comparison of Proportional Hazards and Accelerated Failure Time Models*. Saskatchewan: University of Saskatchewan.
- Zhao, G. (2008). *Nonparametric and Parametric Survival Analysis of Censored Data with Possible Violation of Method Assumptions*. Greensboro: University of North Carolina.

LAMPIRAN

Lampiran 1. Data pasien leukemia

No.	<i>time (x)</i>	TRT (<i>Z</i>)	<i>event (δ)</i>	No.	<i>time (x)</i>	TRT (<i>Z</i>)	<i>event (δ)</i>
1	1	0	1	1	6	1	1
2	1	0	1	2	6	1	1
3	2	0	1	3	6	1	1
4	2	0	1	4	6	1	0
5	3	0	1	5	7	1	1
6	4	0	1	6	9	1	0
7	4	0	1	7	10	1	1
8	5	0	1	8	10	1	0
9	5	0	1	9	11	1	0
10	8	0	1	10	13	1	1
11	8	0	1	11	16	1	1
12	8	0	1	12	17	1	0
13	8	0	1	13	19	1	0
14	11	0	1	14	20	1	0
15	11	0	1	15	22	1	1
16	12	0	1	16	23	1	1
17	12	0	1	17	25	1	0
18	15	0	1	18	32	1	0
19	17	0	1	19	32	1	0
20	22	0	1	20	34	1	0
21	23	0	1	21	35	1	0

dimana

$$Z = \begin{cases} 1, & \text{treatment} \\ 0, & \text{placebo} \end{cases}$$

$$\delta = \begin{cases} 1, & \text{pengamatan eksak} \\ 0, & \text{pengamatan tersensor kanan} \end{cases}$$

Lampiran 2. Estimasi Kaplan-Meier pada data pasien leukemia

```
> KM <- survfit(Surv(time, event) ~ TRT, leukemia)
> summary(KM)
Call: survfit(formula = Surv(time, event) ~ TRT, data = leukemia)
```

```
TRT=0
time n.risk n.event survival std.err lower 95% CI upper 95% CI
  1    21      2  0.9048  0.0641  0.78754  1.000
  2    19      2  0.8095  0.0857  0.65785  0.996
  3    17      1  0.7619  0.0929  0.59988  0.968
  4    16      2  0.6667  0.1029  0.49268  0.902
  5    14      2  0.5714  0.1080  0.39455  0.828
  8    12      4  0.3810  0.1060  0.22085  0.657
 11     8      2  0.2857  0.0986  0.14529  0.562
 12     6      2  0.1905  0.0857  0.07887  0.460
 15     4      1  0.1429  0.0764  0.05011  0.407
 17     3      1  0.0952  0.0641  0.02549  0.356
 22     2      1  0.0476  0.0465  0.00703  0.322
 23     1      1  0.0000      NA      NA      NA
```

```
TRT=1
time n.risk n.event survival std.err lower 95% CI upper 95% CI
  6    21      3  0.857  0.0764  0.720  1.000
  7    17      1  0.807  0.0869  0.653  0.996
 10    15      1  0.753  0.0963  0.586  0.968
 13    12      1  0.690  0.1068  0.510  0.935
 16    11      1  0.627  0.1141  0.439  0.896
 22     7      1  0.538  0.1282  0.337  0.858
 23     6      1  0.448  0.1346  0.249  0.807
```

Lampiran 3. Estimasi model Weibull AFT pada data pasien leukemia

```
> attach(leukemia)
> weib.aft <- survReg(Surv(time, event) ~ TRT, dist="weib")
> summary(weib.aft)
```

Call:

```
survReg(formula = Surv(time, event) ~ TRT, dist = "weib")
```

	Value	Std. Error	z	p
(Intercept)	2.248	0.166	13.55	8.30e-042
TRT	1.267	0.311	4.08	4.51e-005
Log(scale)	-0.312	0.147	-2.12	3.43e-002

Scale= 0.732

Weibull distribution

Loglik(model)= -106.6 Loglik(intercept only)= -116.4

Chisq= 19.65 on 1 degrees of freedom, p= 9.3e-006

Number of Newton-Raphson Iterations: 5

n= 42

Correlation of Coefficients:

	(Intercept)	TRT
TRT	-0.588	
Log(scale)	-0.271	0.344

Lampiran 4. Estimasi model Cox PH pada data pasien leukemia

```
> cox <- coxph(Surv(time, event) ~ TRT, leukemia)
> summary(cox)
Call:
coxph(formula = Surv(time, event) ~ TRT, data = leukemia)

n= 42

      coef exp(coef) se(coef)      z      p
TRT -1.57    0.208    0.412 -3.81 0.00014

      exp(coef) exp(-coef) lower .95 upper .95
TRT    0.208      4.82    0.0925    0.466

Rsquare= 0.322 (max possible= 0.988 )
Likelihood ratio test= 16.4 on 1 df, p=0.0000526
Wald test               = 14.5 on 1 df, p=0.000138
Score (logrank) test = 17.2 on 1 df, p=0.0000328
```

Lampiran 5. Estimasi fungsi *survival* dari Cox-Snell residual untuk model Weibull AFT pada data pasien leukemia

```
> KM2 <- survfit(Surv(rc), data = AFT)
> summary(KM2)
Call: survfit(formula = Surv(rc), data = AFT)
```

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0.0464	42	2	0.9524	0.0329	0.89011	1.000
0.0950	40	4	0.8571	0.0540	0.75759	0.970
0.1172	36	1	0.8333	0.0575	0.72791	0.954
0.1195	35	2	0.7857	0.0633	0.67092	0.920
0.1653	33	1	0.7619	0.0657	0.64339	0.902
0.1908	32	2	0.7143	0.0697	0.58993	0.865
0.2080	30	1	0.6905	0.0713	0.56391	0.845
0.2174	29	1	0.6667	0.0727	0.53831	0.826
0.2731	28	1	0.6429	0.0739	0.51312	0.805
0.3081	27	2	0.5952	0.0757	0.46386	0.764
0.3627	25	1	0.5714	0.0764	0.43976	0.743
0.3940	24	1	0.5476	0.0768	0.41601	0.721
0.4180	23	2	0.5000	0.0772	0.36951	0.677
0.4587	21	1	0.4762	0.0771	0.34676	0.654
0.4919	20	1	0.4524	0.0768	0.32434	0.631
0.5604	19	1	0.4286	0.0764	0.30225	0.608
0.5954	18	1	0.4048	0.0757	0.28049	0.584
0.6673	17	1	0.3810	0.0749	0.25908	0.560
0.7943	16	4	0.2857	0.0697	0.17712	0.461
0.9349	12	2	0.2381	0.0657	0.13861	0.409
1.0156	10	1	0.2143	0.0633	0.12009	0.382
1.0567	9	1	0.1905	0.0606	0.10211	0.355
1.2272	8	2	0.1429	0.0540	0.06810	0.300
1.3822	6	2	0.0952	0.0453	0.03750	0.242
1.8748	4	1	0.0714	0.0397	0.02401	0.213
2.2244	3	1	0.0476	0.0329	0.01231	0.184
3.1636	2	1	0.0238	0.0235	0.00343	0.165
3.3616	1	1	0.0000	NA	NA	NA

Lampiran 6. Estimasi *survival* dari model Cox PH pada data pasien leukemia

```
> summary(survfit(cox))
Call: survfit.coxph(object = cox)
```

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
1	42	2	0.964	0.0254	0.9154	1.000
2	40	2	0.926	0.0367	0.8571	1.000
3	38	1	0.907	0.0414	0.8292	0.992
4	37	2	0.866	0.0497	0.7739	0.969
5	35	2	0.823	0.0572	0.7182	0.943
6	33	3	0.756	0.0663	0.6368	0.898
7	29	1	0.734	0.0687	0.6112	0.882
8	28	4	0.643	0.0771	0.5084	0.813
10	23	1	0.617	0.0791	0.4801	0.794
11	21	2	0.564	0.0826	0.4235	0.752
12	18	2	0.503	0.0860	0.3601	0.703
13	16	1	0.469	0.0873	0.3258	0.676
15	15	1	0.436	0.0878	0.2942	0.647
16	14	1	0.400	0.0880	0.2602	0.616
17	13	1	0.366	0.0873	0.2293	0.584
22	9	2	0.273	0.0868	0.1466	0.509
23	7	2	0.169	0.0784	0.0681	0.419

Lampiran 7. Estimasi fungsi *survival* dari Cox-Snell residual untuk model Cox PH pada data pasien leukemia

```
> KM3 <- survfit(Surv(rc), data = PH)
> summary(KM3)
Call: survfit(formula = Surv(rc), data = PH)
```

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0.0367	17	1	0.9412	0.0571	0.83572	1.000
0.0769	16	1	0.8824	0.0781	0.74175	1.000
0.0976	15	1	0.8235	0.0925	0.66087	1.000
0.1439	14	1	0.7647	0.1029	0.58746	0.995
0.1948	13	1	0.7059	0.1105	0.51936	0.959
0.2797	12	1	0.6471	0.1159	0.45548	0.919
0.3092	11	1	0.5882	0.1194	0.39521	0.876
0.4416	10	1	0.5294	0.1211	0.33818	0.829
0.4829	9	1	0.4706	0.1211	0.28423	0.779
0.5727	8	1	0.4118	0.1194	0.23329	0.727
0.6872	7	1	0.3529	0.1159	0.18543	0.672
0.7572	6	1	0.2941	0.1105	0.14083	0.614
0.8301	5	1	0.2353	0.1029	0.09987	0.554
0.9163	4	1	0.1765	0.0925	0.06320	0.493
1.0051	3	1	0.1176	0.0781	0.03200	0.432
1.2983	2	1	0.0588	0.0571	0.00879	0.394
1.7779	1	1	0.0000	NA	NA	NA

Lampiran 8. Data berdistribusi Weibull dengan parameter $\lambda = 6$ dan $p = 2$

