JSA Jurnal Statistika dan Aplikasinya
Volume 2 Issue 2,
December 2018
e-ISSN: 2620-8369

Received: 7 November 2018
Revised: 13 November 2018
Accepted: 30 November 2018
Published: 30 December 2018

Analisis *Survival* dengan Model Regresi pada Data Tersensor Berdistribusi Log-logistik

Gatri Eka Kusumawardhani[1, a)], Suyono[2, b)], Vera Maya Santi[2, c)]

[1]*Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta*
[2]*Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta*

Email: [a)]gatrieka82@gmail.com, [b)]synjkt@yahoo.com, [c)]vera.indr4@gmail.com

Abstract

Survival analysis is an analysis used to determine the length of time required by an object in order to survive. That time is influenced by several factors called independent variables. One way to know relationship is through a regression model. The dependent variable is a survival time which is log-logistic distributed. The data used in this study were right censored. Log-logistic regression models for survival data can be expressed by transformation Y=lnT. The parameter of the log-logistic regression models for right censored survival data are estimated with the maximum likelihood method. In this study, the application of log-logistic regression model for survival data is in data of lung cancer patients. Best log-logistic regression model is obtained $y_i=1.92458+0.0242393\, x_{i1}+0.639037\varepsilon_i$.

Keywords*: Regression, right censored data, maximum likelihood.*

Abstrak

Analisis *survival* merupakan analisis yang digunakan untuk mengetahui lama waktu yang dibutuhkan oleh suatu objek untuk dapat bertahan hidup. Waktu tersebut dipengaruhi oleh beberapa faktor yang disebut dengan variabel independen. Salah satu cara untuk mengetahui hubungannya adalah melalui model regresi. Variabel dependen merupakan waktu *survival* yang berdistribusi log-logistik. Data yang digunakan dalam penelitian ini adalah data tersensor kanan. Model regresi log-logistik untuk data *survival* dapat dinyatakan dengan transformasi dengan Y=lnT. Estimasi parameter dari model regresi log-logistik untuk data tahan hidup tersensor kanan menggunakan metode maksimum *likelihood*. Penerapan model regresi log-logistik untuk data *survival* yang digunakan dalam penelitian ini adalah data pasien kanker paru-paru. Berdasarkan olah data yang sudah dilakukan, model regresi log-logistik terbaik $y_i=1.92458+0.0242393\, x_{i1}+0.639037\varepsilon_i$.

Kata-kata kunci: Regresi, Data tersensor kanan, metode maksimum *likelihood*.

PENDAHULUAN

Analisis *survival* adalah analisis yang berfungsi untuk mengetahui ketahanan hidup suatu objek yang diteliti. Variabel yang diperhatikan dalam analisis *survival* adalah waktu sampai terjadinya kejadian khusus. Kumpulan data yang digunakan dapat berupa data tersensor. Data tersensor adalah hasil pengamatan dari objek yang diteliti pada kurun waktu tertentu dan tidak mengalami kegagalan hingga penelitian berakhir. Analisis *survival* melibatkan bentuk data *time to event* atau dapat disebut data *survival*, yaitu suatu hasil pengamatan yang dilakukan pada kurun waktu tertentu sampai terjadinya suatu kejadian. Kumpulan data yang digunakan dapat berupa data lengkap atau tersensor.

Analisis *survival* akan sesuai jika digunakan pada data lengkap. Tetapi, pada kenyataannya terdapat kendala pada suatu penelitian seperti keterbatasan dana dan waktu. Dalam hal ini, diperlukan penyensoran data agar lebih efisien dari segi biaya dan waktu. Jenis penyensoran yang dapat digunakan untuk mempercepat penelitian adalah tersensor kanan dan tersensor kiri. Tersensor kanan terjadi bila objek belum mengalami kegagalan sampai penelitian berakhir, sedangkan tersensor kiri terjadi bila kegagalan pada objek sudah terjadi saat objek masuk dalam pengamatan. Analisis data *survival* dengan data tersensor diperlukan asumsi tertentu tentang distribusi populasinya.

