# ANALISIS *SURVIVAL* UNTUK DATA TERSENSOR MENGGUNAKAN MODEL DISTRIBUSI LOG-LOGISTIK

Skripsi
Disusun untuk melengkapi syarat-syarat
guna memperoleh gelar Sarjana Sains

DEBI OKTAVIANI
3125110491

PROGRAM STUDI MATEMATIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI JAKARTA
2015

# LEMBAR PERSETUJUAN HASIL SIDANG SKRIPSI

## ANALISIS *SURVIVAL* UNTUK DATA TERSENSOR MENGGUNAKAN MODEL DISTRIBUSI LOG-LOGISTIK

Nama : Debi Oktaviani

No. Registrasi : 3125110491

| | | Nama | Tanda Tangan | Tanggal |
|---|---|---|---|---|
| Penanggung Jawab | | | | |
| Dekan | : | Prof. Dr. Suyono, M.Si. | ............... | ............ |
| | | NIP. 19671218 199303 1 005 | | |
| Wakil Penanggung Jawab | | | | |
| Pembantu Dekan I | : | Dr. Muktiningsih, M.Si. | ............... | ............ |
| | | NIP. 19640511 198903 2 001 | | |
| Ketua | : | Drs. Sudarwanto, M.Si., DEA | ............... | ............ |
| | | NIP. 19650325 199303 1 003 | | |
| Sekretaris | : | Dian Handayani, M.Si | ............... | ............ |
| | | NIP. 19740415 199803 2 001 | | |
| Penguji | : | Prof. Dr. Suyono, M.Si. | ............... | ............ |
| | | NIP. 19671218 199303 1 005 | | |
| Pembimbing I | : | Dra. Widyanti Rahayu, M.Si. | ............... | ............ |
| | | NIP. 19661103 200112 2 001 | | |
| Pembimbing II | : | Vera Maya Santi, M.Si. | ............... | ............ |
| | | NIP. 19790531 200501 2 006 | | |

Dinyatakan lulus ujian skripsi tanggal: 23 Juni 2015

# ABSTRACT

**DEBI OKTAVIANI, 3125110491. Survival Analysis for Censored Data Using Log-Logistic Distribution Model. Thesis. Faculty of Mathematics and Natural Science, Jakarta State University. 2015.**

*Parametric survival data analysis is one of the methods which is frequently used in analyzing censored data. Certain assumption about the population's distribution is needed in order to analyze survival data with parametric censored data. This method is applied to estimate the survival and failure time probability of an individual, which is assumed, following a distribution such as Exponential, Weibull, or Log-Logistic distribution. Most frequently used parametric distribution is the Log-Logistic distribution because of its hazard function that is not monotonously increasing or decreasing and also inconstant. Therefore, in this thesis will discuss about the survival model for censored data by using Log-Logistic distribution. Censored data that is used in this thesis is right censored data.*

***Keywords*** *: parametric survival data analysis model, Log-Logistic distribution, right censored data.*

# ABSTRAK

**DEBI OKTAVIANI, 3125110491. Analisis *Survival* untuk Data Tersensor Menggunakan Model Distribusi Log-Logistik. Skripsi. Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta. 2015.**

Metode statistika yang sering digunakan untuk menganalisis data *survival* tersensor adalah model analisis data *survival* parametrik. Untuk menganalisis data *survival* parametrik dengan data tersensor diperlukan asumsi tertentu mengenai distribusi populasinya. Metode analisis seperti ini biasanya diaplikasikan untuk mengestimasi probabilitas *survival* dan probabilitas waktu kegagalan seorang individu yang diasumsikan mengikuti distribusi tertentu seperti distribusi Eksponensial, Weibull, atau distribusi Log-Logistik. Distribusi parametrik yang sering digunakan adalah distribusi Log-Logistik karena memiliki bentuk fungsi kegagalan (*hazard*) yang tidak monoton naik atau turun dan juga tidak konstan. Oleh karena itu, skripsi ini akan membahas mengenai model *survival* untuk data tersensor menggunakan distribusi Log-Logistik. Data tersensor yang digunakan merupakan data tersensor kanan.

**Kata kunci** : model analisis *survival* parametrik, distribusi Log-Logistik, data tersensor kanan.

# PERSEMBAHANKU...

"... dan janganlah kamu berputus asa dari rahmat Allah. Sesungguhnya tiada berputus asa dari rahmat-Nya melainkan kaum kafir" (Qs. Yusuf :12)

"*Anfang gut ende gut. Something which is started with a good will and a good way, will be ended with a good result.* "

Skripsi ini kupersembahkan untuk Papa, Mama, Bunda, Tasya, Febi, dan Siex.

*"Terima kasih atas dukungan, do'a, serta kasih yang selalu diberikan"*.

# KATA PENGANTAR

Segala Puji syukur kepada Allah SWT atas pengetahuan dan kemampuan sehingga penulis dapat menyelesaikan skripsi yang berjudul "Analisis *Survival* untuk Data Tersensor Menggunakan Distribusi Log-Logistik" yang merupakan salah satu syarat dalam memperoleh gelar Sarjana Jurusan Matematika Universitas Negeri Jakarta.

Terselesaikannya skripsi ini tidak terlepas dari adanya bantuan dari berbagai pihak yang telah memberikan arahan, doa, motivasi, dan sebagainya. Oleh karena itu, dalam kesempatan ini penulis ingin menyampaikan terima kasih terutama kepada:

1. Kedua orang tua penulis yang telah merawat, membesarkan, memberikan dukungan, dan doa kepada penulis agar tetap semangat dan pantang menyerah dalam menjalani semua hal sulit yang dihadapi.

2. Untuk Adik perempuanku Tasya dan Kakakku Febi, Terima kasih atas semangat dan doa yang telah diberikan dalam penyelesaian tugas akhir ini.

3. Dra. Widyanti Rahayu, M.Si. selaku Dosen Pembimbing I dan Vera Maya Santi, M.Si selaku Dosen Pembimbing II. Terima kasih telah meluangkan waktu, saran, nasihat, dan pengarahan kepada penulis dalam pengerjaan skripsi ini sehingga menjadi lebih baik. Terima kasih pula karena ibu tidak hanya menjadi pembimbing penulis melainkan sebagai ibu kedua. Mohon maaf atas segala kekurangan, semoga ibu selalu diberkahi oleh Allah SWT.

4. Drs. Mulyono, M.Kom selaku dosen Pembimbing Akademik. Drs. Makmuri, M.Si selaku Ketua Jurusan Matematika, dan juga Ratna Widyanti,

S.Si, M.Kom selaku Ketua Prodi Matematika. Terima kasih atas segala bantuan dan kerja sama Bapak dan Ibu selama pengerjaan skripsi ini.

5. Seluruh dosen Matematika FMIPA UNJ yang telah memberikan ilmunya kepada penulis dan seluruh karyawan FMIPA UNJ yang telah memberikan kemudahan kepada penulis dalam bentuk apapun.

6. Ka Mudita selaku kakak dan penasihatku. Terima kasih banyak atas waktu, semangat, doa, dan bantuan kepada penulis.

7. Sahabat terbaikku CTT. Anti, Firdha, Gia, Pephu dan yang paling utama Danti dan Lina. Terima kasih banyak untuk suka duka yang kita hadapi selama ini.

8. Teman Matematika UNJ 2011. Terutama Puti, Teddy, Depe, Ambar, Dytta, Cika, Syifa, Vira, Idam, Iyus, Hamas, Indah Dwi, Mahdhi, dll. Terima kasih atas perjuangannya selama ini.

9. Sahabatku Nina, teman dari kecilku Putri, dan salah satu teman terbaik Cut Devi. Terima kasih atas dukungan dan doa yang kalian berikan selama ini.

10. Untuk kakakku Ka Rizka, ka Dina, Ka Jefri, dan Ka Joy. Terima kasih atas dukungannya.

11. Dini Adlina, Atinda, Dini Rohmah, Yuda, Heru, Febi. Terima kasih atas segala canda tawa.

12. Bas dan Shan, Terima kasih atas bantuan dan doa yang diberikan.

13. Murid yang sudah seperti adik-adikku tersayang; Nadya, Chairany, Arum, Ifta, Debra, Indira, Mirah, Caca, Fia, Daffa, dan Salma. Terima kasih telah memberi semangat dan senyuman setiap hari.

14. dan yang terakhir, untuk Audi Indrayahya. Terima kasih atas segala dukungan, doa, dan semangat yang diberikan selama ini.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Masukan dan kritikan akan sangat berarti. Semoga skripsi ini dapat bermanfaat bagi pembaca sekalian.

Jakarta, Juli 2015

Debi Oktaviani

# DAFTAR ISI

**ABSTRACT** **i**

**ABSTRAK** **ii**

**KATA PENGANTAR** **iv**

**DAFTAR ISI** **viii**

**DAFTAR TABEL** **ix**

**DAFTAR GAMBAR** **xi**

**I PENDAHULUAN** **1**

1.1 Latar Belakang Masalah . . . . . . . . . . . . . . . . . . . . . . . . 1

1.2 Perumusan Masalah . . . . . . . . . . . . . . . . . . . . . . . . . 3

1.3 Pembatasan Masalah . . . . . . . . . . . . . . . . . . . . . . . . 3

1.4 Tujuan Penulisan . . . . . . . . . . . . . . . . . . . . . . . . . . 3

1.5 Manfaat Penulisan . . . . . . . . . . . . . . . . . . . . . . . . . . 4

1.6 Metode Penelitian . . . . . . . . . . . . . . . . . . . . . . . . . . 4

**II LANDASAN TEORI** **5**

2.1 Data *Survival* . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5

2.1.1 Data *Survival* Tersensor . . . . . . . . . . . . . . . . . . . . 6

2.2 Fungsi *Survival* . . . . . . . . . . . . . . . . . . . . . . . . . . . 10

2.3 Fungsi *Hazard* . . . . . . . . . . . . . . . . . . . . . . . . . . . . 13

2.4 Distribusi Log-Logistik . . . . . . . . . . . . . . . . . . . . . . . 18

2.4.1 Fungsi Densitas Log-Logistik . . . . . . . . . . . . . . . . . . 18
2.4.2 Fungsi *Survival* Log-Logistik . . . . . . . . . . . . . . . . . . 21
2.4.3 Fungsi *Hazard* Log-Logistik . . . . . . . . . . . . . . . . . . 22
2.5 Statistika Terurut . . . . . . . . . . . . . . . . . . . . . . . . . 24
2.6 Distribusi Multinomial . . . . . . . . . . . . . . . . . . . . . . . 25
2.7 Uji Statistik Anderson-Darling . . . . . . . . . . . . . . . . . . . 27
2.8 Matriks Informasi . . . . . . . . . . . . . . . . . . . . . . . . . 28
2.9 Interval Konfidensi . . . . . . . . . . . . . . . . . . . . . . . . . 29

**III PEMBAHASAN** **31**

3.1 Model *Survival* Data Tersensor Kanan . . . . . . . . . . . . . . . . 31
3.2 Estimasi Paramater dengan Maksimum Likelihood Estimator *(MLE)* 34
3.2.1 Fungsi Likelihood dari Distribusi Log-Logistik . . . . . . . . 34
3.2.2 Maksimum Likelihood Estimator *(MLE)* . . . . . . . . . . . 34
3.3 Interval Konfidensi . . . . . . . . . . . . . . . . . . . . . . . . . 40
3.4 Diagram Alir Data Tersensor Kanan Menggunakan Distribusi Log-Logistik . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 44
3.5 Studi Kasus . . . . . . . . . . . . . . . . . . . . . . . . . . . . 45

**IV PENUTUP** **56**

4.1 Kesimpulan . . . . . . . . . . . . . . . . . . . . . . . . . . . . 56
4.2 Saran . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 57

**DAFTAR PUSTAKA** **58**

**LAMPIRAN-LAMPIRAN** **60**

# DAFTAR TABEL

3.1 Tabel Probabilitas Pasien Operasi Jantung . . . . . . . . . . . . . . 55

4.1 DATA KEMATIAN PASIEN SETELAH DILAKUKAN OPERASI JANTUNG DI SALAH SATU RUMAH SAKIT DI JAKARTA PADA TAHUN 2014 . . . . . . . . . . . . . . . . . . . . . . . . . . . . 60

# DAFTAR GAMBAR


2.1 Pengamatan *Survival* Terhadap 3 Individu . . . . . . . . . . . . . . 5

2.2 Individu Tersensor Kiri . . . . . . . . . . . . . . . . . . . . . . . . . 7

2.3 Individu Tersensor Kanan . . . . . . . . . . . . . . . . . . . . . . . 8

2.4 Individu Tersensor Interval . . . . . . . . . . . . . . . . . . . . . . . 8

2.5 Pengamatan *Survival* Terhadap 6 Individu . . . . . . . . . . . . . . 9

2.6 Kurva Fungsi *Survival* Secara Teori . . . . . . . . . . . . . . . . . . 12

2.7 Kurva Fungsi *Survival* Secara Praktik . . . . . . . . . . . . . . . . . 13

2.8 Kurva Fungsi *Hazard* (a) konstan, (b) naik, (c) turun, (d) naik dan turun . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17

2.9 Kurva fungsi densitas peluang dari distribusi Log-Logistik . . . . 19

2.10 Kurva fungsi *hazard* kumulatif Log-Logistik . . . . . . . . . . . . 24

3.1 Ilustrasi model tersensor kanan . . . . . . . . . . . . . . . . . . . . . 31

3.2 Ilustrasi model tersensor kanan . . . . . . . . . . . . . . . . . . . . . 32

3.3 Iterasi Newton Raphson . . . . . . . . . . . . . . . . . . . . . . . . . 39

3.4 Diagram Alir Data Tersensor Kanan Menggunakan Distribusi Log-Logistik . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 44

3.5 Kurva Anderson-Darling Data Kematian Pasien Setelah Operasi Jantung . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 46

3.6 Kurva Fungsi *Survival* dari Data Kematian Pasien Setelah Operasi Jantung . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 51

3.7 Kurva Fungsi *Hazard* dari Data Kematian Pasien Setelah Operasi Jantung . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 52

3.8 Kurva Fungsi *Hazard* Kumulatif dari Data Kematian Pasien Setelah Operasi Jantung . . . . . . . . . . . . . . . . . . . . . . . . . 54

# BAB I

# PENDAHULUAN

## 1.1 Latar Belakang Masalah

Analisis *survival (Survival Analysis)* merupakan suatu metode statistika yang berguna untuk menganalisis data mengenai waktu terjadinya suatu kejadian *(time to event)* atau untuk menganalisis data tentang waktu daya tahan *life time*. Salah satu tujuan dari analisis *survival* adalah untuk menaksir dan mengintepretasikan probabilitas kelangsungan hidup dari data *survival* yang didefinisikan bahwa seorang individu dapat bertahan hidup melebihi suatu waktu tertentu. Analisis *survival* ini banyak digunakan dalam berbagai bidang misalnya pada bidang kesehatan untuk menentukan berapa lama suatu individu dapat bertahan hidup dan dapat digunakan pada bidang-bidang lainnya.

Data *survival* dalam analisis semacam ini melibatkan bentuk data *time to event*, misalnya data waktu sampai terjadinya kematian. Sehimpunan data yang digunakan biasanya berupa data tersensor. Data dikatakan tersensor apabila hanya diketahui beberapa informasi mengenai waktu sampai terjadinya suatu kejadian *(event)* pada individu yang bersangkutan tetapi tidak diketahui waktu kejadiannya secara pasti (Collect, 2003).

Model statistika yang sering digunakan untuk menganalisis data *survival* tersensor adalah model analisis data *survival* parametrik. Untuk menganalisis data *survival* parametrik dengan data tersensor diperlukan asumsi tertentu mengenai distribusi populasinya. Distribusi parametrik yang populer untuk menganalisis data *survival* adalah Distribusi Weibull. Distribusi ini sering digunakan untuk memodelkan data *survival* namun penggunaannya jarang digunakan karena pada kenyataannya fungsi *hazard* pada distribusi ini selalu monoton naik atau turun. Oleh karena itu, beberapa peneliti lebih sering menggunakan distribusi Log−Logistik pada kasus yang tidak mengikuti distribusi Weibull, misalnya pada kasus operasi jantung. Pada pasien yang telah mengalami operasi jantung beberapa hari pertama setelah dilakukan operasi tingkat kegagalan akan tinggi sampai mencapai puncak dan setelah tubuh mengalami penyesuaian selama beberapa waktu maka tingkat kegagalanpun akan menurun. Dalam kasus ini distribusi mengikuti distribusi Log-Logistik karena mempunyai bentuk fungsi kegagalan *(hazard)* yang tidak monoton naik atau turun dan juga tidak konstan (Klein & Kleinbaum, 2005). Selain itu, dengan menggunakan distribusi Log-Logistik para peneliti dapat mengetahui ketahanan hidup seorang individu dalam kurun waktu tertentu.

