

ANALISIS SURVIVAL DAN HAZARD UNTUK DATA TERSENSOR TIPE II MENGGUNAKAN MODEL DISTRIBUSI LOG-LOGISTIK

Farida*¹, Alya Rahma Inneztiana², Muhammad Falah El Fahmi³, Toha Saifudin⁴

^{1,2,3,4} Program Studi Statistika, Departemen Matematika, Fakultas Sains dan Teknologi, Universitas Airlangga.

farida-2020@fst.unair.ac.id¹, alya.rahma.inneztiana-2020@fst.unair.ac.id²,
muhammad.falah.el-2020@fst.unair.ac.id³, tohasaifudin@fst.unair.ac.id⁴

**Corresponding Author*

Received 21 April 2023; revised 22 Mei 2023; accepted 15 Juli 2023.

ABSTRAK

Statistika merupakan salah satu bidang ilmu yang sudah berkembang begitu jauh dalam analisis untuk berbagai keperluan inferensi, estimasi, pengujian dan metode peramalan. Fungsi distribusi tahan hidup yang didasarkan pada pengetahuan atau asumsi tertentu tentang distribusi populasinya termasuk dalam fungsi parametrik. Penelitian ini menggunakan fungsi survival dan hazard berdistribusi Log-logistik, dimana data uji hidup diasumsikan mengikuti Distribusi Log-logistik. Distribusi Log-logistik cukup jarang digunakan dalam analisis survival. Distribusi Log-logistik mempunyai bentuk yang hampir sama dengan Distribusi Log-normal. Fungsi survival dan hazard dapat digunakan untuk meneliti tahan hidup pasien yang terserang penyakit kronis, seperti kanker. Data yang digunakan dalam penelitian ini adalah 50 data umur penyakit hingga pasien meninggal karena penyakit kanker paru. Dari pengamatan yang ada, hanya diambil 30 hasil pengamatan pertama. Banyaknya pengamatan yang diambil telah ditentukan sebelum penelitian dilakukan. Pada data ini hanya diambil 30 pengamatan, sehingga terdapat 20 pengamatan yang tersensor. Model *survival* data tersensor tipe II model distribusi loglogistic $S(t_i) =$

$$\left[1 + \left(\frac{t_i}{69,268}\right)^{3,122}\right]^{-1}, \text{ Model } \textit{hazard} \text{ data tersensor tipe II model distribusi loglogistic}$$

$$h(t_i) = \frac{0,045071\left(\frac{t_i}{69,268}\right)^{2,122}}{\left[1 + \left(\frac{t_i}{69,268}\right)^{3,122}\right]}. \text{ Peluang seorang pasien untuk hidup jika menderita}$$

kanker paru-paru selama 48, 60, dan 72 bulan berturut-turut adalah 0,7386; 0,6103; dan 0,4698.

Kata kunci: kanker paru, log logistik, survivor dan hazard, tersensor tipe II.

ABSTRACT

Statistics is a field of science that has developed so far in analysis for various purposes of inference, estimation, testing and forecasting methods. The survival distribution function which is based on certain knowledge or assumptions about the population distribution is included in the parametric function. This study uses the survival and hazard functions with a log-logistic distribution, where the survival test data is assumed to follow a log-logistic distribution. The log-logistic distribution is rarely used in survival analysis. The log-logistic distribution has almost the same form as the log-normal distribution. The survival and hazard functions can be used to study the survival of patients with chronic diseases, such as cancer. The data used in this study were 50 data on the age of the disease until the patient died from lung cancer. From the existing observations, only the first 30 observations were taken. The number of observations taken has been determined before the research was conducted. In this data only 30 observations were taken, so there were 20 censored observations. Type II censored data survival model loglogistic distribution model $S(t_i) = \left[1 + \left(\frac{t_i}{69,268} \right)^{3,122} \right]^{-1}$, Type II censored data hazard model loglogistic distribution model $h(t_i) = \frac{0,045071 \left(\frac{t_i}{69,268} \right)^{2,122}}{\left[1 + \left(\frac{t_i}{69,268} \right)^{3,122} \right]}$. The probability of a patient surviving lung cancer for 48, 60 and 72 consecutive months is 0.7386; 0.6103; and 0.4698.