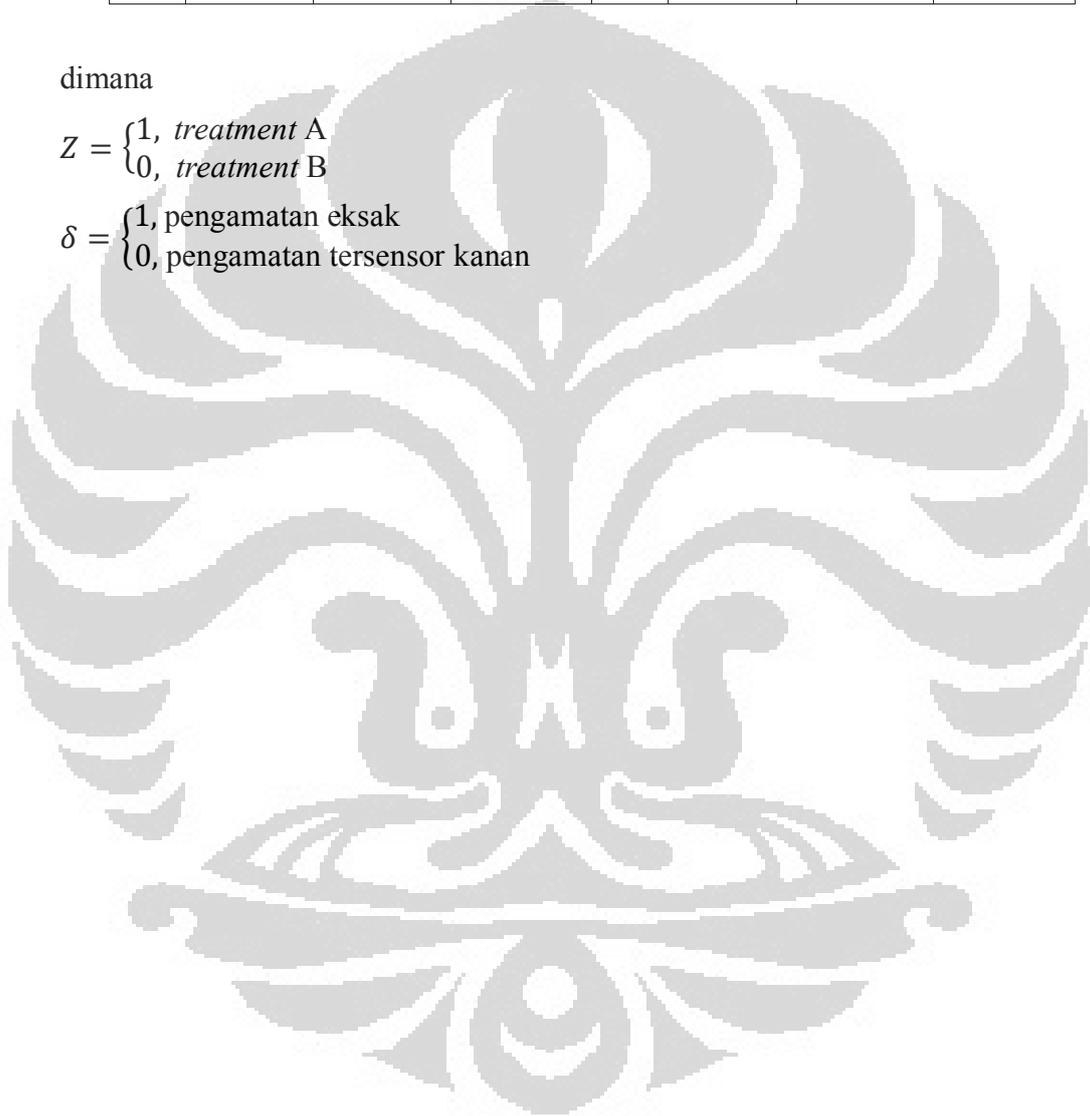
No.	<i>time</i> (<i>x</i>)	TRT (<i>Z</i>)	<i>event</i> (δ)	No.	<i>time</i> (<i>x</i>)	TRT (<i>Z</i>)	<i>event</i> (δ)
1	0.62	0	1	1	1.06	1	1
2	0.70	0	1	2	1.77	1	1
3	0.83	0	1	3	1.99	1	1
4	1.34	0	1	4	2.21	1	1
5	1.40	0	1	5	2.71	1	1
6	1.51	0	0	6	3.05	1	0
7	1.54	0	1	7	3.20	1	1
8	1.55	0	0	8	3.36	1	1
9	1.57	0	1	9	3.41	1	1
10	1.96	0	1	10	3.64	1	1
11	2.14	0	1	11	3.87	1	1
12	2.15	0	0	12	4.19	1	1
13	2.21	0	1	13	4.43	1	1
14	2.24	0	0	14	4.44	1	0
15	2.40	0	1	15	4.82	1	1
16	2.42	0	1	16	4.88	1	1
17	2.45	0	1	17	5.07	1	0
18	2.79	0	1	18	5.09	1	1
19	2.81	0	1	19	5.21	1	0
20	3.08	0	1	20	5.25	1	1
21	3.17	0	1	21	5.56	1	1
22	3.24	0	1	22	5.69	1	0
23	3.27	0	1	23	5.70	1	1
24	3.34	0	1	24	5.77	1	0
25	3.35	0	0	25	6.45	1	0
26	3.38	0	1	26	6.49	1	1
27	3.46	0	0	27	7.20	1	1
28	3.58	0	0	28	7.32	1	1
29	3.68	0	1	29	7.33	1	0
30	4.14	0	1	30	7.95	1	1
31	4.35	0	1	31	8.12	1	1
32	4.42	0	1	32	8.15	1	1
33	4.53	0	1	33	8.17	1	1
34	4.71	0	1	34	8.21	1	1
35	4.80	0	0	35	8.27	1	1
36	4.94	0	1	36	8.30	1	1
37	4.96	0	1	37	8.63	1	1
38	5.01	0	1	38	8.79	1	1
39	5.04	0	1	39	8.87	1	1
40	5.05	0	1	40	9.00	1	1
41	5.21	0	1	41	9.84	1	1

42	5.23	0	1	42	10.52	1	0
43	5.34	0	1	43	10.76	1	1
44	5.67	0	0	44	11.01	1	1
45	6.07	0	1	45	11.97	1	1
46	6.28	0	1	46	12.02	1	1
47	6.33	0	0	47	13.44	1	1
48	6.74	0	1	48	15.09	1	1
49	6.76	0	1	49	17.58	1	0
50	8.15	0	1	50	17.99	1	1

dimana

$$Z = \begin{cases} 1, & \text{treatment A} \\ 0, & \text{treatment B} \end{cases}$$

$$\delta = \begin{cases} 1, & \text{pengamatan eksak} \\ 0, & \text{pengamatan tersensor kanan} \end{cases}$$



Lampiran 9. Data berdistribusi lognormal dengan parameter $\mu = 2.5$ dan $\sigma = 1$

No.	<i>time</i> (<i>x</i>)	TRT (<i>Z</i>)	<i>event</i> (δ)	No.	<i>time</i> (<i>x</i>)	TRT (<i>Z</i>)	<i>event</i> (δ)
1	0.75	0	1	1	0.78	1	1
2	1.05	0	0	2	2.33	1	1
3	1.49	0	1	3	3.20	1	0
4	1.59	0	1	4	3.52	1	1
5	1.89	0	1	5	3.71	1	1
6	1.91	0	1	6	4.01	1	1
7	1.97	0	1	7	4.39	1	1
8	2.01	0	0	8	5.01	1	1
9	2.01	0	1	9	5.14	1	1
10	2.02	0	1	10	5.42	1	1
11	2.21	0	1	11	5.60	1	1
12	2.28	0	1	12	5.66	1	1
13	3.33	0	1	13	6.10	1	1
14	4.28	0	1	14	6.70	1	1
15	4.38	0	0	15	9.45	1	0
16	4.42	0	1	16	9.61	1	1
17	4.74	0	1	17	9.83	1	1
18	5.10	0	0	18	9.90	1	1
19	6.84	0	1	19	11.65	1	1
20	7.26	0	1	20	12.43	1	1
21	7.54	0	1	21	13.53	1	0
22	8.14	0	0	22	14.17	1	1
23	8.18	0	1	23	14.18	1	1
24	8.36	0	1	24	14.49	1	1
25	9.06	0	1	25	16.62	1	1
26	9.75	0	1	26	20.51	1	1
27	9.97	0	1	27	24.37	1	1
28	10.17	0	1	28	24.90	1	1
29	10.46	0	0	29	25.84	1	0
30	11.30	0	1	30	25.95	1	1
31	11.46	0	1	31	28.68	1	1
32	12.66	0	1	32	28.79	1	1
33	14.23	0	0	33	29.80	1	1
34	14.48	0	1	34	30.19	1	1
35	15.71	0	1	35	33.19	1	1
36	16.58	0	0	36	35.39	1	0
37	17.50	0	0	37	35.39	1	1
38	18.51	0	1	38	35.88	1	1
39	18.81	0	1	39	40.74	1	0
40	19.76	0	1	40	41.82	1	1
41	19.85	0	1	41	47.90	1	1
42	21.79	0	1	42	60.89	1	0
43	23.53	0	1	43	61.44	1	1