Beberapa distribusi yang dapat digunakan untuk menggambarkan waktu *survival* antara lain distribusi eksponensial, distribusi Weibull, distribusi gamma, distribusi log-logistik, dan lain-lain (Lawless, 1982). Dalam penulisan ini, hanya mengkaji fungsi *survival* berdistribusi log-logistik. Distribusi log-logistik memiliki kelebihan yang tidak dimiliki oleh distribusi lain, yaitu dapat digunakan pada analisis *survival* penyakit. Waktu tahan hidup tiap individu yang berdistribusi log-logistik dapat dipengaruhi oleh beberapa faktor atau variabel independen. Faktor-faktor tersebut dilihat pada setiap individu saat pengamatan, misalnya umur, jenis kelamin, t*reatment* dan lainnya. Maka, pada penulisan ini akan dikaji ulang tentang analisis *survival* dengan model regresi pada data tersensor kanan berdistribusi log-logistik.

METODE

Analisis Survival

Analisis *survival* merupakan prosedur statistika untuk keperluan analisis data, dimana data yang digunakan berupa waktu sampai terjadinya suatu kejadian tertentu. Kleinbaum (2005) mengemukakan bahwa ada tiga elemen yang perlu diperhatikan dalam menentukan waktu *survival*, yaitu titik awal yaitu waktu dimulainya suatu penelitian, kejadian yang menjadi inti dari penelitian, dan ukuran waktu yang digunakan. Misalkan T adalah variabel acak yang menunjukkan waktu *survival* maka T bernilai nonnegatif, T > 0.

Kuantitas Dasar Analisis Survival

Analisis *survival* ditunjang oleh beberapa kuantitas dasar yaitu sebagai berikut

1. Fungsi Kepadatan Peluang
   Fungsi kepadatan peluang adalah peluang kegagalan individu pada suatu interval (t, t+Δt) persatuan waktu. Fungsi kepadatan peluang dinyatakan dengan f(t)
   $$f(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq (t+\Delta t))}{\Delta t} \tag{1}$$
2. Fungsi *Survival*
   Fungsi *survival* adalah peluang suatu individu yang masih dapat bertahan hidup sampai dengan waktu t, dimana t > 0. Fungsi *survival* dapat dinyatakan sebagai
   $$S(t) = P(T > t) \tag{2}$$
3. Fungsi *Hazard*
   Fungsi *hazard* adalah *rate* suatu individu mengalami kegagalan dalam interval waktu t sampai (t, t+Δt), jika individu tersebut masih bertahan hidup sampai dengan waktu t.
   $$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t+\Delta t \mid T > t)}{\Delta t} \tag{3}$$

Data Tersensor

Data dikatakan tersensor jika pengamatan waktu tahan hidup hanya sebagian, tidak sampai waktu kejadian. Ada tiga tipe penyensoran yang dapat digunakan, yaitu

1. Sensor Kanan
   Dikatakan tersensor kanan pada titik k, apabila nilai observasi yang digunakan t, dan t ≤ k.
2. Sensor Kiri

Suatu data atau observasi dikatakan tersensor kiri pada titik k, apabila nilai observasi yang digunakan adalah t, dan $t \geq k$.

3. Sensor Interval
Sensor interval terjadi ketika hanya diketahui bahwa suatu event yang diinginkan terjadi dalam suatu periode waktu. Data tersensor kiri dan tersensor kanan merupakan kasus khusus dari data tersensor interval.

Distribusi Log-logistik

Distribusi log-logistik adalah distribusi peluang dari variabel acak yang logaritmanya memiliki distribusi logistik. Suatu variabel acak Y dikatakan memiliki distribusi logistik dengan parameter (μ,σ), jika fungsi kepadatan peluangnya sebagai berikut

$$f(y) = \frac{\exp(\frac{y-\mu}{\sigma})}{\sigma(1+\exp(\frac{y-\mu}{\sigma}))^2} \tag{4}$$

Variabel acak T dikatakan mengikuti distribusi log-logistik dengan parameter β dan parameter bentuk α, jika mempunyai fungsi densitas

$$f(t) = \frac{(\beta/\alpha)(t/\alpha)^{\beta-1}}{\{1+(t/\alpha)^{\beta}\}^2} \tag{5}$$

Fungsi distribusi kumulatif dari distribusi log-logistik adalah

$$F(t) = \frac{t^{\beta}}{\alpha^{\beta}+t^{\beta}} \tag{6}$$

Fungsi *survival* dari distribusi log-logistik adalah

$$S(t) = \frac{1}{1+(t/\alpha)^{\beta}} \tag{7}$$

Fungsi *hazard* dari distribusi log-logistik adalah

$$h(t) = \frac{(\beta/\alpha)(t/\alpha)^{\beta-1}}{\{1+(t/\alpha)^{\beta}\}} \tag{8}$$