Beberapa penelitian sebelumnya telah mencoba menerapkan analisis data tersensor pada kasus yang berbeda dengan menggunakan model distribusi yang berbeda pula. Annahar (2011) menerapkan analisis data *survival* tersensor kanan dengan menggunakan *hazard* proposional untuk mengetahui resiko kematian pada penderita penyakit kanker tenggorokan. Sementara itu, Kirana (2013) menerapkan analisis data *survival* pada data tersensor dengan menggunakan perluasan model regresi cox *hazard* non−proposional untuk mengetahui ketahanan hidup pecandu heroin. Dalam tulisan ini akan dibahas mengenai model *survival* untuk data tersensor berdasarkan model distribusi Log−Logistik dan bagaimana cara mengestimasi paramaternya serta contoh penerapan model distribusi Log−Logistik.

## 1.2 Perumusan Masalah

Perumusan masalah yang akan dikaji adalah sebagai berikut:

1. Bagaimana model *survival* untuk data tersensor kanan berdasarkan model distribusi Log−logistik?

2. Bagaimana cara mengestimasi parameter model *survival* untuk data tersensor kanan berdasarkan model distribusi Log−logistik?

3. Bagaimana penerapan model *survival* untuk data tersensor kanan berdasarkan model distribusi Log−logistik?

## 1.3 Pembatasan Masalah

Pembatasan masalah dalam penulisan ini adalah data *survival* yang diamati merupakan data tersensor kanan *(right censored data)*.

## 1.4 Tujuan Penulisan

Tujuan dari penulisan ini diantaranya adalah :

1. Mendapatkan model *survival* untuk data tersensor kanan berdasarkan model distribusi Log−logistik.

2. Mendapatkan estimasi parameter model *survival* untuk data tersensor kanan berdasarkan model distribusi log−Logistik.

3. Menerapkan model *survival* untuk data tersensor kanan berdasarkan model distribusi Log−logistik

## 1.5 Manfaat Penulisan

Manfaat yang diharapkan dari tulisan ini adalah dapat memberikan pengetahuan kepada pembaca mengenai analisis *survival* untuk data tersensor kanan menggunakan model distribusi Log−logistik.

## 1.6 Metode Penelitian

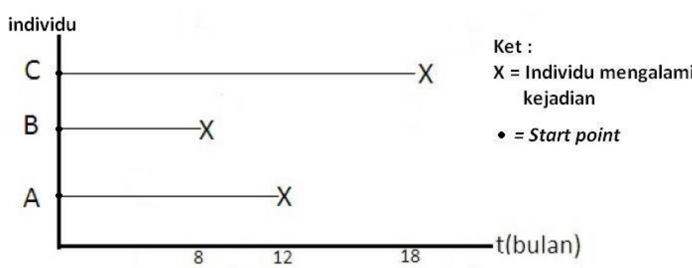
Skripsi ini merupakan kajian teori tentang analisis *survival* untuk data tersensor menggunakan model distribusi Log-logistik yang didasarkan pada buku-buku dan jurnal-jurnal tentang analisis *survival* untuk data tersensor menggunakan model distribusi Log-logistik. Referensi utama yang digunakan Steve Bennet (1983).

# BAB II

# LANDASAN TEORI

## 2.1 Data *Survival*

Dalam analisis *survival*, waktu ketahanan hidup suatu objek sering disebut *survival time* yang merupakan suatu objek dapat bertahan selama jangka waktu pengamatan. Peristiwa yang diamati dalam analisis survival diantaranya adalah kematian, penyakit, dan peristiwa-peristiwa negatif lainnya. Data *survival* adalah data tentang pengamatan jangka waktu dari awal pengamatan sampai terjadinya peristiwa homogen tertentu. Peristiwa itu dapat berupa kegagalan, kematian, dan lain-lain (Lee, 2003).




Gambar 2.1: Pengamatan *Survival* Terhadap 3 Individu

### 2.1.1 Data *Survival* Tersensor

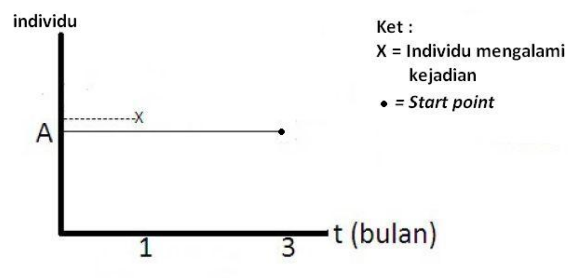
Perbedaan antara analisis *survival* dengan analasis statistik lainnya adalah adanya data tersensor. Data dikatakan tersensor jika pengamatan waktu bertahan hanya sebagian dan tidak sampai waktu kegagalannya *(failure time)* (Pyke & Thompson, 1986). Penyebab terjadinya data tersensor diantaranya adalah ;

- *Loss to follow up*, terjadi bila objek berpindah, meninggal atau menolak untuk berpartisipasi.
- *Drop out*, terjadi apabila perlakuan dihentikan karena alasan tertentu.
- *Termination*, terjadi apabila masa penelitian berakhir sementara objek yang diobservasi belum mencapai kegagalan *(failure time)*.

Berikut adalah beberapa jenis data *survival* tersensor antara lain;

1. Data *Survival* Tersensor Kiri *(left censored)*

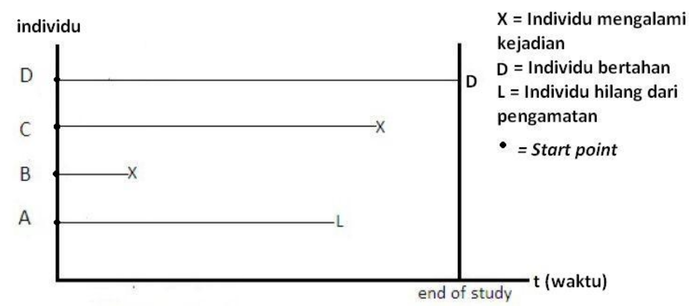
   Data dikatakan tersensor kiri *(left censored)* apabila waktu *survival* sesungguhnya kurang dari waktu *survival* pada saat pengamatan. Sebagai contoh yaitu pengamatan terhadap beberapa pasien yang menjalani operasi pengangkatan tumor. Pada pengamatan ini peristiwa kegagalan yang diamati ialah tumbuh kembalinya tumor tesebut pada tubuh pasien. Empat bulan kemudian setelah dilakukan operasi, dilakukan pemeriksaan terhadap pasien penderita tumor. Didapati tumor tumbuh kembali pada sebagian pasien. Untuk beberapa pasien tertentu, waktu sesungguhnya tumor tumbuh kembali adalah kurang dari empat bulan. Pada kejadian ini, waktu survival tersensor kiri *(left censored)* terjadi ketika kegagalan yang diamati pada individu telah terjadi sebelum *start point*.




Gambar 2.2: Individu Tersensor Kiri

2. Data *Survival* Tersensor Kanan *(right censored)*

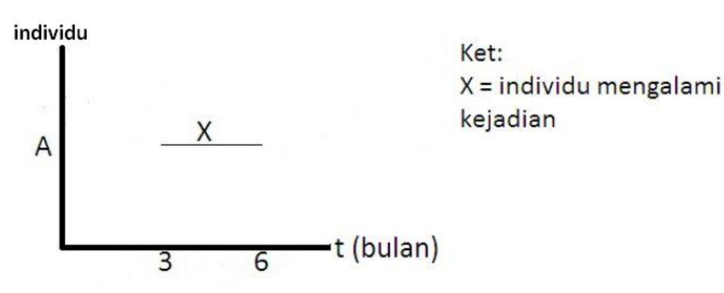
   Data dikatakan tersensor kanan *(right censored)* apabila individu tersebut hilang dari pengamatan *(loss to follow up)* atau individu tersebut masih bertahan sampai waktu pengamatan berakhir. Menurut Lawless (1982) data tersensor merupakan suatu data waktu kematian atau waktu tahan hidup yang hanya terdapat r buah observasi dalam sampel random yang berukuran n dengan $1 \leq r \leq n$. Waktu survival tersensor kanan *(right censored)* individu kurang dari waktu *survival* sesungguhnya individu tersebut, namun waktu *survival* sesungguhnya tidak diketahui. Dalam suatu penelitian, data tersensor kanan lebih sering digunakan dibandingkan data tersensor yang lainnya (Collect, 2004).




Gambar 2.3: Individu Tersensor Kanan

3. Data *Survival* Tersensor Interval

   Tipe lain dari sensor adalah sensor interval. Dalam hal ini, individu mengalami kegagalan pada selang waktu tertentu. Menggunakan contoh yang telah ada sebelumnya, jika pasien telah bebas dari tumor pada saat pemeriksaan di bulan ketiga namun, 3 bulan berikutnya pada saat dilakukan pemeriksaan didapati tumor tersebut kembali tumbuh pada tubuh pasien. Waktu survival sesungguhnya pasien tersebut berada di antara 3 bulan dan 6 bulan. Pada keadaan seperti inilah data survival suatu individu dikatakan tersensor interval.

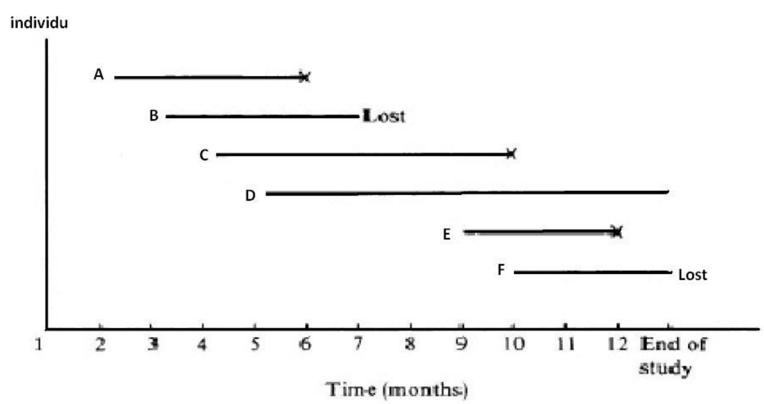



Gambar 2.4: Individu Tersensor Interval

4. *Progressive Censoring*

Dalam beberapa pengamatan, individu yang diamati masuk ke dalam pengamatan dalam waktu yang berbeda. Setelah individu masuk ke dalam pengamatan, mereka diamati sampai mengalami kegagalan atau sampai waktu pengamatan berakhir. Beberapa individu ada yang mengalami kegagalan sehingga waktu *survival*-nya diketahui. Sementara itu, beberapa individu hilang dari pengamatan atau bertahan sampai pengamatan berakhir. Misal terdapat 6 individu yang menderita leukimia masuk kedalam pengamatan. Pengamatan berlangsung selama satu tahun. Pasien-pasien tersebut diberikan perawatan terhadap penyakit leukimia. Peristiwa yang diamati adalah memburuknya keadaan pasien setelah diberikan perawatan.

Berdasarkan Gambar 2.5, waktu *survival* 6 individu tersebut secara berurut adalah sebagai berikut :




Gambar 2.5: Pengamatan *Survival* Terhadap 6 Individu

$4, 4^+, 6, 8, 3,$ dan $3^+$ bulan. Pada individu B dan F yang hilang dari pengamatan *(lost)* maka waktu *survival* akan ditandai dengan tanda + yang menandakan tersensor. Misalkan terdapat n individu yang diamati, r adalah banyaknya individu yang mengalami kegagalan sebelum atau pada waktu T dan ada sebanyak (n - r) individu yang belum mengalami kegagalan pada waktu T. Penyusunan r amatan yang tidak tersensor berdasarkan besarnya adalah

$$t_1 \leq t_2 \leq ... \leq t_r, t^+_{r+1}, ..., t^+_n \tag{2.1}$$

## 2.2 Fungsi *Survival*

Misalkan T merupakan variabel random non-negatif dari waktu tahan hidup suatu individu dalam interval $[0, \infty)$. Fungsi kepadatan peluang *(probability density function)* $f(t)$ adalah peluang terjadinya suatu kejadian *(event)* seperti kegagalan dalam interval waktu dari t sampai $(t + \Delta t)$ dan dinyatakan dengan

$$f(t) = lim_{\Delta t \to 0} \left[ \frac{P(t \leq T < (t + \Delta t))}{\Delta t} \right] \tag{2.2}$$

maka fungsi distribusi kumulatif waktu hidup suatu individu dalam interval $[0, \infty)$ didefinisikan sebagai

$$\begin{aligned} F(t) &= P(T \leq t), t \geq 0 \\ &= \int_0^t f(u)du \end{aligned}$$

dan F(t) merupakan peluang waktu hidup suatu individu kurang dari sama dengan waktu t.

**Definisi 2.2.1.** Misalkan T merupakan waktu hingga suatu kegagalan terjadi dimana $T \geq 0$. Fungsi *survival* adalah peluang seseorang dapat bertahan hingga atau lebih dari waktu t (mengalami kejadian sesudah t)

$$\begin{aligned} S(t) &= P(T > t) \\ &= 1 - P(T \leq t) \\ &= 1 - F(t) \end{aligned} \tag{2.3}$$

fungsi *survival* merupakan fungsi tidak naik dengan $S(0) = 1$ dan $lim_{t \to \infty} S(t) = 0$ maka fungsi F(t) dan S(t) memenuhi hubungan

$$F(t) + S(t) = 1$$

Berdasarkan fungsi distribusi kumulatif, $f(t)$ dapat dinyatakan sebagai turunan dari $F(t)$ terhadap t. Maka dapat diperoleh hubungan antara $f(t)$ dan $S(t)$ sebagai berikut:

$$\begin{aligned} f(t) &= \frac{d}{dt} F(t) \\ &= \frac{d}{dt}(1 - S(t)) \\ &= -\frac{d}{dt} S(t) \\ &= -S'(t) \end{aligned} \tag{2.4}$$

Berikut adalah kurva fungsi *survival* secara teoritis dan fungsi *survival* dalam praktiknya :

1. Fungi *survival* secara teoritis

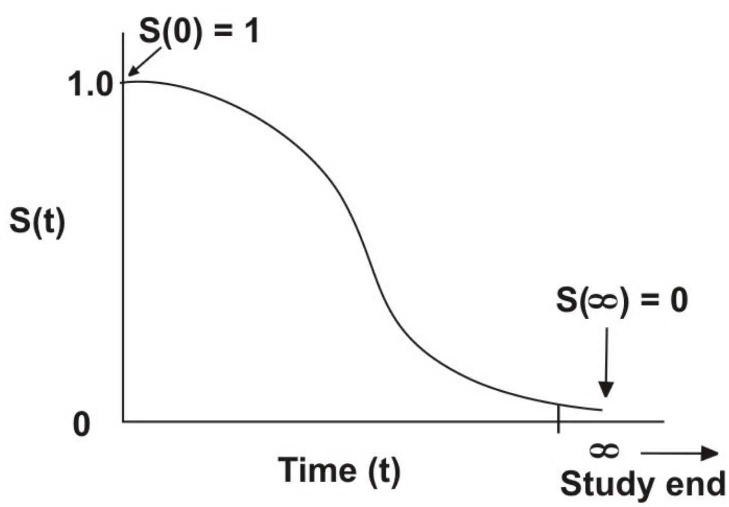
   Berdasarkan Gambar 2.6 jika periode pengamatan bertambah tanpa batas, pada akhir studi tidak seorangpun yang akan bertahan hidup sehingga kurva *survival* harus menuju nol.