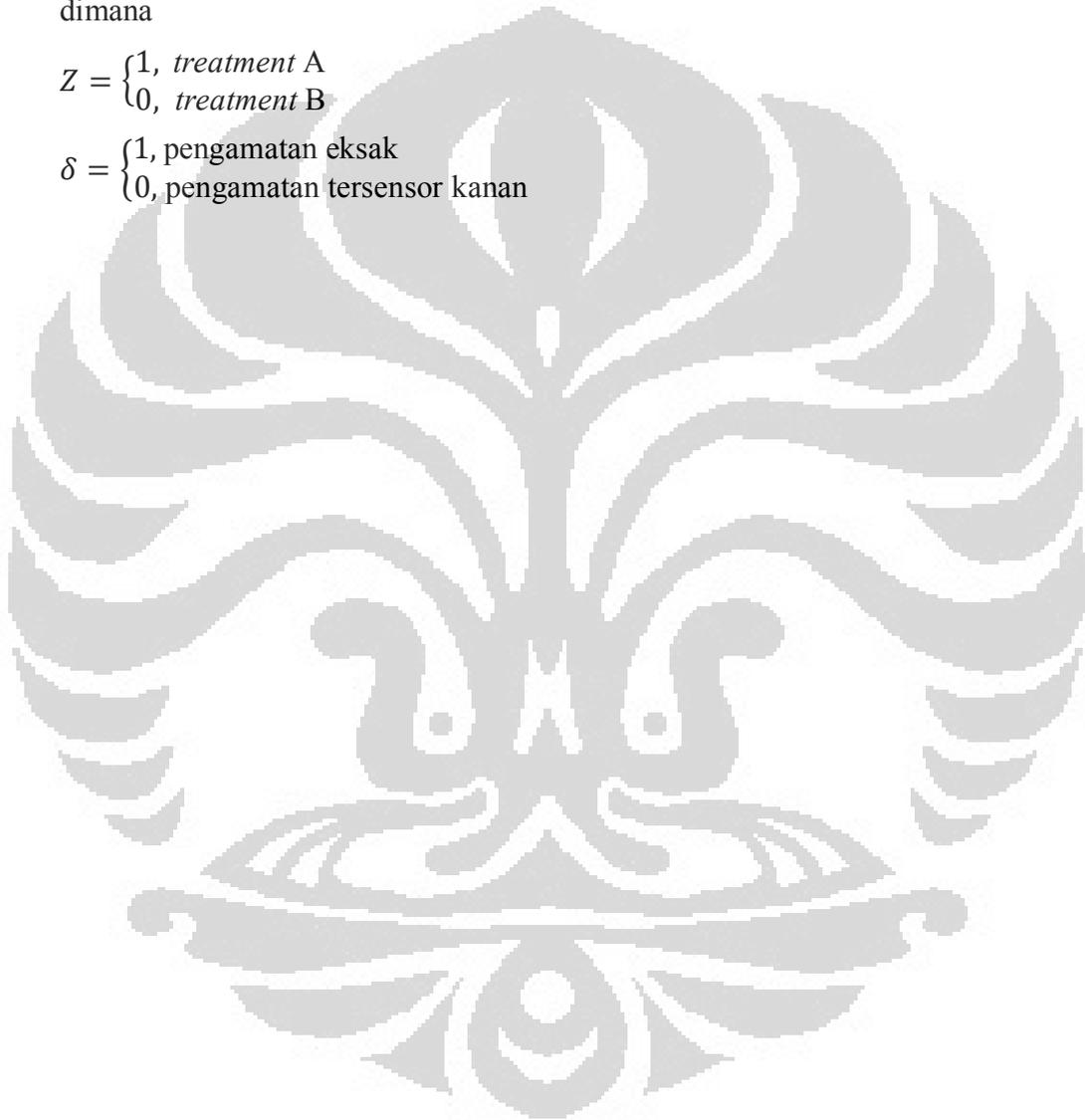
Universitas Indonesia

44	24.79	0	1	44	64.08	1	1
45	25.01	0	1	45	65.89	1	0
46	26.90	0	1	46	69.20	1	1
47	31.15	0	0	47	80.97	1	1
48	31.59	0	1	48	83.98	1	0
49	36.12	0	1	49	86.38	1	1
50	42.59	0	1	50	129.47	1	0

dimana

$$Z = \begin{cases} 1, & \text{treatment A} \\ 0, & \text{treatment B} \end{cases}$$

$$\delta = \begin{cases} 1, & \text{pengamatan eksak} \\ 0, & \text{pengamatan tersensor kanan} \end{cases}$$

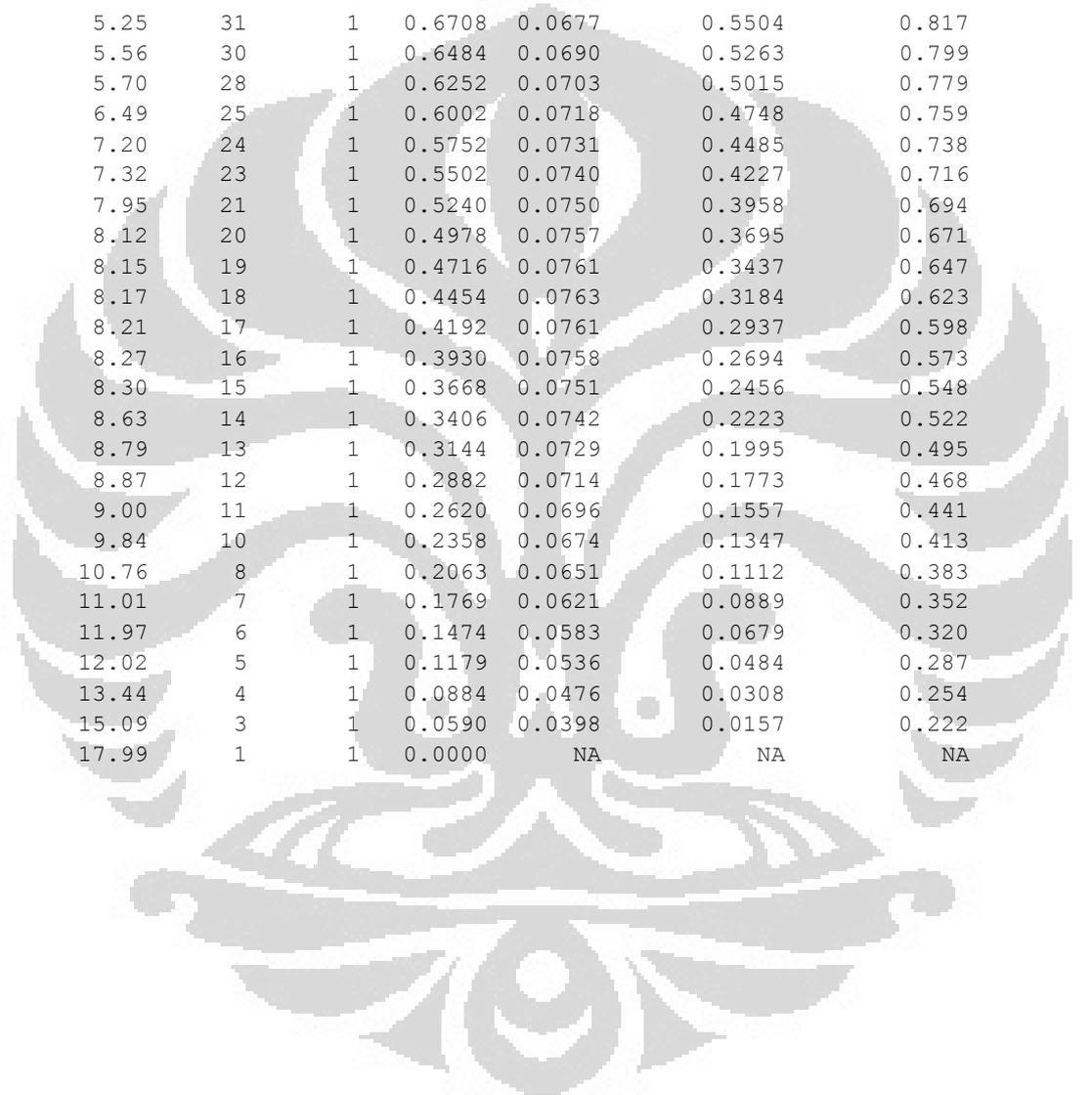


Lampiran 10. Estimasi Kaplan-Meier pada data Weibull

```
> KM.wei <- survfit(Surv(time, event) ~ TRT, weibull)
> summary(KM.wei)
Call: survfit(formula = Surv(time, event) ~ TRT, data = weibull)
```

```
TRT=0
time n.risk n.event survival std.err lower 95% CI upper 95% CI
0.62 50 1 0.9800 0.0198 0.94195 1.000
0.70 49 1 0.9600 0.0277 0.90719 1.000
0.83 48 1 0.9400 0.0336 0.87643 1.000
1.34 47 1 0.9200 0.0384 0.84779 0.998
1.40 46 1 0.9000 0.0424 0.82057 0.987
1.54 44 1 0.8795 0.0461 0.79362 0.975
1.57 42 1 0.8586 0.0496 0.76677 0.961
1.96 41 1 0.8377 0.0526 0.74068 0.947
2.14 40 1 0.8167 0.0553 0.71524 0.933
2.21 38 1 0.7952 0.0579 0.68954 0.917
2.40 36 1 0.7731 0.0603 0.66351 0.901
2.42 35 1 0.7510 0.0625 0.63800 0.884
2.45 34 1 0.7290 0.0645 0.61297 0.867
2.79 33 1 0.7069 0.0662 0.58836 0.849
2.81 32 1 0.6848 0.0677 0.56416 0.831
3.08 31 1 0.6627 0.0690 0.54032 0.813
3.17 30 1 0.6406 0.0702 0.51683 0.794
3.24 29 1 0.6185 0.0711 0.49367 0.775
3.27 28 1 0.5964 0.0719 0.47083 0.756
3.34 27 1 0.5743 0.0726 0.44830 0.736
3.38 25 1 0.5514 0.0732 0.42498 0.715
3.68 22 1 0.5263 0.0741 0.39942 0.693
4.14 21 1 0.5012 0.0747 0.37432 0.671
4.35 20 1 0.4762 0.0750 0.34967 0.648
4.42 19 1 0.4511 0.0751 0.32546 0.625
4.53 18 1 0.4260 0.0750 0.30169 0.602
4.71 17 1 0.4010 0.0747 0.27835 0.578
4.94 15 1 0.3743 0.0743 0.25357 0.552
4.96 14 1 0.3475 0.0737 0.22936 0.527
5.01 13 1 0.3208 0.0727 0.20574 0.500
5.04 12 1 0.2941 0.0714 0.18272 0.473
5.05 11 1 0.2673 0.0697 0.16034 0.446
5.21 10 1 0.2406 0.0677 0.13862 0.418
5.23 9 1 0.2139 0.0652 0.11763 0.389
5.34 8 1 0.1871 0.0623 0.09743 0.359
6.07 6 1 0.1559 0.0592 0.07408 0.328
6.28 5 1 0.1248 0.0550 0.05259 0.296
6.74 3 1 0.0832 0.0500 0.02562 0.270
6.76 2 1 0.0416 0.0386 0.00675 0.256
8.15 1 1 0.0000 NA NA NA
```

```
TRT=1
time n.risk n.event survival std.err lower 95% CI upper 95% CI
1.06 50 1 0.9800 0.0198 0.9420 1.000
1.77 49 1 0.9600 0.0277 0.9072 1.000
1.99 48 1 0.9400 0.0336 0.8764 1.000
2.21 47 1 0.9200 0.0384 0.8478 0.998
```