Model Regresi Data *Survival*

Model regresi digunakan pada analisis *survival* untuk memodelkan hubungan antara waktu *survival* sebagai variabel dependen dengan faktor-faktor yang berpengaruh sebagai variabel independen, dengan asumsi bahwa populasi tahan hidup yang diteliti berdistribusi tertentu. Menurut Lawless (1982), persamaan regresi parametrik pada analisis *survival* dinyatakan dalam model skala, yaitu waktu *survival* T ditransformasikan logaritma sehingga menjadi Y=ln T, dan diperoleh persamaan regresi berikut

$$Y = \ln T = \mu(x) + \sigma\varepsilon \tag{9}$$

Metode Maksimum *Likelihood*

Metode untuk mengestimasi nilai parameter distribusi dari data dalam fungsi tahan hidup adalah dengan menggunakan metode maksimum *likelihood*. Misalkan $X_1, X_2, \ldots, X_n$ adalah variabel acak dari populasi dengan fungsi kepadatan peluang $f(x \mid \theta_1, \theta_2, \ldots, \theta_m)$, fungsi *likelihood* didefinisikan dengan

$$L(\theta_1, \theta_2, \ldots, \theta_m \mid X) = \prod_{i=1}^{n} f(x_i \mid \theta_1, \theta_1, \ldots, \theta_m) \tag{10}$$

Bila fungsi likelihood terdeferensialkan dalam $\theta_1, \theta_2, \ldots, \theta_m$, maka estimator maksimum *likelihood* yang mungkin adalah

$$\frac{\partial}{\partial\theta_k} L(\theta_1, \theta_2, \ldots, \theta_m \mid X) = 0 \tag{11}$$

Uji *Anderson Darling*

Misalkan $x_1, x_2, \ldots, x_n$ adalah data yang akan diuji distribusi, maka uji *Anderson Darling* dapat diperoleh dengan menggunakan rumus sebagai

$$A = -n - \frac{1}{n}\sum_{i=1}^{n}[2i-1]\,[\ln(F(x_i)) - \ln(1-F(n_{n-i+1}))] \tag{12}$$

Dalam hal ini, pendugaan distribusi yang sesuai dipilih berdasarkan nilai *Anderson Darling* terkecil. Beberapa distribusi yang dapat digunakan ketika melakukan uji distribusi yaitu *Smallest extreme value*, Exponensial, Log-logistik, Logistik, Normal, dan Weibull (Lawless, 1982). Distribusi yang memiliki nilai *Anderson Darling* (AD) terkecil adalah distribusi yang paling cocok atau mendekati variabel respon yang berupa waktu *survival*.

Uji Signifikansi Parameter

Uji signifikansi parameter dilakukan setelah diperoleh estimasi model regresi, dilakukan untuk mengetahui variabel independen yang memiliki pengaruh signifikan terhadap model. Pengujian yang dilakukan adalah uji secara simultan yaitu menggunakan uji *partial likelihood ratio* dan pengujian secara parsial yaitu menggunakan uji Wald.

Tahapan Analisis

1. Memilih data tahan *survival* tersensor kanan.
2. Menguji kecocokan data apakah mengikuti distribusi log-logistik.
3. Mengestimasi parameter dengan metode maksimum *likelihood*.
4. Membentuk model regresi *survival.*
5. Menguji signifikansi koefisien regresi secara simultan dan parsial.
6. Mendapatkan model regresi terbaik.

HASIL DAN PEMBAHASAN

Model Regresi pada Data *Survival* Berdistribusi Log-logistik

Model regresi waktu *survival* dibentuk dalam model skala lokasi. Dalam model ini, waktu *survival* T ditransformasikan logaritma sehingga menjadi Y = ln T dan diperoleh persamaan regresi

$$Y = \ln T = \mu(x) + \sigma\varepsilon$$

dengan $\sigma > 0$ dan ε memiliki distribusi logistik.

Diketahui T adalah variabel acak berdistribusi log-logistik dengan fungsi kepadatan peluag T yang diberikan oleh x tertentu. Jika T ditransformasi menjadi Y = ln T, maka diperoleh fungsi kepadatan peluang Y yang diberikan oleh x tertentu yang identik dengan fungsi kepdatan peluang distribusi logistik, yaitu

$$f(y|x) = \frac{\exp(\frac{y-\mu(x)}{\sigma})}{\sigma(1+\exp(\frac{y-\mu(x)}{\sigma}))^2}$$