Keywords: lung cancer, logistic log, survivor and hazard, type II censored.

PENDAHULUAN

Statistika merupakan salah satu bidang ilmu yang sudah berkembang begitu jauh dalam analisis untuk berbagai keperluan inferensi, estimasi, pengujian dan metode peramalan. Fungsi distribusi tahan hidup yang didasarkan pada pengetahuan atau asumsi tertentu tentang distribusi populasinya termasuk dalam fungsi parametrik. Analisis uji hidup untuk data tersensor tipe II adalah suatu data waktu kematian atau kegagalan yang hanya terdapat r buah pengamatan terkecil dalam sampel acak yang berukuran n dengan $1 \leq r \leq n$. Eksperimen menggunakan penyensoran tipe II lebih sering digunakan, misalnya dalam uji hidup dari total observasi

sebanyak n , tetapi uji akan berhenti pada waktu observasi sampel mempunyai waktu kematian atau kegagalan ke- r untuk $1 \leq r \leq n$.

Penelitian dalam berbagai bidang Teknik, biologi, pertanian, kedokteran dan lainnya berkaitan dengan data uji hidup. Analisis yang digunakan untuk menganalisis data waktu hidup tersebut disebut analisis fungsi survival dan fungsi hazard. Beberapa distribusi yang dapat digunakan untuk menggambarkan waktu hidup antara lain Distribusi Eksponensial, Distribusi Weibull, Distribusi Gamma, Distribusi Rayleigh, dan lain-lain (Lawless, 1982: 26).

Penelitian ini menggunakan fungsi survival dan hazard berdistribusi Log-logistik, dimana data uji hidup diasumsikan mengikuti Distribusi Log-logistik. Distribusi Log-logistik cukup jarang digunakan dalam analisis survival. Distribusi Log-logistik mempunyai bentuk yang hampir sama dengan Distribusi Log-normal. Fungsi survival dan hazard dapat digunakan untuk meneliti tahan hidup pasien yang terserang penyakit kronis, seperti kanker.

METODE PENELITIAN

Data yang digunakan dalam penelitian ini merupakan data sekunder yang bersumber dari skripsi karya Dwi Retno Sari dari Universitas Negeri Yogyakarta yang berjudul *Analisis Survival untuk Data Tersensor Tipe II Menggunakan Model Distribusi Log-Logistik*.

Penelitian ini menggunakan metode penelitian kuantitatif, yakni dengan menggunakan model *survival* dan *hazard* pada distribusi parametrik log-logistik. Adapun langkah-langkah analisis data dalam penelitian ini antara lain:

1. Melakukan sensor data dengan analisis data tersensor tipe II pada data umur penyakit pasien penderita kanker paru-paru
2. Melakukan analisis statistika deskriptif
3. Melakukan analisis untuk fungsi densitas peluang, fungsi *hazard*, dan fungsi *survival* dengan bantuan *software* Minitab
4. Menghitung peluang seorang pasien untuk hidup jika menderita penyakit kanker paru-paru dengan menggunakan fungsi *survival*

5. Pengambilan kesimpulan

HASIL PENELITIAN DAN PEMBAHASAN

Bedasarkan data yang diperoleh pada lampiran 1, maka diperoleh gambaran umum 30 dari 50 data, dimana 20 data tersensor untuk 30 pengamatan data umur penyakit pasien (bulan).

Tabel 4.Error! No text of specified style in document..1 Deskripsi Data Umur Penyakit Pasien (bulan)

<i>Variable</i>	<i>N</i>	<i>N*</i>	<i>Mean</i>	<i>SE Mean</i>	<i>StDev</i>	<i>Minimum</i>	<i>Median</i>	<i>Maximum</i>
Data	30	0	51,69	3,08	16,89	20,84	51,13	74,32

Tabel 4.1 menunjukan bahwa umur penyakit pasien (bulan) adalah 51,69 Bulan . umur terendah adalah 20,84 bulan, Sedangkan, umur tertinggi adalah 74,32 bulan.