2.71	46	1	0.9000	0.0424	0.8206	0.987
3.20	44	1	0.8795	0.0461	0.7936	0.975
3.36	43	1	0.8591	0.0494	0.7676	0.962
3.41	42	1	0.8386	0.0523	0.7422	0.948
3.64	41	1	0.8182	0.0549	0.7174	0.933
3.87	40	1	0.7977	0.0572	0.6932	0.918
4.19	39	1	0.7773	0.0593	0.6694	0.903
4.43	38	1	0.7568	0.0611	0.6460	0.887
4.82	36	1	0.7358	0.0629	0.6222	0.870
4.88	35	1	0.7148	0.0646	0.5988	0.853
5.09	33	1	0.6931	0.0661	0.5749	0.836
5.25	31	1	0.6708	0.0677	0.5504	0.817
5.56	30	1	0.6484	0.0690	0.5263	0.799
5.70	28	1	0.6252	0.0703	0.5015	0.779
6.49	25	1	0.6002	0.0718	0.4748	0.759
7.20	24	1	0.5752	0.0731	0.4485	0.738
7.32	23	1	0.5502	0.0740	0.4227	0.716
7.95	21	1	0.5240	0.0750	0.3958	0.694
8.12	20	1	0.4978	0.0757	0.3695	0.671
8.15	19	1	0.4716	0.0761	0.3437	0.647
8.17	18	1	0.4454	0.0763	0.3184	0.623
8.21	17	1	0.4192	0.0761	0.2937	0.598
8.27	16	1	0.3930	0.0758	0.2694	0.573
8.30	15	1	0.3668	0.0751	0.2456	0.548
8.63	14	1	0.3406	0.0742	0.2223	0.522
8.79	13	1	0.3144	0.0729	0.1995	0.495
8.87	12	1	0.2882	0.0714	0.1773	0.468
9.00	11	1	0.2620	0.0696	0.1557	0.441
9.84	10	1	0.2358	0.0674	0.1347	0.413
10.76	8	1	0.2063	0.0651	0.1112	0.383
11.01	7	1	0.1769	0.0621	0.0889	0.352
11.97	6	1	0.1474	0.0583	0.0679	0.320
12.02	5	1	0.1179	0.0536	0.0484	0.287
13.44	4	1	0.0884	0.0476	0.0308	0.254
15.09	3	1	0.0590	0.0398	0.0157	0.222
17.99	1	1	0.0000	NA	NA	NA

Lampiran 11. Estimasi model Weibull AFT pada data Weibull

```
> attach(weibull)
> weib.aft2 <- survReg(Surv(time, event) ~ TRT, dist="weib")
> summary(weib.aft2)
```

Call:

```
survReg(formula = Surv(time, event) ~ TRT, dist = "weib")
```

	Value	Std. Error	z	p
(Intercept)	1.512	0.0773	19.56	3.62e-085
TRT	0.686	0.1090	6.30	3.05e-010
Log(scale)	-0.719	0.0868	-8.29	1.15e-016

Scale= 0.487

Weibull distribution

Loglik(model)= -203.4 Loglik(intercept only)= -218.1

Chisq= 29.44 on 1 degrees of freedom, p= 5.8e-008

Number of Newton-Raphson Iterations: 5

n= 100

Correlation of Coefficients:

	(Intercept)	TRT
TRT	-0.701	
Log(scale)	-0.085	-0.040

Lampiran 12. Estimasi model Cox PH pada data Weibull

```
> cox2 <- coxph(Surv(time, event) ~ TRT, weibull)
> summary(cox2)
Call:
coxph(formula = Surv(time, event) ~ TRT, data = weibull)

n= 100

      coef exp(coef) se(coef)      z      p
TRT -1.45    0.235    0.281 -5.15 2.6e-007

      exp(coef) exp(-coef) lower .95 upper .95
TRT    0.235      4.26    0.135    0.407

Rsquare= 0.247 (max possible= 0.997 )
Likelihood ratio test= 28.4 on 1 df, p=9.97e-008
Wald test              = 26.5 on 1 df, p=2.58e-007
Score (logrank) test = 30.3 on 1 df, p=3.75e-008
```

Lampiran 13. Estimasi fungsi *survival* dari Cox-Snell residual untuk model Weibull AFT pada data Weibull

```
> KM.wei2 <- survfit(Surv(rc), data = AFT2)
> summary(KM.wei2)
Call: survfit(formula = Surv(rc), data = AFT2)
```

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0.0124	100	1	0.99	0.00995	0.97069	1.0000
0.0169	99	1	0.98	0.01400	0.95294	1.0000
0.0217	98	1	0.97	0.01706	0.93714	1.0000
0.0303	97	1	0.96	0.01960	0.92235	0.9992
0.0356	96	1	0.95	0.02179	0.90823	0.9937
0.0448	95	1	0.94	0.02375	0.89459	0.9877
0.0558	94	1	0.93	0.02551	0.88131	0.9814
0.0816	93	1	0.92	0.02713	0.86833	0.9747
0.0847	92	1	0.91	0.02862	0.85560	0.9679
0.0898	91	1	0.90	0.03000	0.84308	0.9608
0.1043	90	1	0.89	0.03129	0.83074	0.9535
0.1079	89	1	0.88	0.03250	0.81856	0.9461
0.1095	88	1	0.87	0.03363	0.80652	0.9385
0.1105	87	1	0.86	0.03470	0.79461	0.9308
0.1137	86	1	0.85	0.03571	0.78282	0.9229
0.1196	85	1	0.84	0.03666	0.77113	0.9150
0.1320	84	1	0.83	0.03756	0.75955	0.9070
0.1362	83	1	0.82	0.03842	0.74805	0.8989
0.1553	82	1	0.81	0.03923	0.73665	0.8907
0.1761	81	1	0.80	0.04000	0.72532	0.8824
0.1777	80	1	0.79	0.04073	0.71407	0.8740
0.2078	79	1	0.78	0.04142	0.70289	0.8656
0.2142	78	1	0.77	0.04208	0.69178	0.8571
0.2162	77	1	0.76	0.04271	0.68074	0.8485
0.2295	76	1	0.75	0.04330	0.66976	0.8399
0.2328	75	1	0.74	0.04386	0.65884	0.8312
0.2335	74	1	0.73	0.04440	0.64797	0.8224
0.2355	73	1	0.72	0.04490	0.63716	0.8136
0.2708	72	1	0.71	0.04538	0.62641	0.8047
0.2741	71	1	0.70	0.04583	0.61571	0.7958
0.2767	70	1	0.69	0.04625	0.60505	0.7869
0.2819	69	1	0.68	0.04665	0.59445	0.7779
0.2838	68	1	0.67	0.04702	0.58390	0.7688
0.3074	67	1	0.66	0.04737	0.57339	0.7597
0.3102	66	1	0.65	0.04770	0.56293	0.7505
0.3247	65	1	0.64	0.04800	0.55251	0.7413
0.3304	64	1	0.63	0.04828	0.54214	0.7321
0.3682	63	1	0.62	0.04854	0.53181	0.7228
0.3712	62	1	0.61	0.04877	0.52152	0.7135
0.3753	61	1	0.60	0.04899	0.51127	0.7041
0.3891	60	1	0.59	0.04918	0.50107	0.6947
0.3914	59	1	0.58	0.04936	0.49090	0.6853
0.4002	58	1	0.57	0.04951	0.48078	0.6758
0.4509	57	1	0.56	0.04964	0.47069	0.6663
0.4795	56	1	0.55	0.04975	0.46065	0.6567
0.5022	55	1	0.54	0.04984	0.45064	0.6471
0.5040	54	1	0.53	0.04991	0.44068	0.6374

Universitas Indonesia

0.5104	53	1	0.52	0.04996	0.43075	0.6277
0.5123	52	1	0.51	0.04999	0.42086	0.6180
0.5345	51	1	0.50	0.05000	0.41101	0.6083
0.5358	50	1	0.49	0.04999	0.40120	0.5985
0.5475	49	1	0.48	0.04996	0.39142	0.5886
0.5748	48	1	0.47	0.04991	0.38169	0.5787
0.6156	47	1	0.46	0.04984	0.37199	0.5688
0.6318	46	1	0.45	0.04975	0.36233	0.5589
0.6519	45	1	0.44	0.04964	0.35272	0.5489
0.6534	44	1	0.43	0.04951	0.34314	0.5389
0.6555	43	1	0.42	0.04936	0.33360	0.5288
0.7743	42	1	0.41	0.04918	0.32410	0.5187
0.8080	41	1	0.40	0.04899	0.31464	0.5085
0.8146	40	1	0.39	0.04877	0.30522	0.4983
0.8195	39	1	0.38	0.04854	0.29584	0.4881
0.8269	38	1	0.37	0.04828	0.28650	0.4778
0.8279	37	1	0.36	0.04800	0.27721	0.4675
0.8401	36	1	0.35	0.04770	0.26796	0.4572
0.8452	35	1	0.34	0.04737	0.25875	0.4468
0.9162	34	1	0.33	0.04702	0.24959	0.4363
0.9182	33	1	0.32	0.04665	0.24047	0.4258
0.9466	32	1	0.31	0.04625	0.23140	0.4153
0.9506	31	1	0.30	0.04583	0.22238	0.4047
0.9695	30	1	0.29	0.04538	0.21341	0.3941
0.9983	29	1	0.28	0.04490	0.20449	0.3834
0.9987	28	1	0.27	0.04440	0.19561	0.3727
1.0802	27	1	0.26	0.04386	0.18680	0.3619
1.1253	26	1	0.25	0.04330	0.17804	0.3511
1.1940	25	1	0.24	0.04271	0.16933	0.3402
1.1999	24	1	0.23	0.04208	0.16069	0.3292
1.2000	23	1	0.22	0.04142	0.15211	0.3182
1.2252	22	1	0.21	0.04073	0.14359	0.3071
1.2408	21	1	0.20	0.04000	0.13514	0.2960
1.2443	20	1	0.19	0.03923	0.12677	0.2848
1.3268	19	1	0.18	0.03842	0.11847	0.2735
1.3405	18	1	0.17	0.03756	0.11025	0.2621
1.3763	17	1	0.16	0.03666	0.10211	0.2507
1.3998	16	1	0.15	0.03571	0.09407	0.2392
1.4396	15	1	0.14	0.03470	0.08613	0.2276
1.5105	14	1	0.13	0.03363	0.07830	0.2158
1.5795	13	1	0.12	0.03250	0.07058	0.2040
1.7917	12	1	0.11	0.03129	0.06299	0.1921
1.8080	11	1	0.10	0.03000	0.05554	0.1800
1.8173	10	1	0.09	0.02862	0.04826	0.1678
1.9513	9	1	0.08	0.02713	0.04116	0.1555
1.9851	8	1	0.07	0.02551	0.03426	0.1430
2.2526	7	1	0.06	0.02375	0.02762	0.1303
2.2691	6	1	0.05	0.02179	0.02128	0.1175
2.2754	5	1	0.04	0.01960	0.01531	0.1045
2.8861	4	1	0.03	0.01706	0.00984	0.0914
3.3278	3	1	0.02	0.01400	0.00507	0.0789
3.9482	2	1	0.01	0.00995	0.00142	0.0703
4.1373	1	1	0.00	NA	NA	NA

