Perbandingan Aplikasi Metode Parametrik (Distribusi Log logistik) dan Non Parametrik (Nelson-Aalen Estimator) dalam Analisis Data Uji Hidup (Studi Kasus: Lama Waktu Kelulusan Mahasiswa Program Studi Statistika FMIPA UNMUL Angkatan 2005)

Application of Survival Analysis : Comparison Between Parametric Method (Log Logistic Distribution) and Nonparametric Method (Nelson-Aalen Estimator) (Case Study: Duration Study of Students of Mulawarman University Faculty of Math and Natural Science Majoring in Statistics 2005)

Riki Herliansyah[1], Yuki Novia Nasution[2], Sri Wahyuningsih[3]
[1]Mahasiswa Program Studi Statistika
[2,3]Dosen Program Studi Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Mulawarman
Email: rickiherliansyah@gmail.com[1], yuki.novia.n@gmail.com[2], swahyuningsih@gmail.com[3]

Abstract

Survival time can be defined broadly as the time to the occurrence of a given event. In this research, we do survival analysis which use parametric (Log Logistic distribution) and nonparametric (Nelson-Aalen Estimator) estimation methods. The purpose of this research is to compare variance between Nelson-Aalen Estimator and Log Logistic Distribution to obtain Relative Efficiency value. Chi-Square method is used to test goodness of fit of data distribution. Parameter of Log Logistic distribution is estimated with algoritm EM and Nelson-Aalen Estimator is estimated with help of R 3.0.l. Analysis results show that 32 from 38 students majoring in statistic 2005 have been already graduated. The highest probability for students majoring in statistic 2005 finish their study is at 63rd month. Survival function show that the longer duration study the the higher probability for them to finish the study. The relative efficiency value is less than 1, so that Nelson-Aalen Estimator is more suitable estimation method than Log Logistic distribution for data duration study of students of Mulawarman University Faculty of Math and Natural Science majoring in Statistics 2005.

Keywords : Survival analysis, log logistic, Nelson-Aalen estimator, variance, relative efficiency.

Pendahuluan

Banyak peristiwa terjadi dalam kehidupan sehari-hari yang terkait dengan waktu, misalnya waktu yang diperlukan oleh seorang *sales* untuk menjual sejumlah barang tertentu, lamanya waktu yang diperlukan oleh seseorang untuk memperoleh kesembuhan dari penyakit yang dideritanya dan lain-lain. Data *survival* adalah data tentang pengamatan jangka waktu dari awal pengamatan sampai terjadinya suatu peristiwa (*event*) (Lawless, 1994).

Banyak peneliti menganggap analisis *survival* untuk sekadar penerapan dua model statistik konvensional untuk tipe khusus masalah yaitu parametrik jika distribusi waktu *survival* diketahui normal dan non parametrik jika distribusi tidak diketahui. Asumsi ini akan benar jika waktu kelangsungan hidup semua objek tepat dan diketahui, namun beberapa waktu *survival* tidak. Diketahui bahwa distribusi *survival* sering miring atau jauh dari keadaan normal. Metode sederhana tidak dapat dilakukan jika beberapa objek masih bertahan pada saat melakukan analisis dan karenanya waktu *survival* objek tersebut tidak diketahui.

Metode non parametrik atau metode bebas distribusi cukup mudah untuk dipahami dan diterapkan. Tetapi metode non parametrik kurang efisien dibandingkan metode parametrik ketika waktu *survival* mengikuti sebuah distribusi teoritis dan lebih efisien ketika tidak ada distribusi teoritis yang cocok diketahui (Lee dan Wang, 2003).

Fungsi *Survival* dan Fungsi *Hazard*

Jika T melambangkan waktu *survival* dan mempunyai distribusi peluang $f(t)$, maka distribusi kumulatif dinyatakan sebagai berikut (Collet, 1994) :

$$F(t) = P(T \le t) = \int_0^t f(u)\,du \qquad (1)$$

Fungsi *survival*, $S(t)$, didefinisikan sebagai probabilitas suatu objek bertahan setelah waktu ke– *t*, yaitu,

$$S(t) = P(T > t) = 1 - P(T \le t) = 1 - F(t) \qquad (2)$$

Fungsi *hazard* merupakan laju *failure* sesaat dengan asumsi telah bertahan sampai waktu ke – *t*, yaitu :

$$h(t) = \lim_{\delta t \to 0}\left\{\frac{P(t \le T < t + \delta t \mid T \ge t)}{\delta t}\right\} \qquad (3)$$