Data tersensor tipe II adalah suatu data waktu kematian atau waktu tahan hidup yang hanya terdapat r buah observasi dalam sampel random yang berukuran n dengan $1 \leq r \leq n$. Maka akan dibahas Fungsi Survival dan Fungsi Hazard dari data tersensor tipe II berdasarkan model distribusi log-logistik.

Menurut Lawless (2003), misalkan T merupakan variabel random tahan hidup dengan T adalah variabel random kontinu non negatif yang menunjukkan tahan hidup individu-individu dalam suatu populasi yang berdistribusi Log-logistik dengan fungsi densitas peluang yaitu :

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

Fungsi Survival dari data tersensor tipe II berdasarkan model distribusi log-logistik yaitu :

$$S(t; \gamma, \beta) = \left[1 + \left(\frac{t}{\gamma} \right)^{\beta} \right]^{-1}$$

Fungsi Hazard dari data tersensor tipe II berdasarkan model distribusi log-logistik yaitu :

$$h(t; \gamma, \beta) = \frac{\left(\frac{\beta}{\gamma} \right) \left(\frac{t}{\gamma} \right)^{\beta-1}}{\left[1 + \left(\frac{t}{\gamma} \right)^{\beta} \right]}$$

Berikut merupakan 50 data umur penyakit hingga pasien meninggal karena penyakit kanker paru. Dari pengamatan yang ada, hanya diambil 30 hasil pengamatan pertama. Banyaknya pengamatan yang diambil telah ditentukan sebelum penelitian dilakukan. Pada data ini hanya diambil 30 pengamatan, sehingga terdapat 20 pengamatan yang tersensor.

Tabel 4.2 Data Umur Penyakit Pasien (Bulan)

20,835	42,985	64,562	75,121*	78,343*
23,364	43,959	65,448	75,321*	79,121*
27,957	45,787	67,540	76,111*	79,173*
30,830	46,238	69,055	76,345*	79,345*
31,395	49,080	71,817	77,115*	80,349*
33,600	53,179	72,157	77,345*	81,134*
36,917	56,004	72,896	78,000*	81,456*
37,794	59,128	72,992	78,010*	82,435*
40,327	61,168	73,202	78,121*	83,123*
41,869	64,449	74,316	78,233*	85,123*

*: menandakan bahwa data tersensor

Untuk mempermudah perhitungan fungsi Survivar dan Hazard, dapat menggunakan software Minitab. Dalam makalh ini menggunakan software Minitab18. Dalam software ini, fungsi densitas peluang yang digunakan adalah

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

Dengan $\mu = \text{location parameter}$

$\sigma = \text{Scale parameter}$

Tabel 4.3 Parameter Lokasi dan Skala

Parameter	Estimate	Standard	95,0% Normal CI	
		Error	Lower	Upper
Location	4,23798	0,0830154	4,07527	4,40069
Scale	0,320325	0,0500249	0,235862	0,435033

Dari Output minitab tersebut diperoleh Parameter Location 4,23798 dan Scale parameter adalah 0,320325. Sehingga diperoleh :

$$\hat{\gamma} = \exp (\text{location parameter})$$

$$\hat{\gamma} = e^{\mu}$$

$$\hat{\gamma} = e^{4,23798}$$

$$\hat{\gamma} = 69,268$$

$$\hat{\beta} = \frac{1}{\text{Scale parameter}}$$

$$\hat{\beta} = \frac{1}{0,320325}$$

$$\hat{\beta} = 3,122$$

Berikut ini adalah bentuk fungsi densitas peluang dari data, sedangkan bentuk kurvanya ditunjukkan pada gambar 4.1:

Analisis Survival dan Hazard Untuk Data Tersensor Tipe II Menggunakan Model Distribusi Log-Logistik

$$f(t_i) = \frac{\left(\frac{3,122}{69,268}\right) \left(\frac{t_i}{69,268}\right)^{3,122-1}}{\left[1 + \left(\frac{t_i}{69,268}\right)^{3,122}\right]^2}$$

$$f(t_i) = \frac{0,045071 \left(\frac{t_i}{69,268}\right)^{2,122}}{\left[1 + \left(\frac{t_i}{69,268}\right)^{3,122}\right]^2}$$