Lampiran 14. Estimasi *survival* dari model Cox PH pada data Weibull

```
> summary(survfit(cox2))
```

```
Call: survfit.coxph(object = cox2)
```

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0.62	100	1	0.992184	0.00781	9.77e-001	1.0000
0.70	99	1	0.984302	0.01110	9.63e-001	1.0000
0.83	98	1	0.976352	0.01364	9.50e-001	1.0000
1.06	97	1	0.968333	0.01581	9.38e-001	0.9998
1.34	96	1	0.960347	0.01770	9.26e-001	0.9957
1.40	95	1	0.952291	0.01942	9.15e-001	0.9911
1.54	93	1	0.944015	0.02108	9.04e-001	0.9863
1.57	91	1	0.935506	0.02269	8.92e-001	0.9811
1.77	90	1	0.926914	0.02421	8.81e-001	0.9756
1.96	89	1	0.918362	0.02562	8.70e-001	0.9700
1.99	88	1	0.909725	0.02696	8.58e-001	0.9641
2.14	87	1	0.901131	0.02823	8.47e-001	0.9582
2.21	85	2	0.883388	0.03070	8.25e-001	0.9457
2.40	82	1	0.874290	0.03190	8.14e-001	0.9391
2.42	81	1	0.865091	0.03307	8.03e-001	0.9324
2.45	80	1	0.855786	0.03420	7.91e-001	0.9255
2.71	79	1	0.846373	0.03530	7.80e-001	0.9185
2.79	78	1	0.837013	0.03635	7.69e-001	0.9114
2.81	77	1	0.827540	0.03736	7.57e-001	0.9041
3.08	75	1	0.817896	0.03837	7.46e-001	0.8967
3.17	74	1	0.808130	0.03936	7.35e-001	0.8891
3.20	73	1	0.798237	0.04032	7.23e-001	0.8813
3.24	72	1	0.788406	0.04123	7.12e-001	0.8735
3.27	71	1	0.778443	0.04213	7.00e-001	0.8655
3.34	70	1	0.768343	0.04300	6.89e-001	0.8574
3.36	68	1	0.757809	0.04390	6.76e-001	0.8489
3.38	67	1	0.747350	0.04475	6.65e-001	0.8404
3.41	66	1	0.736733	0.04557	6.53e-001	0.8317
3.64	63	1	0.725533	0.04646	6.40e-001	0.8225
3.68	62	1	0.714421	0.04728	6.28e-001	0.8134
3.87	61	1	0.703123	0.04808	6.15e-001	0.8040
4.14	60	1	0.691917	0.04882	6.03e-001	0.7945
4.19	59	1	0.680514	0.04955	5.90e-001	0.7849
4.35	58	1	0.669210	0.05022	5.78e-001	0.7752
4.42	57	1	0.657698	0.05086	5.65e-001	0.7653
4.43	56	1	0.645968	0.05149	5.53e-001	0.7552
4.53	54	1	0.634244	0.05209	5.40e-001	0.7450
4.71	53	1	0.622285	0.05267	5.27e-001	0.7346
4.82	51	1	0.609561	0.05328	5.14e-001	0.7235
4.88	50	1	0.596973	0.05384	5.00e-001	0.7124
4.94	49	1	0.584519	0.05432	4.87e-001	0.7013
4.96	48	1	0.571777	0.05479	4.74e-001	0.6899
5.01	47	1	0.558726	0.05522	4.60e-001	0.6782
5.04	46	1	0.545341	0.05563	4.47e-001	0.6660
5.05	45	1	0.531598	0.05601	4.32e-001	0.6535
5.09	43	1	0.517281	0.05638	4.18e-001	0.6405
5.21	42	1	0.503166	0.05668	4.03e-001	0.6275
5.23	40	1	0.488410	0.05697	3.89e-001	0.6139
5.25	39	1	0.473164	0.05723	3.73e-001	0.5997

5.34	38	1	0.458166	0.05739	3.58e-001	0.5857
5.56	37	1	0.442628	0.05750	3.43e-001	0.5710
5.70	34	1	0.425895	0.05768	3.27e-001	0.5554
6.07	32	1	0.409183	0.05779	3.10e-001	0.5397
6.28	31	1	0.391712	0.05784	2.93e-001	0.5232
6.49	28	1	0.370884	0.05818	2.73e-001	0.5044
6.74	27	1	0.350643	0.05824	2.53e-001	0.4856
6.76	26	1	0.329078	0.05827	2.33e-001	0.4656
7.20	25	1	0.305897	0.05831	2.11e-001	0.4445
7.32	24	1	0.283588	0.05799	1.90e-001	0.4234
7.95	22	1	0.261334	0.05745	1.70e-001	0.4021
8.12	21	1	0.240017	0.05656	1.51e-001	0.3809
8.15	20	2	0.198720	0.05400	1.17e-001	0.3385
8.17	18	1	0.177187	0.05253	9.91e-002	0.3168
8.21	17	1	0.156925	0.05065	8.34e-002	0.2954
8.27	16	1	0.137929	0.04840	6.93e-002	0.2744
8.30	15	1	0.120195	0.04580	5.70e-002	0.2537
8.63	14	1	0.103716	0.04291	4.61e-002	0.2333
8.79	13	1	0.088486	0.03976	3.67e-002	0.2135
8.87	12	1	0.074501	0.03639	2.86e-002	0.1940
9.00	11	1	0.061753	0.03284	2.18e-002	0.1751
9.84	10	1	0.050234	0.02916	1.61e-002	0.1567
10.76	8	1	0.038808	0.02517	1.09e-002	0.1383
11.01	7	1	0.028896	0.02104	6.94e-003	0.1204
11.97	6	1	0.020484	0.01688	4.07e-003	0.1030
12.02	5	1	0.013555	0.01280	2.13e-003	0.0863
13.44	4	1	0.008090	0.00894	9.27e-004	0.0706
15.09	3	1	0.004065	0.00545	2.94e-004	0.0562
17.99	1	1	0.000516	0.00131	3.52e-006	0.0755

Lampiran 15. Estimasi fungsi *survival* dari Cox-Snell residual untuk model Cox PH pada data Weibull

```
> KM.wei3 <- survfit(Surv(rc), data = PH2)
> summary(KM.wei3)
Call: survfit(formula = Surv(rc), data = PH2)
```

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0.00785	78	1	0.9872	0.0127	0.96253	1.0000
0.01582	77	1	0.9744	0.0179	0.93991	1.0000
0.02393	76	1	0.9615	0.0218	0.91979	1.0000
0.03218	75	1	0.9487	0.0250	0.90101	0.9990
0.04046	74	1	0.9359	0.0277	0.88309	0.9919
0.04888	73	1	0.9231	0.0302	0.86580	0.9841
0.05761	72	1	0.9103	0.0324	0.84899	0.9759
0.06667	71	1	0.8974	0.0344	0.83257	0.9674
0.07589	70	1	0.8846	0.0362	0.81648	0.9584
0.08516	69	1	0.8718	0.0379	0.80067	0.9492
0.09461	68	1	0.8590	0.0394	0.78511	0.9398
0.10410	67	1	0.8462	0.0409	0.76976	0.9301
0.12399	66	1	0.8333	0.0422	0.75460	0.9203
0.13434	65	1	0.8205	0.0435	0.73962	0.9103
0.14492	64	1	0.8077	0.0446	0.72480	0.9001
0.15573	63	1	0.7949	0.0457	0.71013	0.8897
0.16680	62	1	0.7821	0.0467	0.69559	0.8793
0.17792	61	1	0.7692	0.0477	0.68119	0.8687
0.18930	60	1	0.7564	0.0486	0.66690	0.8579
0.20102	59	1	0.7436	0.0494	0.65274	0.8471
0.21303	58	1	0.7308	0.0502	0.63868	0.8361
0.22535	57	1	0.7179	0.0510	0.62472	0.8251
0.23774	56	1	0.7051	0.0516	0.61086	0.8139
0.25046	55	1	0.6923	0.0523	0.59710	0.8027
0.26352	54	1	0.6795	0.0528	0.58343	0.7914
0.27732	53	1	0.6667	0.0534	0.56985	0.7799
0.29122	52	1	0.6538	0.0539	0.55635	0.7684
0.30553	51	1	0.6410	0.0543	0.54294	0.7568
0.32085	50	1	0.6282	0.0547	0.52961	0.7452
0.33628	49	1	0.6154	0.0551	0.51636	0.7334
0.35222	48	1	0.6026	0.0554	0.50319	0.7216
0.36829	47	1	0.5897	0.0557	0.49009	0.7097
0.38491	46	1	0.5769	0.0559	0.47707	0.6977
0.40166	45	1	0.5641	0.0561	0.46413	0.6856
0.41901	44	1	0.5513	0.0563	0.45125	0.6735
0.43701	43	1	0.5385	0.0564	0.43845	0.6613
0.45532	42	1	0.5256	0.0565	0.42573	0.6490
0.47436	41	1	0.5128	0.0566	0.41307	0.6367
0.49502	40	1	0.5000	0.0566	0.40049	0.6242
0.51588	39	1	0.4872	0.0566	0.38798	0.6117
0.53697	38	1	0.4744	0.0565	0.37554	0.5992
0.55901	37	1	0.4615	0.0564	0.36317	0.5866
0.58210	36	1	0.4487	0.0563	0.35087	0.5739
0.60634	35	1	0.4359	0.0561	0.33864	0.5611
0.63187	34	1	0.4231	0.0559	0.32649	0.5482
0.65917	33	1	0.4103	0.0557	0.31441	0.5353
0.68684	32	1	0.3974	0.0554	0.30241	0.5223