$$h(t) = \frac{f(t)}{S(t)} \qquad (4)$$

dimana fungsi kumulatif *hazard* adalah

$$H(t) = \int_0^t h(t)\,dt \qquad (5)$$

maka hubungan antara fungsi kumulatif *hazard* ($H(t)$), dan fungsi *survival* ($S(t)$), adalah :

$$H(t) = -\ln S(t) \qquad (6)$$

Estimasi Median Data *Survival*

Oleh karena distribusi waktu *survival* cenderung positif menceng ke kanan, maka median lebih dipilih dapat menyimpulkan pengukuran lokasi dari distribusi data. Setelah diperoleh nilai estimasi fungsi *survival*, maka langkah selanjutnya adalah mengestimasi nilai median. Median merupakan waktu pengamatan objek dalam populasi melebihi 50 % diharapkan bertahan (*survive*), yaitu dimana S{t}=0,5.

Dikarenakan estimasi non parametrik dari S(t) merupakan fungsi tangga, biasanya jarang diperoleh nilai fungsi *survival* tepat sama dengan 0,5. Maka, nilai estimasi median dari waktu *survival*,

$$\hat{t}(50) = \min\left\{t_i \,\middle|\, \hat{S}(t_i) \le 0{,}5\right\}$$

dimana t_i adalah waktu *survival* yang teramati pada objek ke-*i*, dengan *i* = 1, 2, …, *n*.

Distribusi Log Logistik

Jika sebuah variabel acak T berdistribusi Log Logistik yang merupakan distribusi kontinu dimana distribusi ini mempunyai dua parameter yaitu α sebagai parameter lokasi dan β sebagai parameter skala dimana α, β > 0. Fungsi kepadatan peluang distribusi Log Logistik adalah (Cox dan Oakes, 1984) :

$$f(t) = \frac{\kappa t^{\kappa-1}\rho^{\kappa}}{\left\{1+(t\rho)^{\kappa}\right\}^2} \qquad (7)$$

di mana :

$$\rho = \frac{1}{e^{\alpha}}$$

dan $\kappa = \frac{1}{\beta}$.

a. Fungsi Kepadatan Kumulatif ($F(t)$)

Distribusi Log Logistik merupakan distribusi yang kontinu, maka fungsi kepadatan kumulatif *F(t)* adalah sebagai berikut:

$$F(t) = \int_0^t \frac{\kappa u^{\kappa-1}\rho^{\kappa}}{\left\{1+(u\rho)^{\kappa}\right\}^2}du = -\frac{1}{1+(t\rho)^{\kappa}} + 1 \qquad (8)$$

b. Fungsi *Survival* ($S(t)$)

Berdasarkan persamaan (2) dan persamaan (8) maka diperoleh fungsi *survival* yaitu sebagai berikut:

$$S(t) = 1 - F(t) = \frac{1}{1+(t\rho)^{\kappa}} \qquad (9)$$

c. Fungsi *Hazard* ($h(t)$)

Berdasarkan pada perasamaan (4) maka fungsi *hazard* adalah sebagai berikut:

$$h(t) = \frac{f(t)}{S(t)} = \frac{\kappa t^{\kappa-1}\rho^{\kappa}}{1+(t\rho)^{\kappa}} \qquad (10)$$

d. Fungsi *Cumulative Hazard* ($H(t)$)

Berdasarkan persamaan (6) maka fungsi *cumulative hazard* H(t) adalah:

$$H(t) = -\ln S(t) = \ln(1+(t\rho)^{\kappa}) \qquad (11)$$

Nelson-Aalen Estimator

Metode ini juga kadang disebut sebagai *empirical cumulative hazard function*, tetapi lebih umum dikenal sebagai *Nelson-Aalen (NA)* estimator, yang dikemukakan oleh Nelson pada tahun1969 dan Aalen dalam sebuah tesis pada tahun 1972. Secara umum estimasi Nelson-Aalen adalah sebagai berikut (Lawless, 1994):

$$\hat{H}(t) = \sum_{j \le t} \frac{d_j}{n_j} \qquad (12)$$

Variansi Distribusi Log Logistik

Variansi distribusi Log Logistik dapat diestimasi dengan terlebih dahulu mencari $E[X]$:

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)dx = \frac{1}{\rho}\frac{\pi/\kappa}{\sin(\frac{\pi}{\kappa})}$$

Selanjutnya akan dicari $E[X^2]$ sebagai berikut :

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx = \frac{1}{\rho^2}\frac{2\pi/\kappa}{\sin(\frac{2\pi}{\kappa})}$$