Secara teori, fungsi *survival* dapat diplot sebagai kurva *survival* yang menggambarkan peluang ketahanan hidup individu pada titik-titik waktu t antara 0 sampai $\infty$. Semua fungsi *survival* memiliki karakteristik seperti berikut :

- merupakan fungsi monoton tak naik (kurva turun) dengan t semakin naik
- pada saat $t = 0$, $(S(t) = S(0) = 1)$

  artinya pada awal pengamatan karena belum ada individu yang telah mengalami suatu kejadian *(event)* maka probabilitas *survival* pada saat itu adalah 1.
- pada saat $t \to \infty$, $S(t) = S(\infty) = 0$

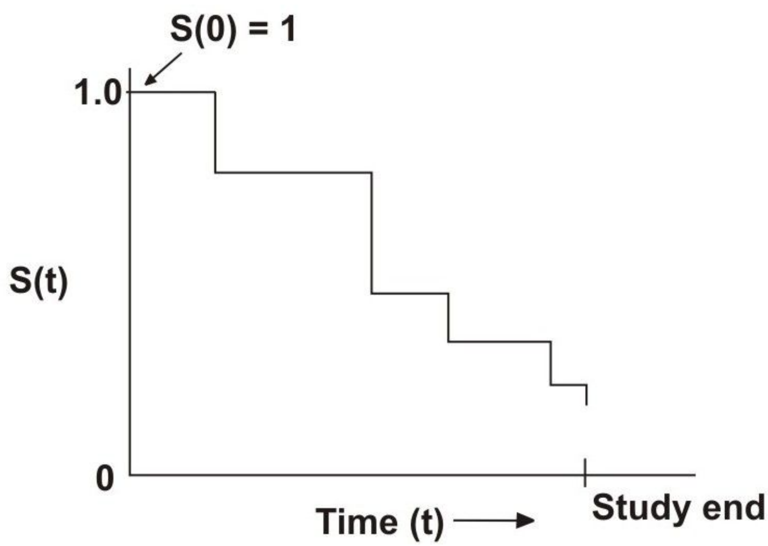
  artinya secara teori jika periode suatu pengamatan bertambah tanpa batas, pada akhirnya tidak ada seorangpun yang akan bertahan hidup sehingga kurva *survival* harus menuju 0.




Gambar 2.6: Kurva Fungsi *Survival* Secara Teori

2. Fungsi *survival* dalam praktiknya

   Berdasarkan Gambar 2.7 ketika digunakan data yang nyata (realitas), akan diperoleh grafik *survival* yang merupakan fungsi tangga. Sementara itu, lamanya periode suatu pengamatan tidak mungkin sampai menuju tak berhingga, mungkin tidak setiap individu yang diamati mengalami suatu kejadian *(event)* sehingga tidak semua fungsi *survival* (yang diestimasi) akan sama dengan 0 pada akhir masa pengamatan.




Gambar 2.7: Kurva Fungsi *Survival* Secara Praktik

## 2.3 Fungsi *Hazard*

Fungsi *hazard* adalah peluang seseorang mengalami suatu resiko atau kejadian *(event)* seperti kegagalan pada waktu t dengan syarat bahwa seseorang tersebut telah bertahan hingga waktu t dalam interval $t + \Delta t$ (Lawless, 2003).

**Definisi 2.3.1.** Misalkan T merupakan waktu hingga suatu kegagalan terjadi dimana $T \geq 0$. Maka fungsi hazard diberikan

$$h(t) = lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \tag{2.5}$$

Berdasarkan persamaan (2.5) $h(t)$ diartikan sebagai limit dan interval waktu $\Delta t$ menuju nol dengan perbandingan dua kuantitas sebagai berikut :

1. Peluang bahwa kejadian akan disebabkan antara waktu t dan $t + \Delta t$ dengan waktu *survival* $T \geq t$

2. Waktu dalam interval $\Delta t$

Fungsi $h(t)$ disebut juga *hazard rate* atau laju kegagalan. Fungsi *hazard* dapat diintepretasikan sebagai tingkat kecepatan suatu individu mengalami suatu kegagalan per satuan waktu. Berdasarkan definisi (2.5) maka diperoleh hubungan antara fungsi *survival* dan fungsi *hazard* sebagai berikut :

$$\begin{aligned} h(t) &= lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \\ &= lim_{\Delta t \to 0} \frac{P(t \leq T < t | \Delta t)}{P(T \geq t)} \left( \frac{1}{\Delta t} \right) \\ &= lim_{\Delta t \to 0} \frac{P(T < t + \Delta t) - P(T < t)}{P(T \geq t)} \left( \frac{1}{\Delta t} \right) \\ &= lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} \left( \frac{1}{P(T \geq t)} \right) \\ &= lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} \left( \frac{1}{S(t)} \right) \\ &= \frac{f(t)}{S(t)} \end{aligned} \tag{2.6}$$

sehingga

$$\begin{aligned} h(t) &= \frac{f(t)}{S(t)} \\ &= \frac{\frac{d}{dt}F(t)}{S(t)} \\ &= \frac{\frac{d}{dt}(1 - S(t))}{S(t)} \\ &= -\frac{S'(t)}{S(t)} \end{aligned} \tag{2.7}$$

Berdasarkan persamaan (2.4) dan (2.6) maka diperoleh

$$\begin{aligned} h(t) &= \frac{f(t)}{S(t)} \\ &= -\frac{S'(t)}{S(t)} \\ &= -\frac{d(\log S(t))}{dt} \end{aligned} \tag{2.8}$$

dengan mengintegralkan kedua sisi pada persamaan (2.8) akan memberikan fungsi *survival* yang dinyatakan dalam fungsi *hazard*

$$\begin{aligned} h(t) &= -\frac{d(\log S(t))}{dt} \\ \int_0^t h(u)du &= \int_0^t -\frac{d(\log S(t))du}{du} \\ -\int_0^t h(u)du &= -\log S(u) \mid_0^t \\ - \mid_0^t h(u)du &= \log S(t) \\ S(t) &= \exp\left[-\int_0^t h(u)du\right] \end{aligned} \tag{2.9}$$

dengan menggunakan persamaan (2.9) dapat diperoleh fungsi *hazard* kumulatifnya sebagai berikut :

$$\begin{aligned} S(t) &= \exp\left[-\int_0^t h(u)du\right] \\ S(t) &= \exp\left[-H(t)\right] \\ logS(t) &= -H(t) \\ H(t) &= -\log S(t) \end{aligned} \tag{2.10}$$

Seperti fungsi *survival*, fungsi *hazard* juga dapat diplot sebagai kurva fungsi *hazard* terhadap terhadap nilai t namun berbeda dengan fungsi *survival*, kurva h(t) tidak harus dimulai dari 1 dan bergerak ke bawah menuju 0, tetapi kurva $h(t)$ dapat dimulai dari nilai berapapun ($h(t) \geq 0$) dan bergerak ke atas dan ke bawah terhadap waktu t. Dengan kata lain, untuk suatu nilai tertentu t, fungsi *hazard* $h(t)$ mempunyai karakteristik sebagai berikut :

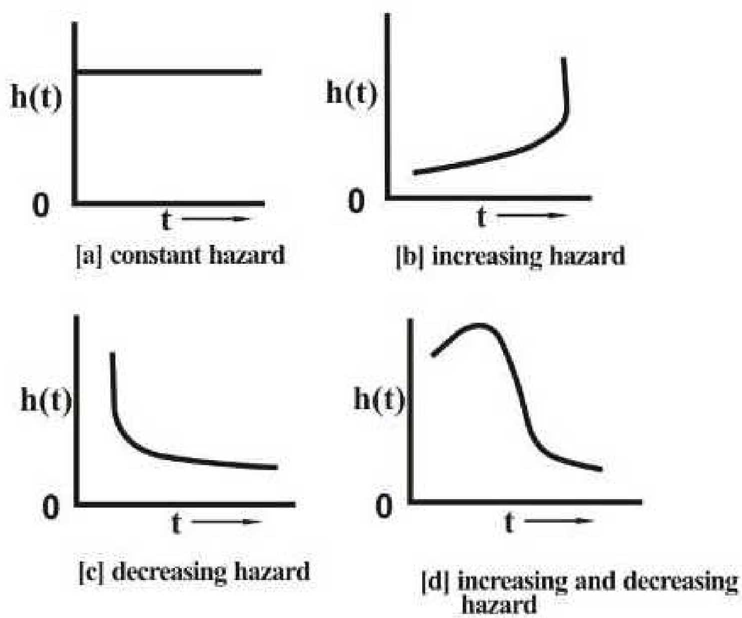
1. Selalu bernilai non negatif, $h(t) \geq 0$.
2. Tidak memiliki nilai batas.

Pada Gambar 2.8 akan ditunjukkan beberapa bentuk kurva fungsi *hazard* dan berikut adalah penjelasan singkat mengenai bentuk fungsi *hazard*

1. Untuk kurva konstan *(constant hazard)*, biasanya fungsi mengikuti model distribusi eksponensial dimana fungsi seperti ini jarang ditemui dalam kehidupan nyata.
2. Untuk kurva *hazard* turun *(decreasing hazard)*, biasanya fungsi mengikuti model distribusi Weibull misalnya saja tingkat kematian sesaat pada populasi bayi yang baru lahir. Cenderung tinggi pada awal setelah kelahiran dan seiring berjalannya waktu tingkat kematian pun akan semakin turun dan stabil.

3. Untuk kurva *hazard* naik *(increasing hazard)*, biasanya fungsi mengikuti model distribusi Weibull misalnya tingkat kematian sesaat pada penderita kanker. Pada awal terdeteksi, tingkat *hazard* masih rendah dan semakin lama akan semakin tinggi tingkat kematian pada penderita kanker tersebut.

4. Untuk kurva *hazard* naik turun *(increasing and decreasing hazard)*, biasanya fungsi mengikuti model distribusi Log-Logistik misalnya tingkat kematian pada individu setelah dilakukan operasi jantung. Kemudian pada resiko awalnya dapat terjadi infeksi atau komplikasi lain sesaat setelah prosedur operasi berjalan, kemudian resikonya berkurang seiring dengan kesembuhan individu.




Gambar 2.8: Kurva Fungsi *Hazard* (a) konstan, (b) naik, (c) turun, (d) naik dan turun

Pada penelitian ini akan digunakan fungsi *hazard* naik turun sesuai dengan ciri khas dari distribusi Log-Logistik.

## 2.4 Distribusi Log-Logistik

Distribusi Log-Logistik merupakan salah satu distribusi peluang kontinu untuk variabel random non-negatif dalam ilmu statistika. Distribusi ini biasa digunakan dalam analisis tahan hidup seperti untuk meneliti waktu penyembuhan suatu penyakit sebagai model parametrik karena sangat mirip dengan distribusi normal.

Bentuk distribusi Log-Logistik sangat mirip dengan distribusi Log-Normal tetapi distribusi Log-Logistik lebih cocok digunakan untuk analisis *survival* daripada distribusi Log-Normal ketika menggunakan data *survival* tersensor karena distribusi Log-Normal tidak dapat digunakan secara langsung pada data tersensor. Untuk fungsi *hazard* yang naik atau turun para peneliti biasanya menggunakan distribusi Weibull sementara itu untuk fungsi *hazard* yang naik turun lebih sering menggunakan distribusi Log-logistik dibanding distribusi Log-Normal dan distribusi Log-Logistik adalah alternatif yang baik untuk digunakan dalam model parametrik dibanding distribusi lainnya.

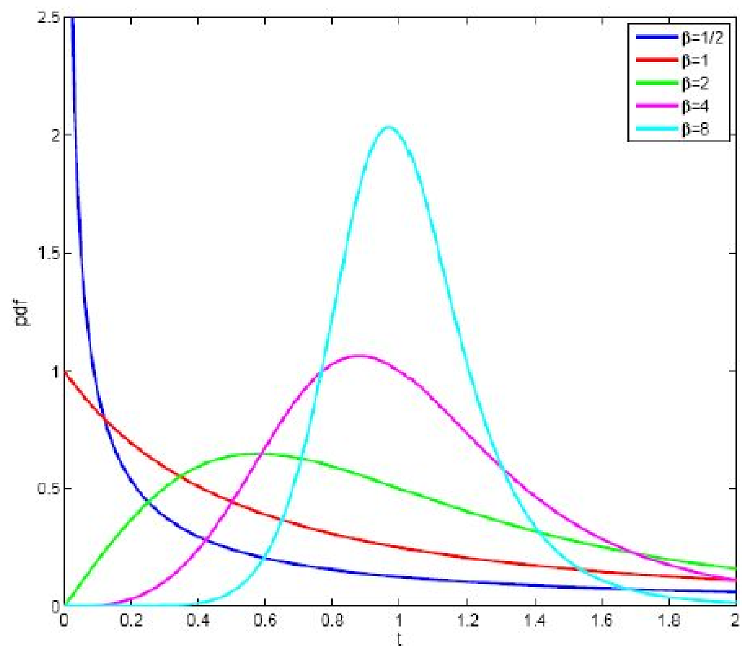
### 2.4.1 Fungsi Densitas Log-Logistik

**Definisi 2.4.1.** Misalkan T merupakan variabel random yang mengikuti distribusi Log-Logistik dengan paramater $\gamma$ dan paramater *shape* $\beta$, maka fungsi densitas dapat didefinisikan sebagai berikut

$$f(t;\gamma;\beta) = \frac{\left[\frac{\beta}{\gamma}\right]\left[\frac{t}{\gamma}\right]^{\beta-1}}{\left[1+\left[\frac{t}{\gamma}\right]^{\beta}\right]^{2}}, t > 0 \tag{2.11}$$

dimana $\gamma > 0$ dan $\beta > 0$, untuk selanjutnya dinotasikan sebagai $T \sim L(\gamma, \beta)$.

Nilai paramater *shape* yaitu $\beta$ yang menyatakan suatu bentuk yang bermacam-macam dari kurva fungsi densitas yakni naik, turun, atau mendatar sehingga kondisi ini sangat cocok digunakan untuk berbagai model data *survival*. Fungsi densitas peluang dari distribusi Log-Logistik dengan $\gamma = 1$ ditunjukkan




Gambar 2.9: Kurva fungsi densitas peluang dari distribusi Log-Logistik

pada Gambar 2.9 diatas untuk nilai $\beta$ yang berbeda. Untuk $0 \leq \beta \leq 1$ fungsi densitas peluangnya menurun, sedangkan untuk $\beta > 1$ fungsi densitas peluangnya merupakan fungsi naik dengan sebuah puncak. Semakin besar nilai $\beta$, puncak dari kurva fungsi densitas peluangnya semakin runcing dan bentuknya semakin simetris.

Maka berdasarkan definisi (2.11) Fungsi distribusi kumulatifnya adalah sebagai berikut

$$F(t) = \int_0^t \frac{\left[\frac{\beta}{\gamma}\right]\left[\frac{y}{\gamma}\right]^{\beta-1}}{\left[1+\left[\frac{y}{\gamma}\right]^{\beta}\right]^2} dy \tag{2.12}$$

misal :

$$\begin{aligned} u &= 1 + \left[\frac{y}{\gamma}\right]^{\beta} \\ du &= \left[\frac{\beta}{\gamma}\right]\left[\frac{y}{\gamma}\right]^{\beta-1} dy \\ dy &= \frac{du}{\left[\frac{\beta}{\gamma}\right]\left[\frac{y}{\gamma}\right]^{\beta-1}} \end{aligned} \tag{2.13}$$

dan

$$\begin{aligned} y = 0 &\rightarrow u = 1 \\ y = t &\rightarrow u = 1 + \left(\frac{t}{\gamma}\right)^{\beta} \end{aligned} \tag{2.14}$$

maka dengan mensubtitusi persamaan (2.13) dan (2.14) diperoleh

$$\begin{aligned} F(t) &= \int_0^t \frac{\left[\frac{\beta}{\gamma}\right]\left[\frac{y}{\gamma}\right]^{\beta-1}}{\left[1+\left[\frac{y}{\gamma}\right]\right]^2} dy \\ &= \int_0^t \frac{\left[\frac{\beta}{\gamma}\right]\left[\frac{y}{\gamma}\right]^{\beta-1}}{u^2} \frac{du}{\left[\frac{\beta}{\gamma}\right]\left[\frac{y}{\gamma}\right]^{\beta-1}} \end{aligned}$$

$$
\begin{aligned}
&= \int_1^{1+\left(\frac{t}{\gamma}\right)^\beta} u^{-2} du \\
&= -u^{-1} \Big|_1^{1+\left(\frac{t}{\gamma}\right)^\beta} \\
&= -\frac{1}{1+\left[\frac{y}{\gamma}\right]^\beta} \Big|_0^t \\
&= -\frac{1}{1+\left[\frac{y}{\gamma}\right]^\beta} + 1 \\
&= 1 - \frac{1}{1+\left[\frac{t}{\gamma}\right]^\beta} \qquad (2.15) \\
&= \frac{1+\left[\frac{t}{\gamma}\right]^\beta - 1}{1+\left[\frac{t}{\gamma}\right]^\beta} \\
&= \frac{\left[\frac{t}{\gamma}\right]^\beta}{1+\left[\frac{t}{\gamma}\right]^\beta} \\
F(t) &= \frac{t^\beta}{\gamma^\beta + t^\beta}
\end{aligned}
$$

dimana $\gamma > 0$ dan $\beta > 0$.