Fungsi distribusi kumulatifnya adalah

$$F(y|x) = 1 - \frac{1}{1+\exp(\frac{y-\mu(x)}{\sigma})}$$

Fungsi *survival* sebagai berikut

$$S(y|x) = \frac{1}{1+\exp(\frac{y-\mu(x)}{\sigma})} \tag{13}$$

dan fungsi *hazard*-nya adalah

$$f(y|x) = \frac{\exp(\frac{y-\mu(x)}{\sigma})}{\sigma(1+\exp(\frac{y-\mu(x)}{\sigma}))^2} \tag{14}$$

Estimasi Parameter pada Data Tersensor Berdistribusi Log-logistik

Fungsi *likelihood* untuk data tersensor kanan adalah sebagai berikut

$$L = \prod_{i=1}^{n} f(t_i)^{\delta_i} [S(t_i)]^{(1-\delta_i)} \quad (15)$$

Maka, fungsi *likelihood* untuk mengestimasi parameter distribusi log-logistik adalah

$$L(\alpha,\beta) = \prod_{i=1}^{n} [\frac{(\beta/\alpha)(t_i/\alpha)^{\beta-1}}{\{1+(t_i/\alpha)^{\beta}\}^2}]^{\delta_i} [\frac{1}{1+(t_i/\alpha)^{\beta}}]^{(1-\delta_i)} \quad (16)$$

Estimasi parameter distribusi log-logistik dilakukan dengan menurunkan persamaan likelihood di atas terhadap α dan β, yaitu sebagai berikut

$$\sum_{i=1}^{n} \delta_i - \sum_{i=1}^{n} \frac{(\delta_i+1)(t_i/\alpha)^{\beta}\log(t_i/\alpha)}{1+(t_i/\alpha)^{\beta}} = 0 \quad (17)$$

dan

$$\frac{\sum_{i=1}^{n} \delta_i}{\beta} + \sum_{i=1}^{n} \delta_i \log(t_i/\alpha) - \sum_{i=1}^{n} \frac{(\delta_i+1)(t_i/\alpha)^{\beta}\log(t_i/\alpha)}{1+(t_i/\alpha)^{\beta}} = 0 \quad (18)$$

Karena kedua persamaan tersebut tidak dapat diselesaikan secara langsung, maka diperlukan metode numerik untuk mendapatkan hasil estimasinya.

Fungsi *likelihood* untuk data tersensor kanan adalah sebagai berikut

$$L = \prod_{i=1}^{n} f(y_i)^{\delta_i} [S(y_i)]^{(1-\delta_i)} \quad (19)$$

Maka, fungsi *likelihood* untuk mengestimasi parameter distribusi log-logistik adalah

$$L(\alpha,\beta) = \prod_{i=1}^{n} [\frac{\exp(\frac{y_i-\mu(x)}{\sigma})}{\sigma[1+(\frac{y_i-\mu(x)}{\sigma})]}]^{\delta_i} [\frac{1}{1+\exp(\frac{y_i-\mu(x)}{\sigma})}]^{(1-\delta_i)} \quad (20)$$

Estimasi parameter model regresi distribusi log-logistik dilakukan dengan menurunkan persamaan likelihood di atas terhadap μ dan σ, yaitu sebagai berikut

$$\sum_{i=1}^{n} \frac{-\delta_i}{\sigma} + \sum_{i=1}^{n} \frac{(\delta_i+1)(1/\sigma)\exp(z_i)}{1+\exp(z_i)} = 0 \quad (21)$$

dan

$$-\frac{1}{\sigma}\sum_{i=1}^{n}(\delta_i(z_i)) - \sum_{i=1}^{n} \frac{\delta_i}{\sigma} + \sum_{i=1}^{n} \frac{(\delta_i+1)(1/\sigma)\exp(z_i)}{1+\exp(z_i)} = 0 \quad (22)$$

Karena kedua persamaan tersebut tidak dapat diselesaikan secara langsung, maka diperlukan metode numerik untuk mendapatkan hasil estimasinya.

Contoh Penerapan

Penerapan dilakukan pada data pasien kanker paru-paru yang diperoleh dari *The Statistical Analysis of Failure Time Data. 2nd Edition. 2002. John Wiley & Sons, Inc*. Variabel dependen adalah waktu tahan hidup pasien pasca operasi. Variabel independennya terdiri dari *performance status* dan umur pasien. Data yang diperoleh dianalisis menggunakan perangkat lunak Minitab 16, dengan langkah-langkahnya sebagai berikut