Sehingga variansi dari X adalah :

$$Var\,(X) = E(X^2) - [E(X)]^2 = \frac{1}{\rho^2}\frac{2\pi/\kappa}{\sin(\frac{2\pi}{\kappa})} - \frac{1}{\rho^2}\frac{\pi^2/\kappa^2}{\sin^2(\frac{\pi}{\kappa})}$$

Jika $0 < \kappa \le 1$ maka $Var\,(X)$ tidak ada. Jika $1 < \kappa \le 2$ maka $Var\,(X) = \infty$. Dan jika $\kappa > 2$ maka :

$$Var(X) = \frac{1}{\rho^2}\frac{2\pi/\kappa}{\sin(\frac{2\pi}{\kappa})} - \frac{1}{\rho^2}\frac{\pi^2/\kappa^2}{\sin^2(\frac{\pi}{\kappa})} \qquad (13)$$

Variansi Nelson-Aalen Estimator

Dengan menggunakan persamaan (12) maka variansi Nelson-Aalen estimator dapat diestimasi sebagai berikut dengan :

$$Var\left(\hat{H}(t)\right) = E\,[\hat{H}(t) - E\left(\hat{H}(t)\right)]^2$$

Sehingga variansi dari Nelson-Aalen estimator adalah

$$Var\left(\hat{H}(t)\right) = \sum_{j<t} \frac{\frac{d_j}{n_j}\left[\frac{n_j - d_j}{n_j}\right]}{n_j - 1} \quad (14)$$

dan

$$Var\left(\hat{H}(t)\right) = \sum_{j<t} \frac{d_j}{n_j^2}, \text{ (jika no } ties\text{)} \quad (15)$$

***Relative Efficiency* (RE)**

Misalkan $\hat{\theta}$ dan $\hat{\theta}^*$ merupakan estimator tak bias untuk θ. *Relative efficiency* dari $\hat{\theta}$ terhadap $\hat{\theta}^*$ didefinisikan sebagai :

$$re\left(\hat{\theta}, \hat{\theta}^*\right) = \frac{Var\,(\hat{\theta}^*)}{Var\,(\hat{\theta})} \quad (16)$$

$\hat{\theta}^*$ dikatakan efisien jika $re\left(\hat{\theta}, \hat{\theta}^*\right) \leq 1$ untuk semua estimator tak bias $\hat{\theta}$ untuk θ dan semua $\theta \in \Omega$ (Herrhyanto dan Gantini, 2009).

Uji *Goodness of Fit*

Uji *Goodness of Fit* mengukur keserasian dari sebuah sampel acak terhadap suatu fungsi distribusi peluang teoritis. Dengan kata lain, uji ini menunjukkan bagaimana suatu distribusi yang dipilih cocok untuk suatu data. Uji *Goodness of Fit* digunakan untuk menguji hipotesis sebagai berikut:

H_0 : Waktu survival berdistribusi tertentu

H_1 : Waktu survival tidak berdistribusi tertentu

Beberapa uji *Goodness of Fit* antara lain adalah uji Kolmogorov-Smirnov (KS), uji Anderson-Darling dan uji *Chi-Square* (χ^2). Adapun uji yang digunakan dalam penelitian ini adalah uji *Chi-Square* (χ^2).

Uji *Chi-Square* digunakan untuk memutuskan apakah suatu sampel berasal dari suatu populasi dengan suatu distribusi tertentu. Uji ini hanya dapat digunakan untuk data sampel yang kontinu. Statistik uji ini dapat dinyatakan dalam formula berikut :

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} \quad (16)$$

Dimana O_i menyatakan frekuensi sesungguhnya dan E_i menyatakan frekuensi harapan. H_0 ditolak jika nilai statistik $\chi^2 > \chi^2_{\alpha;n-1}$.