### 2.4.2 Fungsi *Survival* Log-Logistik

**Definisi 2.4.2.** Misalkan T merupakan variabel random dari waktu tahan hidup dari suatu individu. Fungsi *survival* dari $T \sim L_L(\gamma, \beta)$ didefinisikan sebagai peluang suatu individu bertahan sampai waktu t. Maka berdasarkan persamaan

(2.15) fungsi *survival* didefinisikan sebagai berikut:

$$\begin{aligned} S(t) &= 1 - F(t) \\ &= 1 - \frac{t^{\beta}}{\gamma^{\beta} + t^{\beta}} \\ &= \frac{\gamma^{\beta}}{\gamma^{\beta} + t^{\beta}} \\ &= \frac{1}{1 + \left[\frac{t}{\gamma}\right]^{\beta}} \end{aligned} \tag{2.16}$$

### 2.4.3 Fungsi *Hazard* Log-Logistik

**Definisi 2.4.3.** Misalkan T merupakan variabel random dari waktu tahan hidup suatu individu. Fungsi *hazard* $h(t)$ menyatakan peluang suatu individu mengalami kegagalan pada waktu t. Maka berdasarkan persamaan (2.11) dan (2.16) fungsi *hazard* didefinisikan sebagai berikut:

$$\begin{aligned} h(t) &= \frac{f(t)}{S(t)} \\ &= \frac{\left[\frac{\beta}{\gamma}\right]\left[\frac{t}{\gamma}\right]^{\beta-1}}{\left[1 + \left[\frac{t}{\gamma}\right]^{\beta}\right]^{2}} \frac{\left[1 + \left[\frac{t}{\gamma}\right]^{\beta}\right]}{1} \\ &= \frac{\left[\frac{\beta}{\gamma}\right]\left[\frac{t}{\gamma}\right]^{\beta-1}}{1 + \left[\frac{t}{\gamma}\right]^{\beta}} \end{aligned} \tag{2.17}$$

Berdasarkan persamaan (2.17) maka akan didapat fungsi *hazard* kumulatif H(t) sebagai berikut

$$H(t) = \int_{0}^{t} \frac{\left[\frac{\beta}{\gamma}\right]\left[\frac{y}{\gamma}\right]^{\beta-1}}{1 + \left[\frac{y}{\gamma}\right]^{\beta}} dy \tag{2.18}$$

misal:

$$
\begin{aligned}
u &= 1 + \left[\frac{y}{\gamma}\right]^{\beta} \\
du &= \left[\frac{\beta}{\gamma}\right]\left[\frac{y}{\gamma}\right]^{\beta-1} dy \\
dy &= \frac{du}{\left[\frac{\beta}{\gamma}\right]\left[\frac{y}{\gamma}\right]^{\beta-1}}
\end{aligned} \tag{2.19}
$$

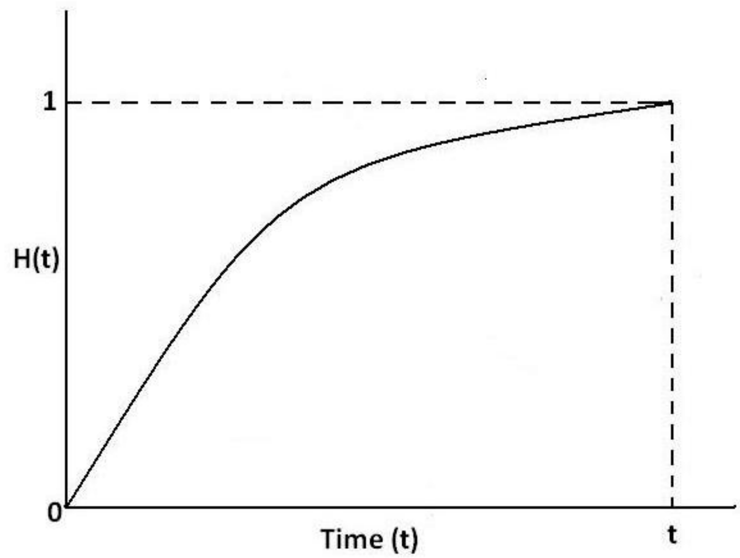
dan

$$
\begin{aligned}
y = 0 &\rightarrow u = 1 \\
y = t &\rightarrow u = 1 + \left(\frac{t}{\gamma}\right)^{\beta}
\end{aligned} \tag{2.20}
$$

maka dengan mensubtitusi persamaan (2.18) dan (2.19) diperoleh

$$
\begin{aligned}
H(t) &= \int_0^t \frac{\left[\frac{\beta}{\gamma}\right]\left[\frac{y}{\gamma}\right]^{\beta-1}}{1 + \left[\frac{y}{\gamma}\right]^{\beta}} dy \\
&= \int_0^{1+\left(\frac{t}{\gamma}\right)^{\beta}} u^{-1} du \\
&= \ln|u| \Big|_1^{1+\left(\frac{t}{\gamma}\right)^{\beta}} \\
&= \ln\left|1 + \left(\frac{t}{\gamma}\right)^{\beta}\right| - \ln|1| \\
&= \ln\left|1 + \left(\frac{t}{\gamma}\right)^{\beta}\right|
\end{aligned} \tag{2.21}
$$

dengan bentuk fungsi *hazard* kumulatif Log-Logistik seperti Gambar 2.10 dan H(t) adalah peluang suatu individu mengalami kegagalan seperti kematian pada waktu awal pengamatan *(start point)* hingga waktu t dengan nilai kurang dari




Gambar 2.10: Kurva fungsi *hazard* kumulatif Log-Logistik

sama dengan 1 yang artinya peluang kegagalan suatu individu semakin lama akan semakin bertambah hingga waktu t.

## 2.5 Statistika Terurut

**Definisi 2.5.1.** Misalkan himpunan variabel random $X_1, X_2, ..., X_n$ merupakan sampel random yang berukuran n dari suatu populasi dengan fungsi densitas $f(x)$ maka fungsi densitas peluang bersama dari variabel random independennya adalah sebagai berikut (Bain and Engelhardt, 1992)

$$f(x_1, x_2, ..., x_n) = f(x_1)f(x_2)...f(x_n) \tag{2.22}$$

Misalkan $X_1, X_2, ..., X_n$ adalah sampel random yang berukuran n dari fungsi densitas peluang $f(x)$ dimana untuk $f(x)$ kontinu dan $f(x) > 0$, $a < x < b$

maka fungsi densitas peluang dari statistik terurut ke-k adalah sebagai berikut

$$g_k(y_k) = \frac{n!}{(k-1)!(n-k)!} [F(y_k)]^{k-1} [1 - F(y_k)]^{n-k} f(y_k), a < y_k < b \tag{2.23}$$

## 2.6 Distribusi Multinomial

**Definisi 2.6.1.** Misal setiap percobaan dapat menghasilkan k macam hasil $x_1, x_2, ..., x_k$ dengan peluang $P_1, P_2, ..., P_k$ maka distribusi peluang acak $X_1, X_2, ..., X_k$ yang menyatakan terjadinya $x_1, x_2, ..., x_k$ dalam n usaha yang bebas maka

$$P(x_1, x_2, ..., x_k) = \frac{n!}{X_1!X_2!...X_k!}(P_1^{X_1})(P_2^{X_2})...(P_k^{X_k}) \tag{2.24}$$

dimana

n = banyak percobaan atau perlakuan

$X_1$ = banyak kejadian peristiwa ke-1

$X_2$ = banyak kejadian peristiwa ke-2

$X_k$ = banyak kejadian peristiwa ke-k

$P_1$ = peluang peristiwa 1 mengalami kejadian

$P_2$ = peluang peristiwa 2 mengalami kejadian

$P_k$ = peluang peristiwa k mengalami kejadian

Distribusi multinomial memiliki rataan $\mu_i$ untuk setiap $X_i$ dan covarians $\mu_i(1 - \mu_i)$ jika $i = j$ dan $-\mu_i\mu_j$ jika $i \neq j$. Selanjutnya, karena fungsi distribusi digunakan untuk peubah acak biner yang diskrit maka didapat

$$\begin{aligned} E[X_i] &= \sum_{x_i \epsilon X_i} x_i f(x_i;\theta) \\ &= 0\mu_1 + 0\mu_2 + ... + 1\mu_i \\ &= \mu_i \end{aligned} \tag{2.25}$$

Selanjutnya akan didapatkan $Var[X_i]$ dan $Cov[X_i]$ menggunakan rumus $Var[X_i] = E[X_i^2] - [E[X_i^2]]^2$ dimana

$$\begin{aligned} E[X_i^2] &= \sum_{x_i \epsilon X_i} x_i^2 f(x_i;\theta) \\ &= 0^2\mu_1 + 0^2\mu_2 + ... + 1^2\mu_i \\ &= \mu_i \\ E[X_i]^2 &= \mu_i^2 \\ Var[X_i] &= E[X_i^2] - E[X_i]^2 \\ &= \mu_i - \mu_i^2 \\ &= \mu_i(1 - \mu_i) \end{aligned} \tag{2.26}$$

untuk mendapatkan $Cov[X_iX_j]$ dengan $i \neq j$ menggunakan rumus $Cov[X_iX_j] = E[X_iX_j] - \{E[X_i].E[X_j]\}$, karena $X_i$ dan $X_j$ saling bebas maka $E[X_iX_j] = 0$ sehingga didapatkan

$$\begin{aligned} Cov[X_iX_j] &= E[X_iX_j] - \{E[X_i].E[X_j]\} \\ &= 0 - \mu_i.\mu_j \\ &= -\mu_i.\mu_j \end{aligned} \tag{2.27}$$

## 2.7 Uji Statistik Anderson-Darling

Uji Statistik-Anderson-Darling merupakan metode yang digunakan untuk menguji apakah sampel data berasal dari distribusi tertentu, seperti distribusi normal, Log-Logistik, dan lain-lain (Lawless, 1982).

1. Uji Hipotesis :

   $H_0: F(t) = F_0(t)$ (Data berdistribusi Log-Logistik)

   $H_1: F(t) \neq F_0(t)$ (Data tidak berdistribusi Log-Logistik)

   dengan statistik uji Anderson Darling ($A^2$)

   $$A^2 = -n - \sum_{i=1}^{n} \frac{(2i-1)}{n}(logF(Z_i) + log(1 - F(Z_{n+1-i}))) \tag{2.28}$$

   dimana

   $$Z_i = \frac{x_i - \bar{x}}{s} \tag{2.29}$$

   dengan

   A = statistik uji untuk metode Anderson-Darling

   n = ukuran sampel

   $x_i$ = data ke-i yang telah diurutkan

   $Z_i$ = data $x_i$ yang distandarisasi

   $\bar{x}$ = rata-rata data

   s = standar deviasi data

   $F(Z_i)$ = nilai fungsi distribusi kumulatif normal baku di $Z_i$

2. Menentukan taraf nyata $\alpha$

3. Keputusan :

   - Terima $H_0$ jika $A^2 > \alpha$
   - Tolak $H_0$ jika $A^2 < \alpha$

## 2.8 Matriks Informasi

**Definisi 2.8.1.** Misalkan $x_1, x_2, ..., x_k$ merupakan sampel random dari suatu distribusi dengan fungsi densitas peluang $f(x, \theta)$ dimana $\theta = (\theta_1, \theta_2, ..., \theta_k)^{'}$ merupakan vektor dari paramater-paramater yang belum diketahui nilainya yang merupakan himpunan dari $\Omega$ maka fungsi likelihood dari $\theta$ adalah

$$L(\theta) = \prod_{i=1}^{n} f(y_i; \theta) \tag{2.30}$$

sehingga persamaan maksimum likelihoodnya adalah sebagai berikut

$$U_j(\theta) = \frac{\partial L(\theta)}{\partial \theta_j}, j = 1, 2, ..., k \tag{2.31}$$

$U(\theta)$ mempunyai rata-rata nol dan matriks kovarian $I(\theta)^{-1}$ dimana

$$I_{ij}(\theta) = E\left[\frac{-\partial^2 L_L(\theta)}{\partial \theta_i \partial \theta_j}\right]; i, j = 1, 2, ..., k \tag{2.32}$$

sehingga matriks $I(\theta)$ disebut matriks informasi untuk distribusi normal (Lawless, 2003).

## 2.9 Interval Konfidensi

**Definisi 2.9.1.** Misalkan $X_1, X_2, ..., X_n$ merupakan variabel random non-negatif. Suatu $l(x_1, x_2, ..., x_n)$ dan $a(x_1, x_2, ..., x_n)$ disebut interval konfidensi 100% untuk $\theta$ jika :

$$P[l(X_1, X_2, ..., X_n) < \theta < a(X_1, X_2, ..., X_n)] = \alpha \tag{2.33}$$

dimana $0 < \alpha < 1$. Sedangkan nilai dari $l(x_1, x_2, ..., x_n)$ dan $a(x_1, x_2, ..., x_n)$ disebut batas konfidensi bawah dan atas (Bain & Engelhardt, 1992).

**Definisi 2.9.2.** *One-sided* Limit Konfidensi

1. Jika $P[l(X_1, X_2, ..., X_n) < \theta] = \alpha$ maka $l(x) = l(x_1, x_2, ..., x_n)$ disebut *one-sided lower* (batas bawah) limit konfidensi untuk $\theta$.

2. Jika $P[a(X_1, X_2, ..., X_n) < \theta] = \alpha$ maka $a(x) = a(x_1, x_2, ..., x_n)$ disebut *one-sided upper* (batas atas) limit konfidensi untuk $\theta$.

**Definisi 2.9.3.** Misalkan $Q = q(X_1, X_2, ..., X_n; \theta)$ adalah variabel random yang merupakan fungsi dari sampel acak $X_1, X_2, ..., X_n$ dan parameter $\theta$, maka Q disebut besaran pivot jika distribusinya tidak tergantung pada $\theta$ maupun parameter lainnya.