1. Pengujian Distribusi
   Beberapa distribusi yang dapat digunakan untuk melakukan uji distribusi yaitu Smallest extreme value, Exponensial, Log-logistik, Logistik, Normal, dan Weibull. Distribusi yang memiliki nilai *Anderson Darling* (AD) terkecil adalah distribusi yang paling cocok. Maka, dapat disimpulkan bahwa data mengikuti distribusi log-logistik, karena nilai *Anderson Darling* terkecil yaitu 0,302 dan *p-value* > 0.250. Berdasarkan *p-value* tersebut maka distribusi yang sesuai adalah distribusi log-logistik.
2. Estimasi Parameter
   Dengan menggunakan metode maksimum *likelihood*, didapat nilai μ=3.85205 dan σ=0.66830. Sehingga diperoleh nilai dari estimasi parameter dari α dan β yaitu sebagai berikut
   $$\alpha = \exp(3.85205) = 47.0895$$
   $$\beta = 1/0.66830\} = 1.49633$$
3. Model Regresi Data *Survival*
   Dengan bantuan perangkat lunak Minitab, model regresi data pasien kanker paru-paru dibentuk dengan menginput semua variabel independen yaitu performance status X_1 dan umur X_2 ke dalam model regresi, sehingga didapatkan hasil sebagai berikut

TABEL 1. HASIL REGRESI WAKTU TAHAN HIDUP BERDISTRIBUSI LOG-LOGISTIK

Variabel	Koefisien	P	Zhitung
Intercept	1.92458	0.189	1.32
X1	0.0242393	0.021	2.31
X2	0.0099162	0.672	0.42
Scale	0.639037		

Model ini menggambarkan hubungan antara waktu *survival* terhadap *performance status* dan umur. Model regresinya adalah

$$t_i = \exp(1.92458+0.0242393x_{1i} + 0.0099162x_{2i} + 0.639037\varepsilon_i) \quad (23)$$

Dengan menggunakan model Y=ln T, persamaannya menjadi

$$y_i = 1.92458+0.0242393x_{1i} + 0.0099162x_{2i} + 0.639037\varepsilon_i \quad (24)$$

4. Uji Signifikansi Parameter
   Uji signifikansi parameter dilakukan setelah diperoleh estimasi model regresi. Pengujian yang dilakukan adalah uji secara simultan dan pengujian secara parsial.
   a. Uji Signifikansi Parameter Secara Simultan
      Hipotesis :
      $H_0 : \theta_1 = \theta_2 = ... = \theta_k = 0$ untuk semua k, dengan k = 1,2
      (secara simultan, koefisien regresi tidak signifikan secara statistik)
      $H_1 : \theta_k \neq 0$ untuk paling sedikit satu k, dengan k = 1,2
      (secara simultan, koefisien regresi signifikan secara statistik)
      Taraf signifikan : $\alpha = 5\% = 0.05$
      Statistik uji : $G = -2\ln L_R/L_f = -2(-536.124+517.868) = 36.512$
      Daerah kritik : H_0 ditolak jika $G > X^2_{0.05,2}$
      Keputusan : H_0 ditolak karena nilai $G > X^2_{0.05,2}$ yaitu 36.512 > 5.99148
      Kesimpulan : Karena H_0 ditolak, maka dapat disimpulkan bahwa variabel X_1, X_2 berpengaruh dalam model.
   b. Uji Signifikansi Parameter Secara Parsial
      Berdasarkan *output* dari perangkat lunak Minitab, diperoleh hasil untuk masing-masing koefisien regresi sebagai berikut

TABEL 2. KOEFISIEN REGRESI DATA *SURVIVAL* TERSENSOR KANAN BERDISTRIBUSI LOG-LOGISTIK

Variabel	Koefisien	P	Zhitung
Intercept	1.92458	0.189	1.32
X1	0.0242393	0.021	2.31
X2	0.0099162	0.672	0.42
Scale	0.639037		

Hipotesis :
$H_0 : \theta_1 = \theta_2 = ... = \theta_k = 0$ untuk semua k, dengan k = 1,2
(secara parsial, koefisien regresi tidak signifikan secara statistik)
$H_1 : \theta_k \neq 0$ untuk paling sedikit satu k, dengan k = 1,2
(secara parsial, koefisien regresi signifikan secara statistik)
Taraf signifikan : $\alpha = 5\% = 0.05$
Daerah kritik : H_0 ditolak jika *p-value* < 5% atau $|Z_{hitung}| > Z_{tabel}$ dimana $Z_{tabel} = Z_{\alpha/2}$
$\theta_1 = |Z_{hitung}| = 2.31$, *p-value* = 0.021 dan $\theta_2 = |Z_{hitung}| = 0.42$, *p-value* = 0.672
Keputusan : H_0 ditolak pada koefisien yang memiliki *p-value* < 5% yaitu pada koefisien θ_2.
Kesimpulan : Dapat disimpulkan bahwa θ_2 tidak signifikan dalam model, sedangkan θ_1 signifikan dalam model. Oleh karena itu, variabel X_2 dikeluarkan dari model.