Universitas Indonesia

0.71660	31	1	0.3846	0.0551	0.29048	0.5093
0.74831	30	1	0.3718	0.0547	0.27863	0.4961
0.78052	29	1	0.3590	0.0543	0.26685	0.4829
0.81503	28	1	0.3462	0.0539	0.25516	0.4696
0.85356	27	1	0.3333	0.0534	0.24354	0.4562
0.89359	26	1	0.3205	0.0528	0.23201	0.4428
0.93723	25	1	0.3077	0.0523	0.22057	0.4292
0.99187	24	1	0.2949	0.0516	0.20922	0.4156
1.04799	23	1	0.2821	0.0510	0.19795	0.4019
1.11146	22	1	0.2692	0.0502	0.18678	0.3881
1.18451	21	1	0.2564	0.0494	0.17571	0.3742
1.26023	20	1	0.2436	0.0486	0.16475	0.3602
1.34196	19	1	0.2308	0.0477	0.15389	0.3461
1.42705	18	1	0.2179	0.0467	0.14315	0.3318
1.61586	17	1	0.2051	0.0457	0.13253	0.3175
1.73055	16	1	0.1923	0.0446	0.12203	0.3031
1.85199	15	1	0.1795	0.0435	0.11168	0.2885
1.98102	14	1	0.1667	0.0422	0.10147	0.2738
2.11864	13	1	0.1538	0.0409	0.09142	0.2589
2.26610	12	1	0.1410	0.0394	0.08155	0.2439
2.42491	11	1	0.1282	0.0379	0.07188	0.2287
2.59694	10	1	0.1154	0.0362	0.06241	0.2133
2.78461	9	1	0.1026	0.0344	0.05320	0.1977
2.99106	8	1	0.0897	0.0324	0.04426	0.1819
3.24913	7	1	0.0769	0.0302	0.03566	0.1659
3.54405	6	1	0.0641	0.0277	0.02745	0.1497
3.88811	5	1	0.0513	0.0250	0.01974	0.1332
4.30100	4	1	0.0385	0.0218	0.01268	0.1167
4.81713	3	1	0.0256	0.0179	0.00653	0.1007
5.50534	2	1	0.0128	0.0127	0.00183	0.0899
7.56940	1	1	0.0000	NA	NA	NA

Lampiran 16. Estimasi Kaplan-Meier pada data lognormal

```
> KM.log <- survfit(Surv(time, event) ~ TRT, lognormal)
> summary(KM.log)
Call: survfit(formula = Surv(time, event) ~ TRT, data = lognormal)
```

```

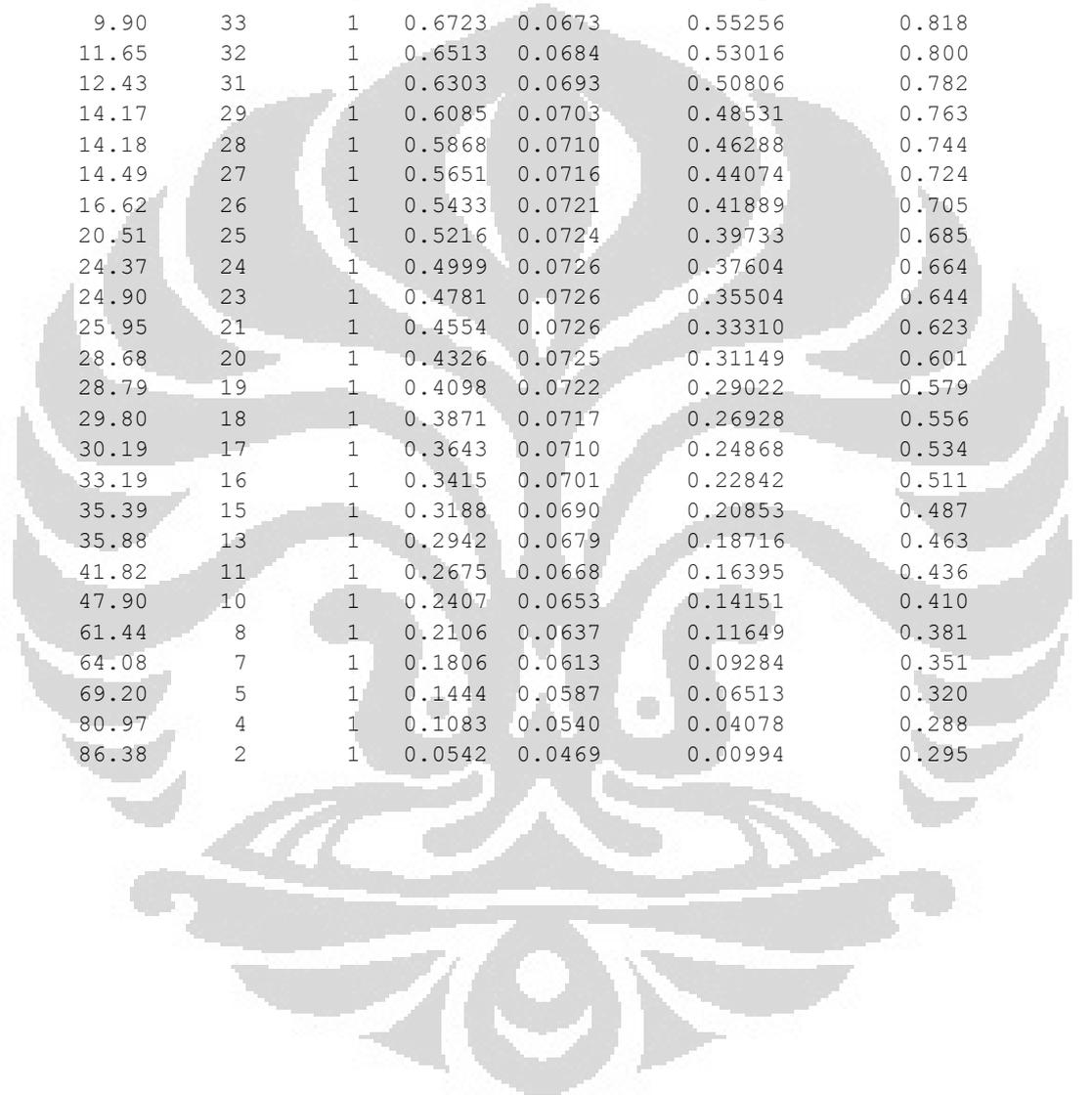
TRT=0
time n.risk n.event survival std.err lower 95% CI upper 95% CI
0.75  50    1    0.9800  0.0198  0.94195  1.000
1.49  48    1    0.9596  0.0280  0.90624  1.000
1.59  47    1    0.9392  0.0340  0.87476  1.000
1.89  46    1    0.9188  0.0389  0.84550  0.998
1.91  45    1    0.8983  0.0431  0.81771  0.987
1.97  44    1    0.8779  0.0467  0.79098  0.974
2.01  43    1    0.8575  0.0499  0.76509  0.961
2.02  41    1    0.8366  0.0529  0.73912  0.947
2.21  40    1    0.8157  0.0555  0.71378  0.932
2.28  39    1    0.7948  0.0579  0.68898  0.917
3.33  38    1    0.7738  0.0600  0.66466  0.901
4.28  37    1    0.7529  0.0620  0.64078  0.885
4.42  35    1    0.7314  0.0638  0.61645  0.868
4.74  34    1    0.7099  0.0655  0.59252  0.851
6.84  32    1    0.6877  0.0671  0.56806  0.833
7.26  31    1    0.6655  0.0685  0.54399  0.814
7.54  30    1    0.6433  0.0697  0.52027  0.796
8.18  28    1    0.6204  0.0709  0.49588  0.776
8.36  27    1    0.5974  0.0719  0.47187  0.756
9.06  26    1    0.5744  0.0727  0.44821  0.736
9.75  25    1    0.5514  0.0733  0.42490  0.716
9.97  24    1    0.5285  0.0738  0.40193  0.695
10.17 23    1    0.5055  0.0741  0.37929  0.674
11.30 21    1    0.4814  0.0744  0.35567  0.652
11.46 20    1    0.4573  0.0744  0.33244  0.629
12.66 19    1    0.4333  0.0743  0.30959  0.606
14.48 17    1    0.4078  0.0742  0.28549  0.582
15.71 16    1    0.3823  0.0738  0.26188  0.558
18.51 13    1    0.3529  0.0737  0.23430  0.532
18.81 12    1    0.3235  0.0732  0.20757  0.504
19.76 11    1    0.2941  0.0722  0.18171  0.476
19.85 10    1    0.2647  0.0707  0.15674  0.447
21.79  9    1    0.2353  0.0687  0.13271  0.417
23.53  8    1    0.2059  0.0661  0.10968  0.386
24.79  7    1    0.1764  0.0629  0.08776  0.355
25.01  6    1    0.1470  0.0589  0.06708  0.322
26.90  5    1    0.1176  0.0539  0.04788  0.289
31.59  3    1    0.0784  0.0482  0.02354  0.261
36.12  2    1    0.0392  0.0367  0.00626  0.246
42.59  1    1    0.0000    NA      NA      NA

```

```

TRT=1
time n.risk n.event survival std.err lower 95% CI upper 95% CI
0.78  50    1    0.9800  0.0198  0.94195  1.000
2.33  49    1    0.9600  0.0277  0.90719  1.000
3.52  47    1    0.9396  0.0338  0.87557  1.000
3.71  46    1    0.9191  0.0388  0.84622  0.998

```



4.01	45	1	0.8987	0.0430	0.81836	0.987
4.39	44	1	0.8783	0.0466	0.79159	0.975
5.01	43	1	0.8579	0.0498	0.76566	0.961
5.14	42	1	0.8374	0.0526	0.74042	0.947
5.42	41	1	0.8170	0.0552	0.71577	0.933
5.60	40	1	0.7966	0.0574	0.69162	0.917
5.66	39	1	0.7762	0.0595	0.66792	0.902
6.10	38	1	0.7557	0.0613	0.64463	0.886
6.70	37	1	0.7353	0.0630	0.62170	0.870
9.61	35	1	0.7143	0.0646	0.59831	0.853
9.83	34	1	0.6933	0.0660	0.57527	0.836
9.90	33	1	0.6723	0.0673	0.55256	0.818
11.65	32	1	0.6513	0.0684	0.53016	0.800
12.43	31	1	0.6303	0.0693	0.50806	0.782
14.17	29	1	0.6085	0.0703	0.48531	0.763
14.18	28	1	0.5868	0.0710	0.46288	0.744
14.49	27	1	0.5651	0.0716	0.44074	0.724
16.62	26	1	0.5433	0.0721	0.41889	0.705
20.51	25	1	0.5216	0.0724	0.39733	0.685
24.37	24	1	0.4999	0.0726	0.37604	0.664
24.90	23	1	0.4781	0.0726	0.35504	0.644
25.95	21	1	0.4554	0.0726	0.33310	0.623
28.68	20	1	0.4326	0.0725	0.31149	0.601
28.79	19	1	0.4098	0.0722	0.29022	0.579
29.80	18	1	0.3871	0.0717	0.26928	0.556
30.19	17	1	0.3643	0.0710	0.24868	0.534
33.19	16	1	0.3415	0.0701	0.22842	0.511
35.39	15	1	0.3188	0.0690	0.20853	0.487
35.88	13	1	0.2942	0.0679	0.18716	0.463
41.82	11	1	0.2675	0.0668	0.16395	0.436
47.90	10	1	0.2407	0.0653	0.14151	0.410
61.44	8	1	0.2106	0.0637	0.11649	0.381
64.08	7	1	0.1806	0.0613	0.09284	0.351
69.20	5	1	0.1444	0.0587	0.06513	0.320
80.97	4	1	0.1083	0.0540	0.04078	0.288
86.38	2	1	0.0542	0.0469	0.00994	0.295