Sebagai ilustrasi, Jika $X_1, X_2, ..., X_n$ adalah sampel acak dari populasi yang berdistribusi normal $N(\mu, \sigma^2)$, dengan $\sigma^2$ diketahui dan $\mu$ tidak diketahui. Jika $\bar{x}$ merupakan MLE bagi $\mu$ maka pivot-nya adalah sebagai berikut :

$$Q = \bar{x} - \mu \tag{2.34}$$

Jika Q dikalikan dengan $\frac{1}{\sigma/n^{\frac{1}{2}}}$ maka

$$
\begin{aligned}
Z &= Q \cdot \frac{1}{\sigma/n^{\frac{1}{2}}} \\
Z &= \frac{\bar{x}-\mu}{\sigma/n^{\frac{1}{2}}} \sim N(0,1)
\end{aligned}
\tag{2.35}
$$

Dengan demikian interval konvidensi untuk $\mu$ adalah sebagai berikut :

$$
\begin{aligned}
P(-z_{\frac{\alpha}{2}} \leq Z \leq z_{\frac{\alpha}{2}}) &= 1-\alpha \\
P(-z_{\frac{\alpha}{2}} \leq \frac{\bar{x}-\mu}{\sigma/n^{\frac{1}{2}}} \leq z_{\frac{\alpha}{2}}) &= 1-\alpha
\end{aligned}
$$

maka diperoleh

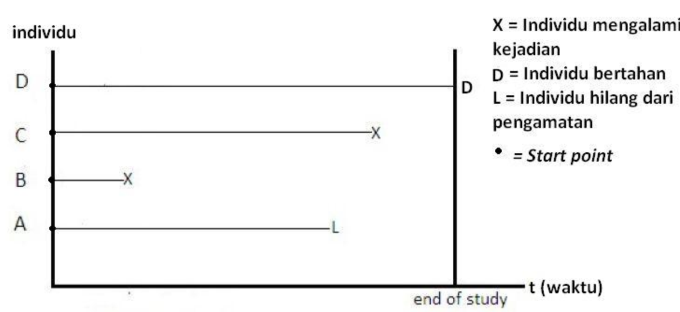
$$
\bar{x} - z_{\frac{\alpha}{2}}\sigma/n^{\frac{1}{2}} \leq \mu \leq \bar{x} + z_{\frac{\alpha}{2}}\sigma/n^{\frac{1}{2}} \tag{2.36}
$$

# BAB III

# PEMBAHASAN

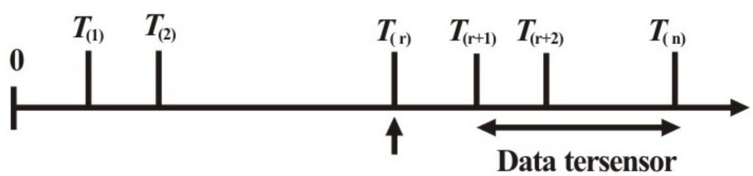
## 3.1 Model *Survival* Data Tersensor Kanan

Pada pembahasan sebelumnya telah dijelaskan mengenai fungsi densitas, fungsi *survival*, dan fungsi *hazard* dari distribusi Log-Logistik berdasarkan persamaan (2.11), (2.16), dan (2.17). Selanjutnya akan dibahas mengenai model *survival* dari data *survival* tersensor kanan *(right censored)*. Data dikatakan tersensor kanan *(right censored)* apabila individu tersebut hilang dari pengamatan *(loss to follow up)* atau individu tersebut masih bertahan sampai waktu pengamatan berakhir. Data tersensor kanan terdapat r buah pengamatan dalam sampel random yang berukuran n dan pengamatan dihentikan setelah kegagalan ke-r yang terjadi sebelum waktu $t_i$, seperti pada Gambar 3.1




Gambar 3.1: Ilustrasi model tersensor kanan

Misalkan T merupakan variabel random yang menyatakan waktu hingga suatu kegagalan terjadi, maka $T_{(1)}$ merupakan waktu hingga kegagalan terjadi pada individu ke-1, $T_{(2)}$ merupakan waktu hingga kegagalan terjadi pada individu ke-2, dan seterusnya hingga $T_{(r)}$ adalah waktu hingga kegagalan terjadi pada individu ke-r dengan data terdiri dari banyaknya individu yang mengalami kegagalan sebanyak r dengan urutan $T_{(1)} \leq T_{(2)} \leq T_{(3)} \leq ... \leq T_{(r)}$ dari sampel random yang berukuran n dengan urutan $T_1, T_2, T_3, ..., T_n$ seperti yang dilustrasikan pada Gambar 3.2 berikut:




Gambar 3.2: Ilustrasi model tersensor kanan

Oleh karena itu, diperoleh $f(t_1)$ yang merupakan fungsi densitas peluang dari variabel random individu ke-1, $f(t_2)$ yang merupakan fungsi densitas peluang dari variabel random individu ke-2, dan seterusnya hingga $f(t_r)$ dengan individu yang mengalami kegagalan yaitu individu ke-1 sampai individu ke-r masing-masing sebanyak satu orang yang mengalami kegagalan, sehingga fungsi densitas untuk satu individu pada waktu ke $t_i$ yang mengikuti distribusi Log-Logistik adalah sebagai berikut :

$$f(t_i) = \frac{\left[\frac{\beta}{\gamma}\right]\left[\frac{t_i}{\gamma}\right]^{\beta-1}}{\left[1+\left[\frac{t_i}{\gamma}\right]^{\beta}\right]^2}; t, \gamma, \beta > 0$$

Sementara itu, individu yang masih bertahan melebihi waktu pengamatan dituliskan dengan $T_{r+1}, T_{r+2}, T_{r+3}, ..., T_n$ sebanyak $(n-r)$ data pengamatan yang disebut dengan data tersensor kanan. Sampel random yang berukuran n dengan jumlah kegagalan sebanyak r memiliki urutan yang mungkin terjadi dan mengikuti distribusi multinomial sehingga didapat $\frac{n!}{1!1...1!(n-r)!}$. Oleh karena itu, fungsi densitas peluang bersama dari $T_1, T_2, T_3, ..., T_n$ dari data yang diamati dapat ditulis sebagai berikut :

$$\begin{aligned}
f(t_1, t_2, ..., t_r) &= \frac{n!}{(n-r)!} f(t_1) f(t_2) ... f(t_r) [P(T_{r+1} \geq t_r) ... P(T_n \geq t_r)] \\
&= \frac{n!}{(n-r)!} \left[ \prod_{i=1}^{r} f(t_i) \right] [(1 - P(T_{r+1} < t_r)) ... (1 - P(T_n < t_r))] \\
&= \frac{n!}{(n-r)!} \left[ \prod_{i=1}^{r} f(t_i) \right] [(1 - F(t_r)) ... (1 - F(t_r))] \\
&= \frac{n!}{(n-r)!} \left[ \prod_{i=1}^{r} f(t_i) \right] [1 - F(t_r)]^{n-r} \\
&= \frac{n!}{(n-r)!} \left[ \prod_{i=1}^{r} f(t_i) \right] [S(t_r)]^{n-r}
\end{aligned}$$

sehingga didapat fungsi densitas peluang bersama data *survival* tersensor kanan dari $t_1, t_2, t_3, ..., t_r$ untuk $r < n$ adalah sebagai berikut :

$$f(t_1, t_2, ..., t_r) = \frac{n!}{(n-r)!} \left[ \prod_{i=1}^{r} f(t_i) \right] [S(t_r)]^{n-r} \tag{3.1}$$

## 3.2 Estimasi Paramater dengan Maksimum Likelihood Estimator *(MLE)*

Pada pembahasan selanjutnya akan didapatkan nilai estimasi parameter dari distribusi Log-Logistik dengan menggunakan metode *MLE*.

### 3.2.1 Fungsi Likelihood dari Distribusi Log-Logistik

Dalam menduga parameter pada distribusi Log-Logistik dengan menggunakan *MLE* diperlukan model *survival* yang ditentukan berdasarkan fungsi *likelihood*-nya. Model *survival* didapat dengan mensubtitusi persamaan (2.11) dan (2.16) pada persamaan (3.1) sehingga didapatkan *likelihood*-nya sebagai berikut :

$$
\begin{aligned}
f(t_1, t_2, ..., t_r) &= \frac{n!}{(n-r)!}\left[\prod_{i=1}^{r} f(t_i)\right][S(t_r)]^{n-r} \\
&= \frac{n!}{(n-r)!}\left[\prod_{i=1}^{r} \frac{(\frac{\beta}{\gamma})^r(\frac{t_i}{\gamma})^{\beta-1}}{[1+(\frac{t_i}{\gamma})^\beta]^2}\right]\left[\frac{1}{1+(\frac{t_r}{\gamma})^\beta}\right]^{n-r} \\
&= \frac{n!}{(n-r)!}\left[(\frac{\beta}{\gamma})^r\prod_{i=1}^{r} \frac{(\frac{t_i}{\gamma})^{\beta-1}}{[1+(\frac{t_i}{\gamma})^\beta]^2}\right]\left[\frac{1}{1+(\frac{t_r}{\gamma})^\beta}\right]^{n-r}
\end{aligned}
$$

maka fungsi *likelihood* dari distribusi Log-Logistik untuk data *survival* tersensor kanan memiliki bentuk

$$
L(\gamma,\beta) = \frac{n!}{(n-r)!}\left[(\frac{\beta}{\gamma})^r\prod_{i=1}^{r} \frac{(\frac{t_i}{\gamma})^{\beta-1}}{[1+(\frac{t_i}{\gamma})^\beta]^2}\right]\left[\frac{1}{1+(\frac{t_r}{\gamma})^\beta}\right]^{n-r} \tag{3.2}
$$

### 3.2.2 Maksimum Likelihood Estimator *(MLE)*

Pada tulisan ini akan digunakan metode maksimum *likelihood* untuk mendapatkan estimasi parameter dari distribusi Log-Logistik. Metode maksimum *likelihood* menggunakan nilai dalam ruang parameter $\Omega$ yang bersesuaian dengan harga kemungkinan maksimum dari data observasi sebagai estimasi dari parameter yang tidak diketahui.

Dalam aplikasinya $L(\gamma, \beta)$ menunjukkan fungsi densitas peluang bersama dari sampel random. Jika $\Omega$ ruang parameter yang merupakan interval terbuka dan $L(\gamma, \beta)$ merupakan fungsi yang dapat diturunkan serta diasumsikan maksimum pada $\Omega$, maka persamaan maksimum *likelihood*-nya adalah

$$\frac{\partial L(\gamma, \beta)}{\partial \gamma} = 0 \tag{3.3}$$

dan

$$\frac{\partial L(\gamma, \beta)}{\partial \beta} = 0 \tag{3.4}$$

Jika penyelesaian dari persamaan (3.3) dan (3.4) ada maka maksimum dari $L(\gamma, \beta)$ dapat terpenuhi tetapi apabila penyelesaian dari persamaan tersebut sulit untuk diselesaikan maka fungsi $L(\gamma, \beta)$ dapat dibuat logaritma naturalnya (selanjutnya ditulis dengan $L_L(\gamma, \beta)$) dengan ketentuan $L_L(\gamma, \beta)$ maksimum.

Dari persamaan (3.2) yakni persamaan *likelihood* sampel tersensor kanan diperoleh fungsi *likelihood* untuk distribusi Log-logistik sebagai berikut :

$$L(\gamma, \beta) = \frac{n!}{(n-r)!} \left[ (\frac{\beta}{\gamma})^r \prod_{i=1}^{r} \frac{(\frac{t_i}{\gamma})^{\beta-1}}{[1 + (\frac{t_i}{\gamma})^{\beta}]^2} \right] \left[ \frac{1}{1 + (\frac{t_r}{\gamma})^{\beta}} \right]^{n-r}$$

Oleh karena $L(\gamma, \beta)$ sulit diturunkan terhadap $\gamma$ ataupun $\beta$ maka perlu dibuat persamaan logaritma natural maksimum *likelihood*-nya, sehingga didapat fungsi *likelihood*-nya sebagai berikut :

$$
\begin{aligned}
L_L(\gamma,\beta) &= \ln\frac{n!}{(n-r)!} + \ln\beta^r - \ln\gamma^r + \sum_{i=1}^{r}\ln\left(\frac{t_i}{\gamma}\right)^{\beta-1} \\
&\quad + \sum_{i=1}^{r}\ln\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^{-2} + \ln\left[1+\left(\frac{t_r}{\gamma}\right)^{\beta}\right]^{r-n} \\
&= \ln\frac{n!}{(n-r)!} + \ln\beta^r - \ln\gamma^r + \sum_{i=1}^{r}(\beta-1)\ln\left(\frac{t_i}{\gamma}\right) \\
&\quad - \sum_{i=1}^{r}2\ln\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right] + (r-n)\ln\left[1+\left(\frac{t_r}{\gamma}\right)^{\beta}\right]
\end{aligned}
\tag{3.5}
$$

Berdasarkan persamaan (3.5) akan didapatkan turunan pertama dan kedua terhadap $\gamma$ dan $\beta$, berikut adalah perhitungan turunan pertama $L_L(\gamma,\beta)$ terhadap $\gamma$ yaitu sebagai berikut :

$$
\begin{aligned}
\frac{\partial L_L(\gamma,\beta)}{\partial\gamma} &= -\frac{r}{\gamma} + \frac{r(1-\beta)}{\gamma} - 2\sum_{i=1}^{r}\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^{-1}\left(\frac{-\beta}{\gamma}\right)\left(\frac{t_i}{\gamma}\right)^{\beta} \\
&\quad - \frac{(r-n)}{1+\left(\frac{t_r}{\gamma}\right)^{\beta}}\left(\frac{\beta}{\gamma}\right)\left(\frac{t_r}{\gamma}\right)^{\beta} \\
&= \frac{-r\beta}{\gamma} + \left(\frac{2\beta}{\gamma}\right)\sum_{i=1}^{r}\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^{-1}\left(\frac{t_i}{\gamma}\right)^{\beta} - \frac{(r-n)}{1+\left(\frac{t_r}{\gamma}\right)^{\beta}}\left(\frac{\beta}{\gamma}\right)\left(\frac{t_r}{\gamma}\right)^{\beta}
\end{aligned}
\tag{3.6}
$$

berdasarkan persamaan (3.6) maka akan didapatkan perhitungan turunan kedua dari $L_L(\gamma,\beta)$ terhadap $\gamma$ yaitu sebagai berikut :

$$\begin{aligned}\frac{\partial^2 L_L(\gamma,\beta)}{\partial\gamma^2} &= \frac{r\beta}{\gamma^2} - 2\left(\frac{\beta}{\gamma}\right)\left[\sum_{i=1}^{r}\frac{\left[\left(\frac{\beta+1}{\gamma}\right)\left(\frac{t_i}{\gamma}\right)^{\beta}\left(1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right)-\left(\frac{\beta}{\gamma}\right)\left(\frac{t_i}{\gamma}\right)^{2\beta}\right]}{\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^2}\right] \\ &\quad - \frac{(r-n)}{\left[1+\left(\frac{t_r}{\gamma}\right)^{\beta}\right]^2}\left[\left(\frac{-\beta-1}{\gamma}\right)\left(\frac{\beta}{\gamma}\right)\left(\frac{t_r}{\gamma}\right)^{\beta}\left(1+\left(\frac{t_r}{\gamma}\right)^{\beta}\right)+\left(\frac{\beta}{\gamma}\right)^2\left(\frac{t_r}{\gamma}\right)^{2\beta}\right]\end{aligned} \tag{3.7}$$

Oleh karena perhitungan turunan kedua $L(\gamma,\beta)$ terhadap $\gamma$ mendapatkan hasil yang negatif maka *MLE* untuk $\gamma$ yang dilambangkan dengan $\hat{\gamma}$ diperoleh dengan menyelesaikan persamaan (3.6) yaitu sebagai berikut :

$$-\frac{r\hat{\beta}}{\hat{\gamma}}+\left(\frac{2\hat{\beta}}{\hat{\gamma}}\right)\sum_{i=1}^{r}\left[1+\left(1+\left(\frac{t_i}{\hat{\gamma}}\right)\right)^{\hat{\beta}}\right]^{-1}\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}}-\frac{(r-n)}{1+\left(1+\left(\frac{t_r}{\hat{\gamma}}\right)\right)^{\hat{\beta}}}\left(\frac{\hat{\beta}}{\hat{\gamma}}\right)\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}}=0 \tag{3.8}$$

Selanjutnya akan didapatkan nilai estimasi paramater $\hat{\beta}$ dengan menghitung turunan dari $L_L(\gamma,\beta)$ terhadap $\beta$ yaitu sebagai berikut :

$$\begin{aligned}\frac{\partial L_L(\gamma,\beta)}{\partial\beta} &= \frac{r}{\beta}+\sum_{i=1}^{r}\ln\frac{t_i}{\gamma}-2\sum_{i=1}^{r}\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^{-1}\left(\frac{t_i}{\gamma}\right)^{\beta}\ln\left(\frac{t_i}{\gamma}\right) \\ &\quad+\frac{(r-n)}{1+\left(\frac{t_r}{\gamma}\right)^{\beta}}\left(\frac{t_r}{\gamma}\right)^{\beta}\ln\left(\frac{t_r}{\gamma}\right)\end{aligned} \tag{3.9}$$