5. Interpretasi Model Terbaik

Berdasarkan model terbaik yang diperoleh, maka diketahui faktor yang berpengaruh signifikan terhadap waktu *survival* adalah variabel *performance status* X_1 dengan nilai signifikansi *p-value* < 0.05. Maka didapatkan model regresi terbaik untuk data pasien kanker paru-paru yang diperoleh dari *The Statistical Analysis of Failure Time Data. 2nd Edition. 2002. John Wiley & Sons, Inc* adalah

$$y_i = 1.92458+0.0242393x_{1i} + 0.639037\varepsilon_i \quad (25)$$

Model regresi di atas dapat diinterpretasikan bahwa variabel X_1 yaitu *performance status* memiliki pengaruh positif terhadap variabel dependen yaitu waktu tahan hidup pasien kanker paru-paru.

KESIMPULAN DAN SARAN

Kesimpulan

1. Model regresi untuk data *survival* yang berdistribusi log-logistik adalah $Y_i = \mu(x) + \sigma\varepsilon_i$ dimana ε mengikuti distribusi logistik.
2. Estimasi parameter yang didapatkan adalah

$$\sum_{i=1}^{n} \frac{-\delta_i}{\sigma} + \sum_{i=1}^{n} \frac{(\delta_i+1)(1/\sigma)\exp(z_i)}{1+\exp(z_i)} = 0 \quad (26)$$

dan

$$-\frac{1}{\sigma}\sum_{i=1}^{n}(\delta_i(z_i)) - \sum_{i=1}^{n} \frac{\delta_i}{\sigma} + \sum_{i=1}^{n} \frac{(\delta_i+1)(1/\sigma)\exp(z_i)}{1+\exp(z_i)} = 0 \quad (27)$$

Karena kedua persamaan tersebut tidak dapat diselesaikan secara langsung, maka diperlukan metode numerik untuk mendapatkan hasil estimasinya.

3. Model regresi terbaik untuk data pasien kanker paru-paru yang diperoleh dari *The Statistical Analysis of Failure Time Data. 2nd Edition. 2002. John Wiley & Sons, Inc* adalah

$$y_i = 1.92458+0.0242393x_{1i} + 0.639037\varepsilon_i \quad (28)$$

Saran

Saran dari penulis yaitu, dapat dilakukan pengembangan lebih lanjut tentang model regresi pada data *survival* dengan distribusi yang lain. Seperti distribusi lognormal, distribusi Weibull, dan lain-lain. Selain itu, data *survival* yang digunakan dapat berupa data tersensor kiri.

UCAPAN TERIMA KASIH

Terima kasih kepada dosen pembimbing dan seluruh pihak yang telah memberikan bantuan kepada penulis dalam menyelesaikan penelitian ini.

REFERENSI

Alakus, K., Erilli, N.A. 2001. *Confidence Intervals Estimation for Survival Function in Log-logistic Distribution and Proportional Odds Regression Based on Censored Survival Time Data*. J Biomet Biostat, 2:116.

Bain, Lee J., Engelhardt, Max. 1991. *Introduction to Probability and Mathematical Statistics 2nd ed*. California: Duxbury Press.

Burnet, Fraix., Gabaud, Valls. 2014. *Surival Data and Regression Models*. EAS Publication Series, 66, 125-147.

Kalbfleisch dan Prentice. 2002. *The Statistical Analysis of Failure Time Data 2nd ed*. Wiley & Sons, Inc. Canada.

Klein, John P., Moeschberger, Melvin L. 2003. *Survival Analysis Techniques for Censored and Truncated Data 2nd ed*. New York: Springer.

Kleinbaum, D.G., Klein, M. 2005. *Survival Analysis: A Self-Learning Text 2nd ed*. New York: Springer.

Lawless, J.F. 1982. *Statistical Model and Methods for Lifetime Data 2nd ed*. New Jersey: John Wiley & Sons.

Sari, Ulfah. 2010. *Penggunaan Metode Hazard dalam Menentukan Loyalitas Pengguna Kartu Seluler GSM Prabayar*. Tesis. Diterbitkan. Institus Pertanian Bogor: Bogor.