Lampiran 17. Estimasi model parametrik Weibull AFT pada data lognormal

```
> attach(lognormal)
> weib.aft3 <- survReg(Surv(time, event) ~ TRT, dist="weib")
> summary(weib.aft3)
```

Call:

```
survReg(formula = Surv(time, event) ~ TRT, dist = "weib")
```

	Value	Std. Error	z	p
(Intercept)	2.7092	0.1521	17.811	5.80e-071
TRT	0.8613	0.2146	4.014	5.97e-005
Log(scale)	-0.0418	0.0884	-0.473	6.36e-001

Scale= 0.959

Weibull distribution

Loglik(model)= -330.5 Loglik(intercept only)= -337.4

Chisq= 13.84 on 1 degrees of freedom, p= 0.0002

Number of Newton-Raphson Iterations: 4

n= 100

Correlation of Coefficients:

	(Intercept)	TRT
TRT	-0.702	
Log(scale)	-0.079	-0.033

Lampiran 18. Estimasi model Cox PH pada data lognormal

```

> cox3 <- coxph(Surv(time, event) ~ TRT, lognormal)
> summary(cox3)
Call:
coxph(formula = Surv(time, event) ~ TRT, data = lognormal)

n= 100

      coef exp(coef) se(coef)      z      p
TRT -0.798    0.45    0.244 -3.28 0.0011

      exp(coef) exp(-coef) lower .95 upper .95
TRT    0.45      2.22    0.279    0.726

Rsquare= 0.102 (max possible= 0.997 )
Likelihood ratio test= 10.8 on 1 df, p=0.00103
Wald test               = 10.7 on 1 df, p=0.00105
Score (logrank) test = 11.2 on 1 df, p=0.000812

```

Lampiran 19. Estimasi fungsi *survival* dari Cox-Snell residual untuk model Weibull AFT pada data lognormal

```
> KM.log2 <- survfit(Surv(rc), data = AFT3)
> summary(KM.log2)
Call: survfit(formula = Surv(rc), data = AFT3)
```

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0.0186	100	1	0.99	0.00995	0.97069	1.0000
0.0437	99	1	0.98	0.01400	0.95294	1.0000
0.0584	98	1	0.97	0.01706	0.93714	1.0000
0.0627	97	1	0.96	0.01960	0.92235	0.9992
0.0811	96	1	0.95	0.02179	0.90823	0.9937
0.0899	95	1	0.94	0.02375	0.89459	0.9877
0.0901	94	1	0.93	0.02551	0.88131	0.9814
0.0948	93	1	0.92	0.02713	0.86833	0.9747
0.0965	92	1	0.91	0.02862	0.85560	0.9679
0.1028	91	1	0.90	0.03000	0.84308	0.9608
0.1131	90	1	0.89	0.03129	0.83074	0.9535
0.1151	89	1	0.88	0.03250	0.81856	0.9461
0.1167	88	1	0.87	0.03363	0.80652	0.9385
0.1202	87	1	0.86	0.03470	0.79461	0.9308
0.1229	86	1	0.85	0.03571	0.78282	0.9229
0.1231	85	1	0.84	0.03666	0.77113	0.9150
0.1233	84	1	0.83	0.03756	0.75955	0.9070
0.1296	83	1	0.82	0.03842	0.74805	0.8989
0.1333	82	1	0.81	0.03923	0.73665	0.8907
0.1358	81	1	0.80	0.04000	0.72532	0.8824
0.1399	80	1	0.79	0.04073	0.71407	0.8740
0.1408	79	1	0.78	0.04142	0.70289	0.8656
0.1456	78	1	0.77	0.04208	0.69178	0.8571
0.1473	77	1	0.76	0.04271	0.68074	0.8485
0.1591	76	1	0.75	0.04330	0.66976	0.8399
0.1755	75	1	0.74	0.04386	0.65884	0.8312
0.2079	74	1	0.73	0.04440	0.64797	0.8224
0.2513	73	1	0.72	0.04490	0.63716	0.8136
0.2558	72	1	0.71	0.04538	0.62641	0.8047
0.2617	71	1	0.70	0.04583	0.61571	0.7958
0.2638	70	1	0.69	0.04625	0.60505	0.7869
0.2702	69	1	0.68	0.04665	0.59445	0.7779
0.2768	68	1	0.67	0.04702	0.58390	0.7688
0.2794	67	1	0.66	0.04737	0.57339	0.7597
0.3002	66	1	0.65	0.04770	0.56293	0.7505
0.3126	65	1	0.64	0.04800	0.55251	0.7413
0.3240	64	1	0.63	0.04828	0.54214	0.7321
0.3345	63	1	0.62	0.04854	0.53181	0.7228
0.3654	62	1	0.61	0.04877	0.52152	0.7135
0.3835	61	1	0.60	0.04899	0.51127	0.7041
0.3837	60	1	0.59	0.04918	0.50107	0.6947
0.3924	59	1	0.58	0.04936	0.49090	0.6853
0.4403	58	1	0.57	0.04951	0.48078	0.6758
0.4528	57	1	0.56	0.04964	0.47069	0.6663
0.4687	56	1	0.55	0.04975	0.46065	0.6567
0.4872	55	1	0.54	0.04984	0.45064	0.6471
0.5280	54	1	0.53	0.04991	0.44068	0.6374

Universitas Indonesia

0.5305	53	1	0.52	0.04996	0.43075	0.6277
0.5428	52	1	0.51	0.04999	0.42086	0.6180
0.5639	51	1	0.50	0.05000	0.41101	0.6083
0.5905	50	1	0.49	0.04999	0.40120	0.5985
0.6376	49	1	0.48	0.04996	0.39142	0.5886
0.6523	48	1	0.47	0.04991	0.38169	0.5787
0.6662	47	1	0.46	0.04984	0.37199	0.5688
0.6748	46	1	0.45	0.04975	0.36233	0.5589
0.6859	45	1	0.44	0.04964	0.35272	0.5489
0.6901	44	1	0.43	0.04951	0.34314	0.5389
0.7175	43	1	0.42	0.04936	0.33360	0.5288
0.7204	42	1	0.41	0.04918	0.32410	0.5187
0.7435	41	1	0.40	0.04899	0.31464	0.5085
0.7547	40	1	0.39	0.04877	0.30522	0.4983
0.7999	39	1	0.38	0.04854	0.29584	0.4881
0.8031	38	1	0.37	0.04828	0.28650	0.4778
0.8324	37	1	0.36	0.04800	0.27721	0.4675
0.8370	36	1	0.35	0.04770	0.26796	0.4572
0.8436	35	1	0.34	0.04737	0.25875	0.4468
0.9312	34	1	0.33	0.04702	0.24959	0.4363
0.9452	33	1	0.32	0.04665	0.24047	0.4258
0.9627	32	1	0.31	0.04625	0.23140	0.4153
0.9957	31	1	0.30	0.04583	0.22238	0.4047
0.9957	30	1	0.29	0.04538	0.21341	0.3941
1.0100	29	1	0.28	0.04490	0.20449	0.3834
1.0480	28	1	0.27	0.04440	0.19561	0.3727
1.1087	27	1	0.26	0.04386	0.18680	0.3619
1.1533	26	1	0.25	0.04330	0.17804	0.3511
1.1728	25	1	0.24	0.04271	0.16933	0.3402
1.1851	24	1	0.23	0.04208	0.16069	0.3292
1.2439	23	1	0.22	0.04142	0.15211	0.3182
1.2647	22	1	0.21	0.04073	0.14359	0.3071
1.3315	21	1	0.20	0.04000	0.13514	0.2960
1.3380	20	1	0.19	0.03923	0.12677	0.2848
1.3654	19	1	0.18	0.03842	0.11847	0.2735
1.4743	18	1	0.17	0.03756	0.11025	0.2621
1.5970	17	1	0.16	0.03666	0.10211	0.2507
1.6866	16	1	0.15	0.03571	0.09407	0.2392
1.7023	15	1	0.14	0.03470	0.08613	0.2276
1.7536	14	1	0.13	0.03363	0.07830	0.2158
1.7698	13	1	0.12	0.03250	0.07058	0.2040
1.8362	12	1	0.11	0.03129	0.06299	0.1921
1.8495	11	1	0.10	0.03000	0.05554	0.1800
1.9039	10	1	0.09	0.02862	0.04826	0.1678
2.0036	9	1	0.08	0.02713	0.04116	0.1555
2.1400	8	1	0.07	0.02551	0.03426	0.1430
2.1715	7	1	0.06	0.02375	0.02762	0.1303
2.3602	6	1	0.05	0.02179	0.02128	0.1175
2.4520	5	1	0.04	0.01960	0.01531	0.1045
2.4972	4	1	0.03	0.01706	0.00984	0.0914
2.5251	3	1	0.02	0.01400	0.00507	0.0789
2.9654	2	1	0.01	0.00995	0.00142	0.0703
3.8507	1	1	0.00	NA	NA	NA

Lampiran 20. Estimasi *survival* dari model Cox PH pada data lognormal

```
> summary(survfit(cox3))
```

```
Call: survfit.coxph(object = cox3)
```