Perhitungan turunan kedua dari $L_L(\gamma,\beta)$ terhadap $\beta$ adalah sebagai berikut :

$$\frac{\partial^2 L_L(\gamma,\beta)}{\partial\beta^2} = -\left(\frac{r}{\beta^2}\right)-2\sum_{i=1}^{r}\frac{\left(\frac{t_i}{\gamma}\right)^{\beta}\left(\ln\left(\frac{t_i}{\gamma}\right)\right)^2}{\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^2}+\frac{(r-n)}{\left[1+\left(\frac{t_r}{\gamma}\right)^{\beta}\right]^2}\left(\frac{t_r}{\gamma}\right)^{\beta}\left(\ln\left(\frac{t_r}{\gamma}\right)\right)^2 \tag{3.10}$$

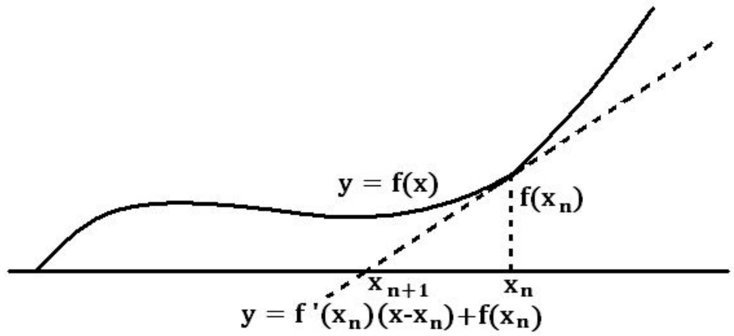
Oleh karena perhitungan turunan kedua $L(\gamma, \beta)$ terhadap $\beta$ mendapatkan hasil yang negatif maka *MLE* untuk $\gamma$ yang dilambangkan dengan $\hat{\gamma}$ diperoleh dengan menyelesaikan persamaan (3.9) yaitu sebagai berikut :

$$\frac{r}{\hat{\beta}}+\sum_{i=1}^{r}\ln\frac{t_i}{\hat{\gamma}}-2\sum_{i=1}^{r}\left[1+\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}}\right]^{-1}\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}}\ln\left(\frac{t_i}{\hat{\gamma}}\right)+\frac{(r-n)}{1+\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}}}\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}}\ln\left(\frac{t_r}{\hat{\gamma}}\right)=0 \tag{3.11}$$

Berdasarkan perhitungan yang telah dilakukan maka *MLE* untuk $\hat{\gamma}$ terdapat pada persamaan (3.8) dan $\hat{\beta}$ pada persamaan (3.11). Kedua persamaan tersebut sulit diselesaikan secara eksplisit karena memiliki bentuk yang sangat kompleks dan tidaklah linier sehingga perlu dilakukan metode numerik Newton-Raphson untuk mendapatkan nilai estimasi paramater dari $\hat{\gamma}$ dan $\hat{\beta}$ pada analisis *survival* menggunakan model distribusi Log-Logistik. Menurut Tirta (2009) algoritma pokok dari metode ini memiliki langkah-langkah sebagai berikut :

1. Tentukan hampiran awal $\gamma_n$ dan $\beta_n$ pada langkah ke-n, n $= 0, 1, 2, ...$

2. Misalkan $x_n$ merupakan hampiran awal untuk $\gamma_n$ maka untuk menentukan hampran awal $\gamma_n$ langkah pertama yang dilakukan adalah menentukan gradien garis singgung terhadap kurva $y = f(x_n)$ di titik $(x_n, f(x_n))$ dimana $f'(x_n) = 0$ merupakan gradien garis. Selanjutnya, tentukan persamaan garis singgungnya yaitu $y = f'(x_n)(x - x_n) + f(x_n)$ seperti pada Gambar 3.3 dan begitu pula untuk $\beta$.




Gambar 3.3: Iterasi Newton Raphson

3. Lakukan iterasi dengan persamaan :

$$\gamma_{n+1} = \gamma_n - \frac{f'(\gamma_n)}{f''(\gamma_n)}$$

untuk menduga nilai parameter $\gamma$ dan begitu pula untuk $\beta$, karena persamaan turunan pertama parameter $\gamma$ dan $\beta$ tidaklah linear dan terlalu kompleks untuk diselesaikan maka fungsi turunan pertamanya haruslah berupa vektor $\mathbf{T}$ dan turunan keduanya disebut matriks *Hessian* H, dengan bentuk multivariat dari metode Newton-Raphson yaitu sebagi berikut :

$$G_{n+1} = G_n - \frac{\mathbf{T}(\gamma, \beta)}{H(\gamma, \beta)}$$

sehingga didapat

$$\mathbf{T}(\gamma, \beta) = \begin{bmatrix} \frac{\partial L_L(\gamma,\beta)}{\partial \gamma} \\ \frac{\partial L_L(\gamma,\beta)}{\partial \beta} \end{bmatrix}$$

dan

$$\mathbf{H}(\gamma, \beta) = \begin{bmatrix} \frac{\partial^2 L_L(\gamma,\beta)}{\partial \gamma^2} & \frac{\partial^2 L_L(\gamma,\beta)}{\partial \gamma \partial \beta} \\ \frac{\partial^2 L_L(\gamma,\beta)}{\partial \beta \partial \gamma} & \frac{\partial^2 L_L(\gamma,\beta)}{\partial \beta^2} \end{bmatrix}$$

dimana

$$\mathbf{G_{n+1}} = \begin{bmatrix} \gamma_{n+1} \\ \beta_{n+1} \end{bmatrix}$$

dan

$$\mathbf{G_n} = \begin{bmatrix} \gamma_n \\ \beta_n \end{bmatrix}$$

4. Hentikan proses iterasi sampai mendapatkan nilai parameter yang konvergen (jika selisih antara $\gamma_{n+1}$ dan $\gamma_n$ kurang dari $10^{-6}$ sehingga didapatkan nilai estimasi parameter yang stabil).

## 3.3 Interval Konfidensi

Setelah $\hat{\gamma}$ dan $\hat{\beta}$ diperoleh, selanjutnya akan ditentukan interval konfidensi untuk $\gamma$ dan $\beta$. Asumsikan pendugaan untuk $\gamma$ dan $\beta$ didasarkan pada sampel besar yang secara aproksimasi berdistribusi normal N(0,1), sehingga dengan menggunakan pendekatan pivot (berdasarkan persamaan (2.35)) dapat ditulis sebagai berikut :

$$Z_1 = \frac{\hat{\gamma} - \gamma}{se(\hat{\gamma})} \tag{3.12}$$

dan

$$Z_2 = \frac{\hat{\beta} - \beta}{se(\hat{\beta})} \tag{3.13}$$

Selanjutnya untuk menentukan $se(\hat{\gamma})$ dan $se(\hat{\beta})$ langkah pertama yang dilakukan adalah menentukan matriks informasi ($I(\gamma, \beta)$) yaitu

$$I(\gamma, \beta) = E\left(\frac{-\partial^2 L_L(\gamma, \beta)}{\partial\gamma\partial\beta}\right) \tag{3.14}$$

dengan $\frac{\partial^2 L_L(\gamma,\beta)}{\partial\gamma^2}$ pada persamaan (3.7) dan $\frac{\partial^2 L_L(\gamma,\beta)}{\partial\beta^2}$ pada persamaan (3.10) maka didapat

$$\mathbf{I}(\gamma,\beta) = \begin{bmatrix} -E\left[\frac{\partial^2 L_L(\gamma,\beta)}{\partial\gamma^2}\right] & -E\left[\frac{\partial^2 L_L(\gamma,\beta)}{\partial\gamma\partial\beta}\right] \\ -E\left[\frac{\partial^2 L_L(\gamma,\beta)}{\partial\beta\partial\gamma}\right] & -E\left[\frac{\partial^2 L_L(\gamma,\beta)}{\partial\beta^2}\right] \end{bmatrix}$$

kemudian berdasarkan persamaan (3.8) yang merupakan $\frac{\partial L_L(\gamma,\beta)}{\partial\gamma}$ akan dihitung $\frac{\partial^2 L_L}{\partial\gamma\partial\beta}$ yaitu sebagai berikut :

$$\begin{aligned}\frac{\partial^2 L_L}{\partial\gamma\partial\beta} &= \frac{2}{\gamma}\left[\sum_{i=1}^{r}\left[\left(\frac{t_i}{\gamma}\right)^{\beta}\left[1+\beta\ln\left(\frac{t_i}{\gamma}\right)\right]\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]-\beta\left(\frac{t_i}{\gamma}\right)^{2\beta}\ln\left(\frac{t_i}{\gamma}\right)\right]\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^{-2}\right] \\ &\quad -\frac{(r-n)}{\left[1+\left(\frac{t_r}{\gamma}\right)^{\beta}\right]^{2}}\left[\left(\frac{1}{\gamma}\left(\frac{t_r}{\gamma}\right)^{\beta}\right)\left[1+\beta\ln\left(\frac{t_r}{\gamma}\right)\right]-\left(\frac{\beta}{\gamma}\right)\left(\frac{t_r}{\gamma}\right)^{2\beta}\right]\ln\left(\frac{t_r}{\gamma}\right)-\left(\frac{r}{\gamma}\right)\end{aligned} \tag{3.15}$$

dan berdasarkan persamaan (3.11) yang merupakan $\frac{\partial L_L(\gamma,\beta)}{\partial\beta}$ akan dihitung $\frac{\partial^2 L_L}{\partial\beta\partial\gamma}$ yaitu sebagai berikut :

$$\begin{aligned}\frac{\partial^2 L_L}{\partial\gamma\partial\beta} &= \frac{2}{\gamma}\left[\sum_{i=1}^{r}\left[\left(\frac{t_i}{\gamma}\right)^{\beta}\left[1+\beta\ln\left(\frac{t_i}{\gamma}\right)\right]\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]-\beta\left(\frac{t_i}{\gamma}\right)^{2\beta}\ln\left(\frac{t_i}{\gamma}\right)\right]\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^{-2}\right] \\ &\quad -\frac{(r-n)}{\left[1+\left(\frac{t_r}{\gamma}\right)^{\beta}\right]^{2}}\left[\left(\frac{1}{\gamma}\left(\frac{t_r}{\gamma}\right)^{\beta}\right)\left[1+\beta\ln\left(\frac{t_r}{\gamma}\right)\right]-\left(\frac{\beta}{\gamma}\right)\left(\frac{t_r}{\gamma}\right)^{2\beta}\right]\ln\left(\frac{t_r}{\gamma}\right)-\left(\frac{r}{\gamma}\right)\end{aligned} \tag{3.16}$$

Selanjutnya berdasarkan persamaan (3.15) dan (3.16) didapatkan

$$\frac{\partial^2 L_L(\gamma,\beta)}{\partial\gamma\partial\beta} = \frac{\partial^2 L_L(\gamma,\beta)}{\partial\beta\partial\gamma}$$

Jika $\hat{\gamma}$ dan $\hat{\beta}$ adalah *MLE* dari $\gamma$ dan $\beta$ dengan invers matriks informasi adalah penduga matriks covarians $I(\gamma,\beta)^{-1}$, maka matriks invers dapat ditulis sebagai berikut :

$$\mathbf{I}(\gamma,\beta)^{\mathbf{-1}} = \begin{bmatrix} var(\hat{\gamma}) & cov(\hat{\gamma}\hat{\beta}) \\ cov(\hat{\beta}\hat{\gamma}) & var(\hat{\beta}) \end{bmatrix}$$

dan

$$\hat{V} = I(\gamma,\beta)^{-1}$$

sehingga didapat

$$\hat{V} = \begin{bmatrix} var(\hat{\gamma}) & cov(\hat{\gamma}\hat{\beta}) \\ cov(\hat{\beta}\hat{\gamma}) & var(\hat{\beta}) \end{bmatrix}$$

dan matriks kovarian dapat ditulis sebagai berikut

$$\hat{V} = \begin{bmatrix} \hat{V}_{11} & \hat{V}_{12} \\ \hat{V}_{21} & \hat{V}_{22} \end{bmatrix}$$

dengan *standard error* untuk $\hat{\gamma}$ adalah

$$se(\hat{\gamma}) = \hat{V}_{11}^{\frac{1}{2}} \tag{3.17}$$

dan *standard error* untuk $\hat{\beta}$ adalah

$$se(\hat{\beta}) = \hat{V}_{22}^{\frac{1}{2}} \tag{3.18}$$

Oleh karena itu, interval konfidensi untuk $\gamma$ dengan taraf signifikansi

$1-\alpha$ adalah sebagai berikut :

$$\begin{aligned} P(-z_{\frac{\alpha}{2}} \leq Z_1 \leq z_{\frac{\alpha}{2}}) &= 1-\alpha \\ P(-z_{\frac{\alpha}{2}} \leq \frac{\hat{\gamma}-\gamma}{se(\hat{\gamma})} \leq z_{\frac{\alpha}{2}}) &= 1-\alpha \end{aligned}$$

maka diperoleh

$$\hat{\gamma} - z_{\frac{\alpha}{2}} se(\hat{\gamma}) \leq \gamma \leq \hat{\gamma} + z_{\frac{\alpha}{2}} se(\hat{\gamma}) \tag{3.19}$$

dan interval konfidensi untuk $\beta$ dengan taraf signifikansi $1-\alpha$ adalah sebagai berikut :

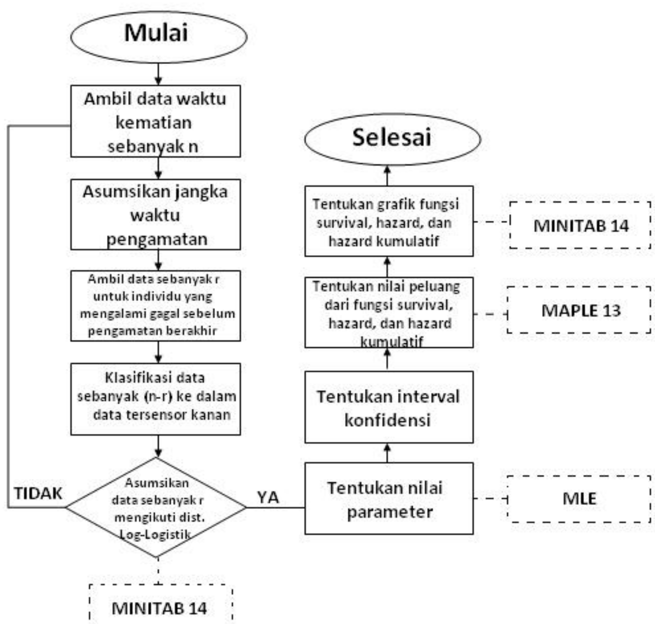
$$\begin{aligned} P(-z_{\frac{\alpha}{2}} \leq Z_2 \leq z_{\frac{\alpha}{2}}) &= 1-\alpha \\ P(-z_{\frac{\alpha}{2}} \leq \frac{\hat{\beta}-\beta}{se(\hat{\beta})} \leq z_{\frac{\alpha}{2}}) &= 1-\alpha \end{aligned}$$

maka diperoleh

$$\hat{\beta} - z_{\frac{\alpha}{2}} se(\hat{\beta}) \leq \beta \leq \hat{\beta} + z_{\frac{\alpha}{2}} se(\hat{\beta}) \tag{3.20}$$

## 3.4 Diagram Alir Data Tersensor Kanan Menggunakan Distribusi Log-Logistik

Setelah didapatkan model *survival* data tersensor kanan menggunakan distribusi Log-Logistik dapat dibuat diagram alir data tersensor kanan menggunakan distribusi Log-Logistik. Berikut gambar diagram alir data tersensor kanan menggunakan distribusi Log-Logistik :




Gambar 3.4: Diagram Alir Data Tersensor Kanan Menggunakan Distribusi Log-Logistik

Selanjutnya, dengan menggunakan langkah-langkah di atas akan diketahui probabilitas *survival*, probabilitas waktu kegagalan individu pada waktu tertentu, dan probabilitas kegagalan individu dari awal pengamatan hingga waktu tertentu dengan melibatkan bentuk data *time to event* seperti data kematian.