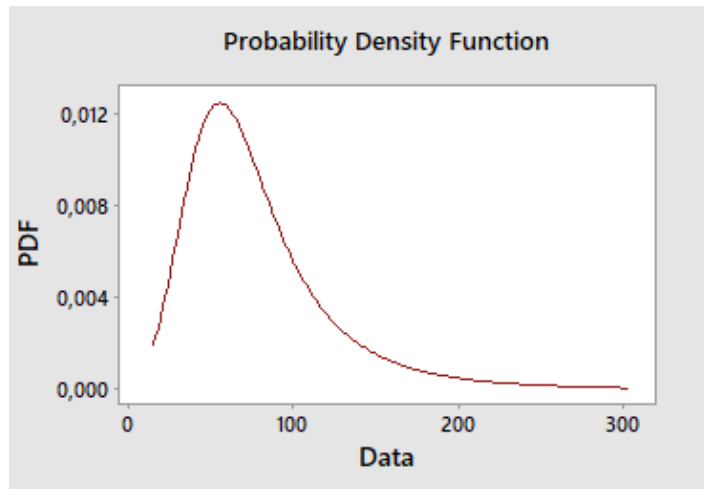


Figure 4. **Error! No text of specified style in document.**1 Kurva fungsi densitas peluang dari data umur penyakit pasien

Bentuk fungsi hazard ditunjukkan pada persamaan sebelumnya dan kurva fungsi hazard dari data pada gambar 4.2.

$$h(t_i) = \frac{\left(\frac{3,122}{69,268}\right) \left(\frac{t_i}{69,268}\right)^{3,122-1}}{\left[1 + \left(\frac{t_i}{69,268}\right)^{3,122}\right]}$$

$$h(t_i) = \frac{0,045071 \left(\frac{t_i}{69,268}\right)^{2,122}}{\left[1 + \left(\frac{t_i}{69,268}\right)^{3,122}\right]}$$

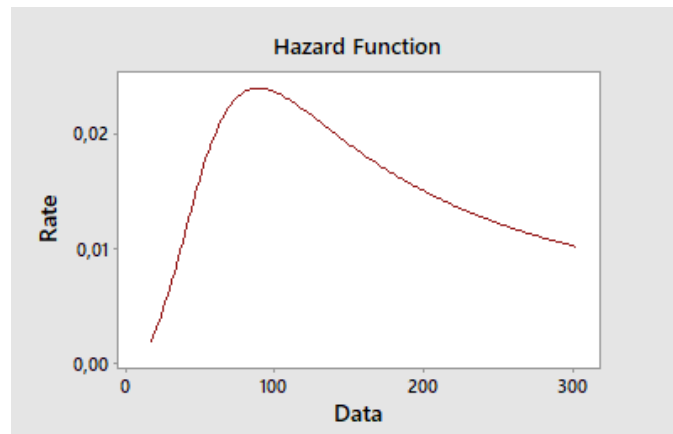


Figure 4. *Error! No text of specified style in document.*2 Kurva fungsi hazard dari data umur penyakit pasien

Bentuk fungsi survivor ditunjukkan pada persamaan sebelumnya dan kurva fungsi survivor dari data pada gambar 4.3.