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0.75	100	1	0.9908	0.00918	0.97296	1.000
0.78	99	1	0.9815	0.01296	0.95646	1.000
1.49	97	1	0.9722	0.01589	0.94154	1.000
1.59	96	1	0.9628	0.01834	0.92751	0.999
1.89	95	1	0.9533	0.02048	0.91403	0.994
1.91	94	1	0.9438	0.02241	0.90094	0.989
1.97	93	1	0.9343	0.02417	0.88813	0.983
2.01	92	1	0.9247	0.02579	0.87553	0.977
2.02	90	1	0.9149	0.02735	0.86286	0.970
2.21	89	1	0.9051	0.02882	0.85034	0.963
2.28	88	1	0.8952	0.03020	0.83794	0.956
2.33	87	1	0.8853	0.03150	0.82563	0.949
3.33	85	1	0.8753	0.03273	0.81342	0.942
3.52	84	1	0.8652	0.03391	0.80127	0.934
3.71	83	1	0.8552	0.03501	0.78930	0.927
4.01	82	1	0.8453	0.03604	0.77750	0.919
4.28	81	1	0.8354	0.03702	0.76586	0.911
4.39	79	1	0.8252	0.03799	0.75399	0.903
4.42	78	1	0.8151	0.03890	0.74227	0.895
4.74	77	1	0.8049	0.03978	0.73057	0.887
5.01	76	1	0.7946	0.04063	0.71886	0.878
5.14	74	1	0.7842	0.04146	0.70704	0.870
5.42	73	1	0.7739	0.04224	0.69535	0.861
5.60	72	1	0.7636	0.04298	0.68379	0.853
5.66	71	1	0.7533	0.04367	0.67237	0.844
6.10	70	1	0.7431	0.04433	0.66107	0.835
6.70	69	1	0.7329	0.04495	0.64989	0.826
6.84	68	1	0.7228	0.04553	0.63882	0.818
7.26	67	1	0.7126	0.04608	0.62772	0.809
7.54	66	1	0.7023	0.04662	0.61660	0.800
8.18	64	1	0.6917	0.04717	0.60517	0.791
8.36	63	1	0.6810	0.04770	0.59370	0.781
9.06	62	1	0.6703	0.04820	0.58220	0.772
9.61	60	1	0.6594	0.04870	0.57051	0.762
9.75	59	1	0.6485	0.04916	0.55895	0.752
9.83	58	1	0.6375	0.04960	0.54735	0.743
9.90	57	1	0.6266	0.04999	0.53588	0.733
9.97	56	1	0.6157	0.05036	0.52454	0.723
10.17	55	1	0.6048	0.05070	0.51315	0.713
11.30	53	1	0.5934	0.05106	0.50135	0.702
11.46	52	1	0.5820	0.05139	0.48950	0.692
11.65	51	1	0.5704	0.05169	0.47758	0.681
12.43	50	1	0.5589	0.05195	0.46582	0.671
12.66	49	1	0.5475	0.05218	0.45421	0.660
14.17	47	1	0.5358	0.05240	0.44233	0.649
14.18	46	1	0.5242	0.05259	0.43060	0.638
14.48	44	1	0.5122	0.05278	0.41859	0.627
14.49	43	1	0.5002	0.05293	0.40649	0.615
15.71	42	1	0.4882	0.05305	0.39458	0.604

16.62	40	1	0.4756	0.05318	0.38205	0.592
18.51	38	1	0.4627	0.05331	0.36914	0.580
18.81	37	1	0.4495	0.05340	0.35614	0.567
19.76	36	1	0.4362	0.05345	0.34302	0.555
19.85	35	1	0.4226	0.05346	0.32979	0.542
20.51	34	1	0.4088	0.05344	0.31644	0.528
21.79	33	1	0.3952	0.05334	0.30335	0.515
23.53	32	1	0.3814	0.05320	0.29013	0.501
24.37	31	1	0.3673	0.05302	0.27674	0.487
24.79	30	1	0.3533	0.05277	0.26367	0.473
24.90	29	1	0.3391	0.05247	0.25041	0.459
25.01	28	1	0.3251	0.05210	0.23749	0.445
25.95	26	1	0.3104	0.05173	0.22387	0.430
26.90	25	1	0.2958	0.05127	0.21065	0.416
28.68	24	1	0.2810	0.05076	0.19719	0.400
28.79	23	1	0.2663	0.05015	0.18415	0.385
29.80	22	1	0.2520	0.04944	0.17154	0.370
30.19	21	1	0.2379	0.04863	0.15935	0.355
31.59	19	1	0.2227	0.04783	0.14623	0.339
33.19	18	1	0.2071	0.04697	0.13276	0.323
35.39	17	1	0.1918	0.04595	0.11992	0.307
35.88	15	1	0.1761	0.04484	0.10691	0.290
36.12	14	1	0.1608	0.04352	0.09463	0.273
41.82	12	1	0.1437	0.04223	0.08076	0.256
42.59	11	1	0.1272	0.04060	0.06803	0.238
47.90	10	1	0.1096	0.03887	0.05467	0.220
61.44	8	1	0.0909	0.03680	0.04115	0.201
64.08	7	1	0.0735	0.03399	0.02970	0.182
69.20	5	1	0.0546	0.03046	0.01827	0.163
80.97	4	1	0.0376	0.02564	0.00987	0.143
86.38	2	1	0.0178	0.01839	0.00236	0.135

Lampiran 21. Estimasi fungsi *survival* dari Cox-Snell residual untuk model Cox PH pada data lognormal

```
> KM.log3 <- survfit(Surv(rc), data = PH3)
> summary(KM.log3)
Call: survfit(formula = Surv(rc), data = PH3)
```

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0.00924	80	1	0.9875	0.0124	0.96345	1.0000
0.01867	79	1	0.9750	0.0175	0.94138	1.0000
0.02819	78	1	0.9625	0.0212	0.92176	1.0000
0.03791	77	1	0.9500	0.0244	0.90342	0.9990
0.04783	76	1	0.9375	0.0271	0.88593	0.9921
0.05784	75	1	0.9250	0.0294	0.86905	0.9846
0.06796	74	1	0.9125	0.0316	0.85264	0.9766
0.07829	73	1	0.9000	0.0335	0.83660	0.9682
0.08894	72	1	0.8875	0.0353	0.82089	0.9595
0.09971	71	1	0.8750	0.0370	0.80545	0.9506
0.11071	70	1	0.8625	0.0385	0.79024	0.9414
0.12183	69	1	0.8500	0.0399	0.77525	0.9320
0.13319	68	1	0.8375	0.0412	0.76044	0.9224
0.14479	67	1	0.8250	0.0425	0.74580	0.9126
0.15642	66	1	0.8125	0.0436	0.73132	0.9027
0.16806	65	1	0.8000	0.0447	0.71698	0.8926
0.17984	64	1	0.7875	0.0457	0.70277	0.8824
0.19213	63	1	0.7750	0.0467	0.68869	0.8721
0.20444	62	1	0.7625	0.0476	0.67473	0.8617
0.21704	61	1	0.7500	0.0484	0.66087	0.8511
0.22992	60	1	0.7375	0.0492	0.64712	0.8405
0.24309	59	1	0.7250	0.0499	0.63347	0.8298
0.25631	58	1	0.7125	0.0506	0.61992	0.8189
0.26971	57	1	0.7000	0.0512	0.60645	0.8080
0.28329	56	1	0.6875	0.0518	0.59308	0.7970
0.29692	55	1	0.6750	0.0524	0.57979	0.7858
0.31075	54	1	0.6625	0.0529	0.56658	0.7747
0.32462	53	1	0.6500	0.0533	0.55345	0.7634
0.33884	52	1	0.6375	0.0537	0.54040	0.7520
0.35339	51	1	0.6250	0.0541	0.52743	0.7406
0.36860	50	1	0.6125	0.0545	0.51453	0.7291
0.38419	49	1	0.6000	0.0548	0.50170	0.7176
0.40003	48	1	0.5875	0.0550	0.48895	0.7059
0.41642	47	1	0.5750	0.0553	0.47627	0.6942
0.43309	46	1	0.5625	0.0555	0.46365	0.6824
0.45020	45	1	0.5500	0.0556	0.45111	0.6706
0.46745	44	1	0.5375	0.0557	0.43863	0.6587
0.48500	43	1	0.5250	0.0558	0.42622	0.6467
0.50286	42	1	0.5125	0.0559	0.41388	0.6346
0.52189	41	1	0.5000	0.0559	0.40161	0.6225
0.54128	40	1	0.4875	0.0559	0.38940	0.6103
0.56142	39	1	0.4750	0.0558	0.37726	0.5981
0.58178	38	1	0.4625	0.0557	0.36519	0.5857
0.60239	37	1	0.4500	0.0556	0.35318	0.5734
0.62399	36	1	0.4375	0.0555	0.34125	0.5609
0.64588	35	1	0.4250	0.0553	0.32938	0.5484
0.66904	34	1	0.4125	0.0550	0.31758	0.5358

Universitas Indonesia

0.69275	33	1	0.4000	0.0548	0.30585	0.5231
0.71703	32	1	0.3875	0.0545	0.29419	0.5104
0.74318	31	1	0.3750	0.0541	0.28260	0.4976
0.77068	30	1	0.3625	0.0537	0.27108	0.4847
0.79962	29	1	0.3500	0.0533	0.25964	0.4718
0.82965	28	1	0.3375	0.0529	0.24828	0.4588
0.86133	27	1	0.3250	0.0524	0.23699	0.4457
0.89453	26	1	0.3125	0.0518	0.22578	0.4325
0.92836	25	1	0.3000	0.0512	0.21466	0.4193
0.96391	24	1	0.2875	0.0506	0.20362	0.4059
1.00158	23	1	0.2750	0.0499	0.19267	0.3925
1.04044	22	1	0.2625	0.0492	0.18181	0.3790
1.08146	21	1	0.2500	0.0484	0.17104	0.3654
1.12362	20	1	0.2375	0.0476	0.16038	0.3517
1.16989	19	1	0.2250	0.0467	0.14982	0.3379
1.21807	18	1	0.2125	0.0457	0.13937	0.3240
1.26940	17	1	0.2000	0.0447	0.12903	0.3100
1.32313	16	1	0.1875	0.0436	0.11882	0.2959
1.37833	15	1	0.1750	0.0425	0.10874	0.2816
1.43590	14	1	0.1625	0.0412	0.09881	0.2672
1.50193	13	1	0.1500	0.0399	0.08903	0.2527
1.57455	12	1	0.1375	0.0385	0.07942	0.2380
1.65130	11	1	0.1250	0.0370	0.07000	0.2232
1.73670	10	1	0.1125	0.0353	0.06079	0.2082
1.82759	9	1	0.1000	0.0335	0.05182	0.1930
1.94003	8	1	0.0875	0.0316	0.04312	0.1776
2.06199	7	1	0.0750	0.0294	0.03474	0.1619
2.21092	6	1	0.0625	0.0271	0.02675	0.1460
2.39800	5	1	0.0500	0.0244	0.01924	0.1300
2.61047	4	1	0.0375	0.0212	0.01236	0.1138
2.90772	3	1	0.0250	0.0175	0.00636	0.0982
3.28075	2	1	0.0125	0.0124	0.00178	0.0877
4.02856	1	1	0.0000	NA	NA	NA