## 3.5 Studi Kasus

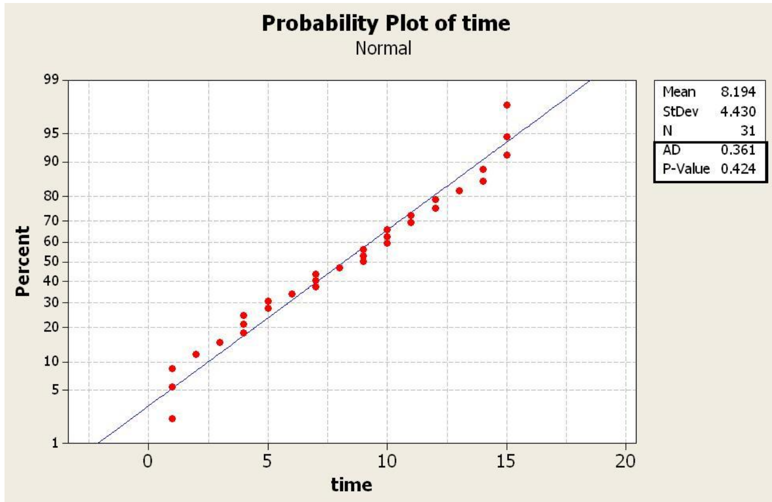
Data yang digunakan dalam tuisan ini adalah data kematian pasien setelah diberi perlakuan berupa operasi jantung di salah satu Rumah Sakit di Jakarta pada tahun 2014. Data terdiri dari 40 pasien setelah melakukan operasi jantung hingga mengalami kegagalan berupa kematian. Waktu *survival* dari pasien adalah waktu (dalam hari) hingga terjadinya kematian. Dari 40 pasien terdapat 31 pasien dengan waktu *survival* tidak tersensor dan 9 pasien dengan waktu *survival* tersensor (pasien yang masih bertahan hingga lebih dari 15 hari waktu pengamatan). Selanjutnya akan dihitung peluang hidup (probabilitas *survival*) dan peluang kegagalan (probabilitas kematian) seorang pasien setelah diberi perlakuan berupa operasi jantung dalam waktu 10 hari. Pengolahan data dilakukan dengan bantuan *software MINITAB* 14 dan *MAPLE* 13.

Langkah pertama dalam aplikasi ini adalah mengasumsikan data kematian 31 pasien setelah melakukan operasi jantung mengikuti distribusi Log-Logistik. Untuk menguji apakah data yang diasumsikan berdistribusi Log-Logistik maka akan dilakukan uji statistik Anderson-Darling dengan hipotesis

$H_0$: $F(t) = F_0(t)$ (Data berdistribusi Log-Logistik)

$H_1$: $F(t) \neq F_0(t)$ (Data tidak berdistribusi Log-Logistik)

Berdasarkan Gambar 3.5 didapat nilai statistik Anderson-Darling ($A^2$) sebesar 0,361 dan p = $F_0(t)$ = 0,424 dengan $F(t)$ merupakan fungsi kumulatif data berdistribusi Log-Logistik. Berdasarkan Gambar 3.4 didapat $A^2 > 0,05$, sehingga $H_0$ diterima yang berarti data mengikuti distribusi Log-Logistik.




Gambar 3.5: Kurva Anderson-Darling Data Kematian Pasien Setelah Operasi Jantung

Langkah berikutnya adalah mencari nilai parameter menggunakan metode *MLE* dengan bantuan software *MINITAB* 14. Fungsi densitas distribusi Logistik yang digunakan adalah

$$f(y) = \frac{\exp\left[\frac{lny-\mu}{\sigma}\right]}{\sigma\left[1+\exp\left[\frac{lny-\mu}{\sigma}\right]^2\right]} \tag{3.21}$$

dengan

$\mu$ = *location parameter*

$\sigma$ = *scale parameter*

Bentuk fungsi densitas pada persamaan (3.22) merupakan hasil generalisasi dari fungsi densitas distribusi Logistik pada persamaan (3.21) (Ojo & Olapade, 2003). Misal variabel random T yang menyatakan waktu hingga kematian terjadi dan T dikatakan mengikuti distribusi Log-Logistik dengan parameter $\gamma$ dan $\beta$

maka fungsi densitas Log-Logistik ialah sebagai berikut:

$$f(t) = \frac{\left[\frac{\beta}{\gamma}\right]\left[\frac{t}{\gamma}\right]^{\beta-1}}{\left[1+\left[\frac{t}{\gamma}\right]^{\beta}\right]^{2}} \tag{3.22}$$

dengan

$\gamma = \exp(\mu)$

$\beta = \frac{1}{\sigma}$

Hasil dugaan untuk $\gamma$ dapat diperoleh dari nilai *location parameter* ($\mu$) yaitu 2,3451 sehingga didapat nilai $\hat{\gamma}$ sebagai berikut :

$$\begin{aligned} \hat{\gamma} &= \exp(\mu) \\ &= \exp(2,3451) \\ &= 10,43431611 \\ &\approx 10,435 \end{aligned} \tag{3.23}$$

dan nilai *scale parameter* ($\sigma$) yaitu 0,785485 sehingga didapat nilai $\hat{\beta}$ sebagai berikut :

$$\begin{aligned} \hat{\beta} &= \frac{1}{\sigma} \\ &= \frac{1}{0,682343} \\ &= 1,465538593 \\ &\approx 1,466 \end{aligned} \tag{3.24}$$

Setelah nilai parameter $\hat{\gamma}$ dan $\hat{\beta}$ diketahui maka akan ditentukan interval konfidensi untuk $\gamma$ dan $\beta$. Dari perhitungan sebelumnya telah didapat nilai $\hat{\gamma}$ pada persamaan (3.23) dan $\hat{\beta}$ persamaan (3.24) dengan bantuan *software MINITAB* 14. Selanjutnya, untuk mendapatkan matriks informasi akan dicari turunan kedua untuk $\gamma$ berdasarkan persamaan (3.7) yaitu

$$\frac{\partial^2 L_L(\gamma,\beta)}{\partial\gamma^2} = -0,27143 \tag{3.25}$$

dan turunan kedua untuk $\beta$ berdasarkan persamaan (3.10) yaitu

$$\frac{\partial^2 L_L(\gamma,\beta)}{\partial\beta^2} = -18,4893 \tag{3.26}$$

selanjutnya akan dicari $\frac{\partial^2 L_L(\gamma,\beta)}{\partial\gamma\partial\beta}$ berdasarkan persamaan (3.15) yaitu

$$\frac{\partial^2 L_L(\gamma,\beta)}{\partial\gamma\partial\beta} = -0,3669 \tag{3.27}$$

dan didapat $\frac{\partial^2 L_L(\gamma,\beta)}{\partial\beta\partial\gamma}$ berdasarkan persamaan (3.16) yaitu

$$\frac{\partial^2 L_L(\gamma,\beta)}{\partial\beta\partial\gamma} = -0,3669 \tag{3.28}$$

sehingga didapat matriks informasi sebagai berikut :

$$\mathbf{I}(\gamma,\beta) = \begin{bmatrix} 0,27143 & 0,3669 \\ 0,3669 & 18,4893 \end{bmatrix}$$

dan matriks kovarian

$$\hat{V} = I(\gamma,\beta)^{-1}$$

sehingga didapat

$$\hat{\mathbf{V}} = \begin{bmatrix} 3,78579 & -0,07515 \\ -0,07515 & 0,05558 \end{bmatrix}$$

Selanjutnya akan didapatkan interval konfidensi untuk $\gamma$. Oleh karena itu, akan dicari nilai dari $se(\hat{\gamma})$. Berdasarkan persamaan (3.17) nilai *standard error* untuk $\hat{\gamma}$ adalah

$$\begin{aligned} se(\hat{\gamma}) &= \hat{V}_{11}^{\frac{1}{2}} \\ &= 1,945710532 \\ &\approx 1,946 \end{aligned} \tag{3.29}$$

sehingga interval konfidensi untuk $\gamma$ didapatkan dari pendekatan nilai pivot berdasarkan persamaan (3.12), dengan taraf signifikansi 0,05. Berdasarkan persamaan (3.19) didapat

$$10,435 - 1,96(1,946) \leq \gamma \leq 10,435 + 1,96(1,946)$$

sehingga diperoleh interval konfidensi untuk $\gamma$ yaitu sebagai berikut :

$$9,97244 \leq \gamma \leq 10,89756$$

dan untuk mendapatkan interval konfidensi untuk $\beta$ maka akan dicari nilai dari $se(\hat{\beta})$. Berdasarkan persamaan (3.18) nilai *standard error* untuk $\hat{\beta}$ adalah sebagai berikut :

$$\begin{aligned} se(\hat{\beta}) &= \hat{V}_{22}^{\frac{1}{2}} \\ &= 0,2357474037 \\ &\approx 0,236 \end{aligned} \tag{3.30}$$

sehingga interval konfidensi untuk $\beta$ didapatkan dari pendekatan nilai pivot berdasarkan persamaan (3.13), dengan taraf signifikansi 0,05. Berdasarkan persamaan (3.20) didapat

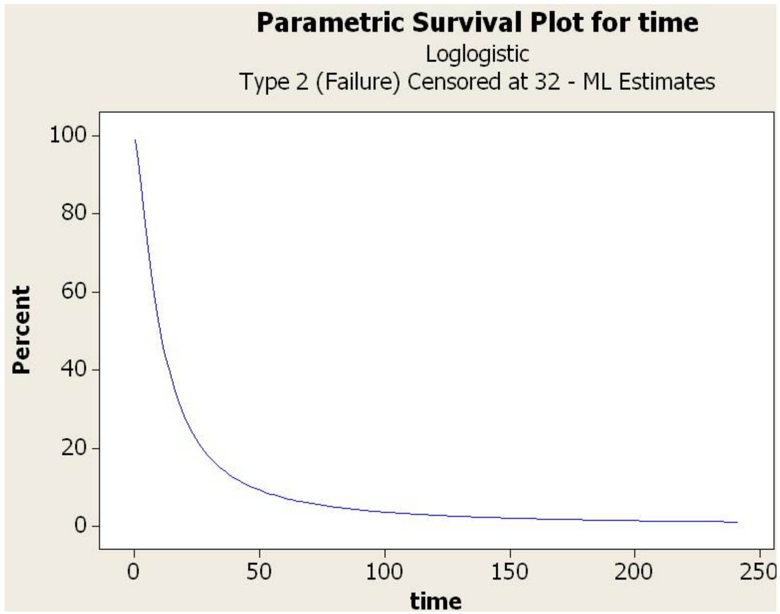
$$1,466 - 1,96(0,236) \leq \beta \leq 1,466 + 1,96(0,236)$$

sehingga diperoleh interval konfidensi untuk $\beta$ yaitu sebagai berikut :

$$1,00344 \leq \beta \leq 1,92856$$

Selanjutnya akan didapat nilai fungsi *survival* seorang pasien yang diintepretasikan sebagai probabilitas seorang pasien dapat bertahan hidup setelah melakukan operasi jantung pada waktu ke-t. Gambar 3.6 menunjukkan kurva fungsi *survival* pada data kematian pasien setelah melakukan operasi jantung.

Berdasarkan hasil yang diperoleh pada Gambar 3.6 dengan menggunakan *software MINITAB* 14 diketahui bahwa pada awal pengamatan karena belum ada pasien yang mengalami kematian maka probabilitas *survival* pada saat itu adalah 1, selanjutnya ketika periode pengamatan bertambah tanpa batas seiring dengan kesembuhan pasien maka probabilitas *survival* akan menurun dan konstan pada hari ke 250 yang berarti peluang hidup pasien konstan setelah hari ke 250 setelah diberi perlakuan yang pada akhirnya tidak ada satu pasienpun yang akan bertahan




Gambar 3.6: Kurva Fungsi *Survival* dari Data Kematian Pasien Setelah Operasi Jantung

hidup pada hari yang tidak dapat ditentukan.

Selanjutnya akan dihitung probabilitas *survival* seorang pasien setelah melakukan operasi jantung pada waktu ke-t menggunakan persamaan (2.16) yaitu

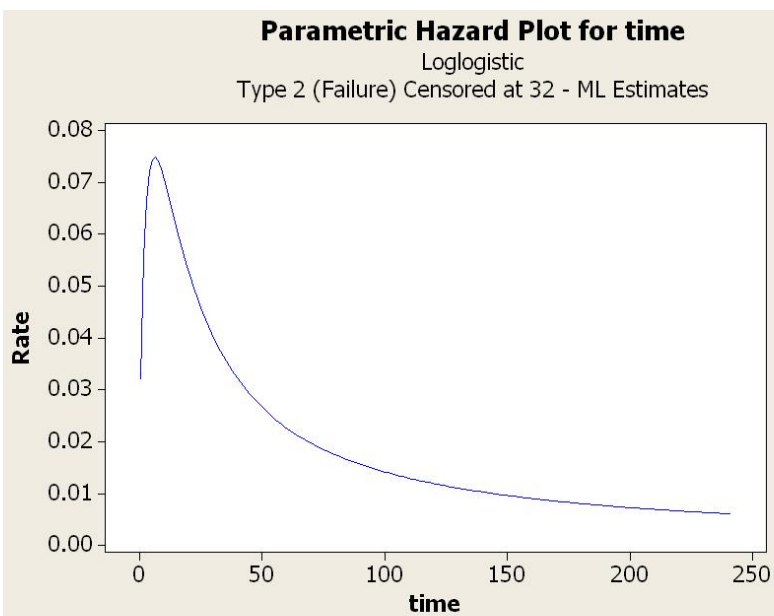
$$S(t_i) = \frac{1}{1 + \left[\frac{t_i}{10{,}435}\right]^{1{,}466}} \tag{3.31}$$

kemudian akan dihitung nilai fungsi *survival* selama 10 hari dengan menggunakan *software MAPLE* 13 yaitu

$$\begin{aligned} S(10) &= \frac{1}{1 + \left[\frac{10}{10{,}435}\right]^{1{,}466}} \\ S(10) &= 0,5156006686 \\ &\approx 0,52 \end{aligned} \tag{3.32}$$

Berdasarkan hasil S(10) pada persamaan (3.32) diketahui bahwa peluang seorang pasien setelah melakukan operasi jantung dan dapat bertahan hingga hari ke-10 adalah sebesar 0,52 atau peluang seorang pasien dapat bertahan hidup selama 10 hari ialah sebesar 52 % yang berarti peluang hidup pasien cukup rendah.

Selanjutnya untuk mengetahui probabilitas kematian pada waktu ke-t akan dihitung fungsi *hazard* yang diintepretasikan sebagai probabilitas kematian pada waktu ke-t. Berdasarkan hasil yang diperoleh pada Gambar 3.7 dengan menggunakan *software MINITAB* 14 diintepretasikan bahwa beberapa hari setelah dilakukan operasi, resiko kematian pada pasien semakin meningkat hingga mencapai puncak pada hari ke 13 namun setelah hari ke 13 seiring prosedur pasca operasi berlangsung maka resiko kematian akan terus berkurang seiring dengan kesembuhan pasien.




Gambar 3.7: Kurva Fungsi *Hazard* dari Data Kematian Pasien Setelah Operasi Jantung

Selanjutnya akan dihitung probabilitas kegagalan (kematian) seorang pasien pada waktu ke-t menggunakan persamaan (2.17) yaitu

$$h(t_i) = \frac{\left[\frac{1{,}466}{10{,}435}\right]\left[\frac{t_i}{10{,}435}\right]^{1{,}466-1}}{1+\left[\frac{t_i}{10{,}435}\right]^{1{,}466}} \tag{3.33}$$

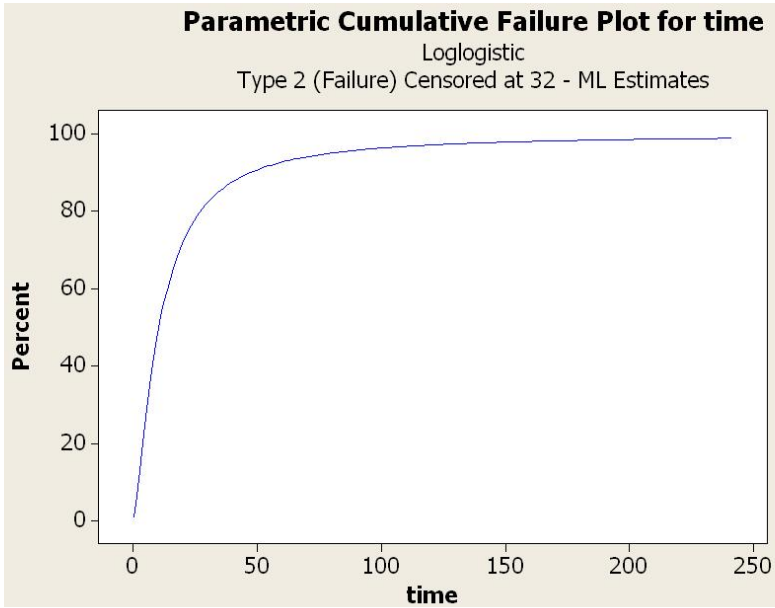
kemudian akan dihitung nilai fungsi *hazard* pada hari ke-10 hari yaitu

$$\begin{aligned} h(10) &= \frac{\left[\frac{1{,}466}{10{,}435}\right]\left[\frac{10}{10{,}435}\right]^{1{,}466-1}}{1+\left[\frac{10}{10{,}435}\right]^{1{,}466}} \\ h(10) &= 0,07101294198 \\ &\approx 0,072 \end{aligned} \tag{3.34}$$

dan berdasarkan hasil h(10) pada persamaan (3.34) dapat diketahui peluang kegagalan (kematian) seorang pasien setelah melakukan operasi pada hari ke-10 yaitu sebesar 0,072 atau peluang kematian pada hari ke-10 adalah sebesar 7,2 % ini berarti peluang kematian pasien cukuplah rendah setelah diberi perlakuan.