$$S(t_i) = \left[1 + \left(\frac{t_i}{69,268} \right)^{3,122} \right]^{-1}$$

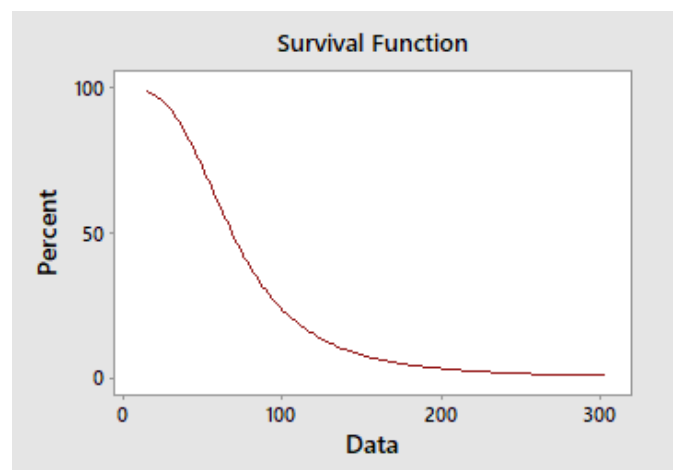


Figure 4.3 Kurva fungsi survivor dari data umur penyakit pasien

Jika fungsi survivor telah didapatkan, maka dapat dihitung peluang seorang pasien untuk hidup jika menderita penyakit kanker paru-paru selama 48, 60 dan 72 bulan.

$$S(48) = \left[1 + \left(\frac{48}{69,268} \right)^{3,122} \right]^{-1}$$

$$S(48) = 0.7586$$

$$S(60) = \left[1 + \left(\frac{60}{69,268} \right)^{3,122} \right]^{-1}$$

$$S(60) = 0.6103$$

$$S(72) = \left[1 + \left(\frac{72}{69,268} \right)^{3,122} \right]^{-1}$$

$$S(72) = 0.4698$$

Jadi peluang seorang pasien untuk hidup jika menderita kanker paru-paru selama 48 bulan atau 4 tahun bulan adalah 0,7386. Peluang pasien untuk hidup jika menderita kanker paru-paru selama 60 bulan atau 5 tahun adalah 0.6103. Peluang pasien untuk hidup jika menderita kanker paru-paru selama 72 bulan atau 6 tahun adalah 0.4698.

SIMPULAN

Model *survival* data tersensor tipe II model distribusi loglogistic $S(t_i) =$

$\left[1 + \left(\frac{t_i}{69,268} \right)^{3,122} \right]^{-1}$, Model *hazard* data tersensor tipe II model distribusi

loglogistic $h(t_i) = \frac{0,045071 \left(\frac{t_i}{69,268} \right)^{2,122}}{\left[1 + \left(\frac{t_i}{69,268} \right)^{3,122} \right]}$. Peluang seorang pasien untuk hidup jika

menderita kanker paru-paru selama 48, 60, dan 72 bulan berturut-turut adalah 0,7386; 0,6103; dan 0,4698.

UCAPAN TERIMA KASIH

Peneliti berterima kasih kepada Dr. Toha Saifudin, S.Si., M.Si. selaku dosen pengampu mata kuliah analisis data uji hidup dalam membimbing, memberikan saran dan masukan untuk penyusunan artikel ilmiah ini.

DAFTAR PUSTAKA

- Bain, L. J., & Engelhardt, M. (1992). *Introduction to probability and mathematical statistics* (Vol. 4). Belmont, CA: Duxbury Press.
- Dixit, A. (2008). *Exact Comparision of Hazard Rate Functions of Log-Logistic Survival Distributions* (Doctoral dissertation).
- Harlan, J. (2017). Analisis Survival. Depok: Gunadarma.
- Lambert, P., Collett, D., Kimber, A., & Johnson, R. (2004). Parametric accelerated failure time models with random effects and an application to kidney transplant survival. *Statistics in medicine*, 23(20), 3177-3192.
- Lawless, J. F. (1982). *Statistical Model and Method for Lifetime Data*. New York: John Willey and Sons.
- Lawless, J. F. (2011). *Statistical models and methods for lifetime data*. John Wiley & Sons.
- Machin, D., Cheung, Y. B., & Parmar, M. (2006). *Survival analysis: a practical approach*. John Wiley & Sons.
- Rao, G. S., Kantam, R. R. L., & Rosaiah, K. (2009, July). Reliability estimation in log-logistic distribution from censored samples. In *Probstat forum* (Vol. 2, pp. 52-67).
- Sari, D. R. (2011). Analisis Survival untuk Data Tersensor Tipe II Menggunakan Model Distribusi Log-Logistik. *Tugas Akhir*.
- Walpole, R. E. (1995). Pengantar Statistika edisi ke-3. *Jakarta: Gramedia Pustaka Utama*.