Sementara itu, untuk mengetahui probabilitas kematian seorang pasien setelah melakukan operasi jantung dari awal pengamatan hingga hari ke-t dapat dilihat pada Gambar 3.8 yang menyatakan grafik fungsi *hazard* kumulatif dengan menggunakan *software MINITAB* 14.

Berdasarkan hasil yang diperoleh pada Gambar 3.8 dapat diketahui bahwa pada awal pengamatan probabilitas kematian pasien adalah nol dimana pasien belum mengalami kematian. Selanjutnya ketika pengamatan bertambah hingga menuju hari ke 250, probabilitas kematian akan terus meningkat yang disebabkan karena menurunnya daya tahan pasien. Oleh karena itu, pada akhir pengamatan probabilitas kematian pasien adalah 100 % pada waktu yang tidak dapat diten-




Gambar 3.8: Kurva Fungsi *Hazard* Kumulatif dari Data Kematian Pasien Setelah Operasi Jantung

tukan.

Selanjutnya akan dihitung peluang kematian seorang pasien setelah melakukan operasi jantung dari awal pengamatan hingga waktu ke-t berdasarkan persamaan (2.21) yaitu

$$H(t_i) = ln\left|1 + (\frac{t_i}{10,435})^{1,466}\right| \tag{3.35}$$

kemudian akan dihitung nilai fungsi *hazard* kumulatif dari awal pengamatan hingga hari ke-10 yaitu

$$\begin{aligned} H(10) &= ln|1 + (\frac{10}{0,3660})^{1,274}| \\ H(10) &= 0,6624227111 \\ &\approx 0,67 \end{aligned} \tag{3.36}$$

dan berdasarkan hasil H(10) pada persamaan (3.36) dapat diketahui peluang kegagalan (kematian) seorang pasien setelah melakukan operasi jantung dari awal pengamatan hingga hari ke-10 yaitu sebesar 0,67 atau peluang kematian dari awal pengamatan hingga hari ke-10 setelah diberi perlakuan berupa operasi jantung ialah sebesar 67 % yang berarti peluang kematian dari awal pengamatan hingga hari ke-10 cukup tinggi pada pasien setelah melakukan operasi jantung.

Berdasarkan perhitungan yang telah dilakukan dengan menggunakan *software MAPLE* 13 pada persamaan (3.32), (3.34), dan (3.36) maka didapat tabel probabilitas untuk seorang pasien yang telah diberi perlakuan berupa operasi jantung dalam waktu 10 hari yaitu seperti berikut :

Tabel 3.1: Tabel Probabilitas Pasien Operasi Jantung

| Probabilitas | Percent (%) |
|---|---|
| S(10) | 52 |
| h(10) | 7,2 |
| H(10) | 67 |

Jadi berdasarkan tabel di atas dapat disimpulkan bahwa probabilitas *survival* seorang pasien dapat bertahan hidup selama 10 hari S(10) adalah sebesar 52 % yang berarti peluang hidup seorang pasien setelah melakukan operasi cukup rendah, probabilitas kematian seorang pasien pada hari ke-10 h(10) adalah sebesar 7,2 % yang berarti tingkat kematian pada pasien cukuplah rendah , dan probabilitas kematian dari awal pengamatan berlangsung hingga hari ke-10 H(10) adalah sebesar 67 % yang berarti tingkat kematian dari awal pengamatan hingga hari ke-10 cukuplah tinggi.

# BAB IV

# PENUTUP

## 4.1 Kesimpulan

1. Model *survival* untuk data tersensor kanan menggunakan distribusi Log-Logistik yaitu sebagai berikut :

$$L(\gamma,\beta)=\frac{n!}{(n-r)!}\left[(\frac{\beta}{\gamma})^r\prod_{i=1}^{r}\frac{(\frac{t_i}{\gamma})^{\beta-1}}{[1+(\frac{t_i}{\gamma})^{\beta}]^2}\right]\left[\frac{1}{1+(\frac{t_r}{\gamma})^{\beta}}\right]^{n-r}$$

2. Metode *maximum likelihood estimator* (MLE) digunakan untuk mengestimasi parameter $\gamma$ pada persamaan

$$-\frac{r\hat{\beta}}{\hat{\gamma}}+\left(\frac{2\hat{\beta}}{\hat{\gamma}}\right)\sum_{i=1}^{r}\left[1+\left(1+\left(\frac{t_i}{\hat{\gamma}}\right)\right)^{\hat{\beta}}\right]^{-1}\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}}-\frac{(r-n)}{1+\left(1+\left(\frac{t_r}{\hat{\gamma}}\right)\right)^{\hat{\beta}}}\left(\frac{\hat{\beta}}{\hat{\gamma}}\right)\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}}=0$$

dan parameter $\beta$ pada persamaan

$$\frac{r}{\hat{\beta}}+\sum_{i=1}^{r}ln\frac{t_i}{\hat{\gamma}}-2\sum_{i=1}^{r}\left[1+\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}}\right]^{-1}\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}}ln\left(\frac{t_i}{\hat{\gamma}}\right)+\frac{(r-n)}{1+\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}}}\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}}ln\left(\frac{t_r}{\hat{\gamma}}\right)=0$$

Oleh karena kedua persamaan tersebut terlalu sulit untuk diselesaikan secara eksplisit maka diperlukan metode Newton-Raphson untuk mendapatkan nilai estimasinya.

3. Aplikasi analisis *survival* untuk data tersensor kanan menggunakan distribusi Log-Logistik diterapkan pada pasien yang mengalami kematian setelah diberi perlakuan berupa operasi jantung dengan peluang hidup untuk seorang pasien selama 10 hari adalah sebesar 52 %, peluang kematian seorang pasien pada hari ke-10 ialah sebesar 7,2 %, dan peluang kematian dari awal pengamatan hingga hari ke-10 ialah sebesar 67 %.

## 4.2 Saran

Skripsi ini membahas mengenai data *survival* untuk data tersensor kanan menggunakan distribusi parametrik Log-Logistik pada contoh kasus pasien yang mengalami kematian setelah diberi perlakuan berupa operasi jantung di salah satu Rumah Sakit di Jakarta pada tahun 2014. Disarankan untuk menggunakan data tersensor jenis lain seperti data tersensor kiri dengan menggunakan distribusi parametrik lainnya seperti distribusi Weibull atau Eksponensial pada contoh kasus yang berbeda pula.

# DAFTAR PUSTAKA


Annahar,J. 2011. Analisis Data *Survival* Tersensor Kanan dengan Menggunakan Model *Hazard* Proposional. Skripsi. Jakarta : Universitas Negeri Jakarta.

Athoillah, Ibnu. 2012. Model Regresi Data Tahan Hidup Tersensor Tipe III Berdistribusi Log-Logistik. Jurnal. Semarang : Universitas Dipenogoro.

Bain, L.J and Engelhardt. 1992. *Log-Logistic Regression Models for Survival Data*. Applied Statistics, 32.165-171.

Bennet,S. 1983. *Introduction to Probability and Mathematical Statistics*. 2nd ed. California : Duxbury Press.

Collect, David. 2003. *Modelling Survival Data in Medical Research*, 2nd edition. London : Chapman and Hall.

Conkin, Johnny, Houston and Lyndon B. Johnson. 2001. *A Log Logistic Survival Model Applied to Hypobaric Decompression Sickness*. Texas : NASA/TP-2001-210775.

Dixit, Asha. 2008. *Exact Comparison of Hazard Rate Functions of Log-Logistic Survival Distributions* [Thesis]. Alabama : Auburn University.

Kirana, RR Mudita Candra, 2014. Perluasan Model Regresi *Cox Hazard* Non-Proposional. Skripsi. Jakarta : Universitas Negeri Jakarta.

Kleinbaum, D.G, and Klein, M. 2005. *Survival Analysis - A Self Learning Text*, Second edition. New York : Springer.

Lawless, J.F. 2003. *Statistical Model and Methods for Lifetime Data*. New York : John Willey and Sons,Icn.

Lee, E.T. 2003. *Statistical Methods for Survival Data Analysis*. 3rd ed. New York : Wiley.

Machin, David, Yin Bun C and Mahesh Parmar. 2006. *Survival Analysis a Practical Approach*. 2nd edition. Chicester : John Wiley and Sons Ltd.

Olapade,A.K and M. O. Ojo. 2003. *On The Generalized Logistic and Log-Logisttic Distribution*. Nigeria : Obafemi Awolowo University.

Rao, G.S Kantam and K. Rosaih. 2009. *Reliability Estimation in Log-Logistic Distribution from Censored Sampels*, Prob. Stat.

Tirta, I.M. 2009. Analisis Regresi dengan R. Jember : *Jember University Press.*

Walpole, Ronald E. 1993. Pengantar Statistika Edisi ke-3. Jakarta : Gramedia Pustaka Umum.

Wardani, Purwita. 2010. Penaksiran Fungsi *Survival* dengan Metode Nonparametrik dan Parametrik. Skripsi. Depok : Universitas Indonesia.

# LAMPIRAN

## LAMPIRAN 1

Tabel 4.1: DATA KEMATIAN PASIEN SETELAH DILAKUKAN OPERASI JANTUNG DI SALAH SATU RUMAH SAKIT DI JAKARTA PADA TAHUN 2014

| Pasien | Waktu Kematian (hari) | Sensor |
|---|---|---|
| 1 | 2 | 0 |
| 2 | 1 | 0 |
| 3 | 1 | 0 |
| 4 | 4 | 0 |
| 5 | 5 | 0 |
| 6 | 10 | 0 |
| 7 | 14 | 0 |
| 8 | 21 | 1 |
| 9 | 7 | 0 |
| 10 | 9 | 0 |
| 11 | 30 | 1 |
| 12 | 11 | 0 |
| 13 | 14 | 0 |
| 14 | 44 | 1 |
| 15 | 15 | 0 |
| 16 | 4 | 0 |
| 17 | 6 | 0 |
| 18 | 11 | 0 |
| 19 | 13 | 0 |
| 20 | 9 | 0 |
| 21 | 15 | 0 |
| 22 | 31 | 1 |
| 23 | 18 | 1 |
| 24 | 49 | 1 |
| 25 | 9 | 0 |

| Pasien | Waktu Kematian (hari) | Sensor |
|---|---|---|
| 26 | 15 | 0 |
| 27 | 7 | 0 |
| 28 | 3 | 0 |
| 29 | 10 | 0 |
| 30 | 5 | 0 |
| 31 | 18 | 1 |
| 32 | 45 | 1 |
| 33 | 1 | 0 |
| 34 | 12 | 0 |
| 35 | 7 | 0 |
| 36 | 28 | 1 |
| 37 | 4 | 0 |
| 38 | 41 | 1 |
| 39 | 12 | 0 |
| 40 | 10 | 0 |

0 = sensor

1 = tidak tersensor

# LAMPIRAN 2

## Distribution Analysis: time

Variable: time

| Censoring Information | Count |
|---|---|
| Uncensored value | 31 |
| Right censored value | 9 |

Type 2 (Failure) Censored at 32

Estimation Method: Maximum Likelihood

Distribution: Loglogistic

Parameter Estimates

| | | Standard | 95.0% Normal CI | |
|---|---|---|---|---|
| Parameter | Estimate | Error | Lower | Upper |
| Location | 2.34951 | 0.185669 | 1.98560 | 2.71341 |
| Scale | 0.682343 | 0.105344 | 0.504184 | 0.923456 |

Log-Likelihood = -116.946

Goodness-of-Fit
Anderson-Darling (adjusted) = 44.609

# LAMPIRAN 3

$$S(10) = \frac{1}{1 + \left(\frac{10}{10.435}\right)^{1.466}}$$

$$S(10) = 0.5156006686$$

$$h(10) = \frac{\left(\frac{1.466}{10.435}\right) \cdot \left(\frac{10}{10.435}\right)^{1.466 - 1}}{1 + \left(\frac{10}{10.435}\right)^{1.466}}$$

$$h(10) = 0.07101294198$$

$$H(10) = \ln\left(1 + \left(\frac{10}{10.435}\right)^{1.466}\right)$$

$$H(10) = 0.6624227111$$

**SURAT PERNYATAAN KEASLIAN SKRIPSI**

Dengan ini saya yang bertanda tangan di bawah ini, mahasiswa Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta:

| | | |
|---|---|---|
| Nama | : | Debi Oktaviani |
| No. Registrasi | : | 3125110491 |
| Jurusan | : | Matematika |
| Program Studi | : | Matematika |

Menyatakan bahwa skripsi ini yang saya buat dengan judul "**Analisis *Survival* untuk Data Tersensor Menggunakan Distribusi Log-Logistik**" adalah :

1. Dibuat dan diselesaikan oleh saya sendiri.
2. Bukan merupakan duplikat skripsi yang pernah dibuat oleh orang lain atau jiplakan karya tulis orang lain.

Pernyataan ini dibuat dengan sesungguhnya dan saya bersedia menanggung segala akibat yang timbul jika pernyataan saya tidak benar.

Jakarta, Juli 2015

Yang membuat pernyataan

**Debi Oktaviani**

## DAFTAR RIWAYAT HIDUP

**DEBI OKTAVIANI.** Lahir di Jakarta, 3 Oktober 1993. Anak kedua dari pasangan Bapak Budi Darmanto dan Ibu Dewi Meilyasari. Saat ini bertempat tinggal di Jalan Swakarsa I B No. 36 A RT 04 RW 03, Jakarta Timur 13450.

| | | |
|---|---|---|
| No. Ponsel | : | 089602771310 |
| Email | : | debioktaviani3@ymail.com |

**Riwayat Pendidikan** : Penulis mengawali pendidikan di TK Indah Dahlia selama 2 tahun, dan kemudian melanjutkan pendidikan di SDN 02 Pagi Jakarta pada tahun 1999 - 2005. Setelah itu, penulis melanjutkan ke SMPN 27 Jakarta hingga tahun 2008. Kemudian kembali melanjutkan ke SMA Negeri 12 Jakarta dan lulus tahun 2011. Di Tahun yang sama penulis melanjutkan ke Universitas Negeri Jakarta (UNJ), jurusan Matematika, melalui jalur SNMPTN Undangan. Di pertengahan tahun 2015 penulis telah memperoleh gelar Sarjana Sains untuk Jurusan Matematika, Program Studi Matematika, FMIPA, UNJ.

**Riwayat Organisasi** : Selama di bangku perkuliahan, penulis aktif di berbagai organisasi kemahasiswaaan. Pada tahun 2012, penulis mendapat kepercayaan sebagai staff LLMJ (Lembaga Legislatif Mahasiswa Jurusan) Matematika sebagai pengawas Departemen Advokasi dan Olah raga. Tahun 2012 pula, penulis aktif dalam kegiatan mahasiswa UKM (Unit Kesenian Mahasiswa) Universitas Negeri Jakarta.