Metodologi Penelitian

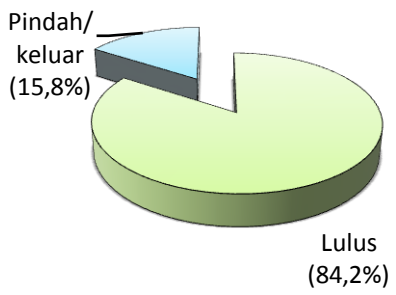
Data yang digunakan dalam penelitian ini adalah data lama waktu kelulusan mahasiswa Program Studi Statistika Fakultas MIPA UNMUL angkatan 2005. Adapun variabel yang digunakan dalam penelitian ini adalah lama kelulusan yaitu lama kelulusan mahasiswa sampai dengan lulus dan berada dalam batas periode penelitian, dalam satuan bulan. Jenis sensor yang digunakan dalam penelitian ini adalah sensor kanan. Jenis sensor ini terjadi jika objek yang diamati pindah/keluar sebelum *failure* teramati atau *failure* terjadi setelah data dikumpulkan/penelitian selesai. Untuk penentuan status sensor digunakan ketentuan sebagai berikut :

1. Jika seorang mahasiswa menempuh studi sampai dengan lulus kurang dari atau sama dengan 14 Desember 2013, maka waktu *survival* (lama studi sampai dengan lulus) dikategorikan sebagai data *survival* teramati tidak tersensor.
2. Jika seorang mahasiswa tidak diketahui lama studinya sampai dengan lulus, maka waktu *survival* adalah lama semester yang ditempuh, dihitung dari awal semester (t = 0) sampai terjadinya hal berikut :
   a. Melebihi batas akhir periode penelitian, yaitu 14 Desember 2013,
   b. Keluar karena pindah ke jurusan/perguruan tinggi lain,

maka, waktu *survival* termasuk dalam kategori data *survival* yang tersensor. Status sensor bersifat biner dengan nilai 0 untuk waktu *survival* yang tersensor dan bernilai 1 untuk waktu *survival* yang tidak tersensor. Selain itu, terdapat variabel lain seperti jenis kelamin yaitu laki-laki dan perempuan dan asal daerah untuk Samarinda dan luar Samarinda yang digunakan untuk memberikan deksriptif mengenai data.

Analisis dan Pembahasan

Hasil analisis menunjukkan bahwa dari 38 orang mahasiswa S1 Program Studi Statistika angkatan 2005 FMIPA UNMUL yang teramati sampai dengan 14 Desember 2013, dapat diketahui bahwa sebesar 84,2% telah mencapai kelulusan.



Gambar 1. Diagram Lingkaran Kelulusan Mahasiswa Program Studi Statistika angkatan 2005.

Identifikasi Distribusi Data

Tabel 1 menunjukan nilai statistik (χ^2) dan χ^2_{tabel} dari uji *Chi-Square* dengan menggunakan bantuan *software EasyFit* dengan hipotesis sebagai berikut :

H_0 : Waktu *survival* berdistribusi log logistik

H_1 : Waktu *survival* tidak berdistribusi log logistik

Tabel 1. Uji Kesesuaian Distribusi Data

Uji	χ^2	*P-Value*	χ^2_{tabel}
Chi-Square	0,218	0,897	5,992

Berdasarkan Tabel 1 dapat diperoleh nilai statistik *Chi-Square* yaitu $\chi^2 = 0{,}218$, karena nilai $\chi^2 < \chi^2_{tabel}$ (5,992) pada taraf signikansi $\alpha = 5\%$, sehingga diputuskan H_0 gagal ditolak. Dengan demikian dapat disimpulkan bahwa waktu *survival* distribusi log logistik.

Pendugaan Parameter Data

Estimasi nilai parameter untuk distribusi Log Logistik dengan bantuan *software* R 3.0.1 adalah sebagai berikut:

Tabel 2. Pendugaan Parameter Distribusi Log Logistik

Data	α	β
Mahasiswa Statistika	4,105	0,116

Berdasarkan Tabel 2 maka diperoleh nilai ρ dan κ yaitu sebagai berikut :

$$\rho = \frac{1}{e^{\alpha}} = \frac{1}{e^{4,105}} == 0{,}016$$

$$\kappa = \frac{1}{\beta} = \frac{1}{0{,}116} = 8{,}63.$$

Analisis *Survival* Menggunakan Distribusi Log Logistik

1. Fungsi Kepadatan Peluang Distribusi Log Logistik

Fungsi kepadatan peluang adalah rumusan matematis untuk menentukan peluang kelulusan seorang mahasiswa untuk yang telah lulus pada waktu ke-*t*. Adapun fungsi kepadatan peluang distribusi Log Logistik berdasarkan persamaan (7) adalah sebagai berikut:

$$f(t) = \frac{8{,}634\, t^{7,634}\, 0{,}016^{8,634}}{(1 + (0{,}016t)^{8,634})^2}$$

Untuk nilai fungsi kepadatan peluang data lamanya masa studi mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005 hingga 14 Desember 2013 dapat dilihat pada Tabel 3.

Tabel 3 menunjukkan peluang kelulusan mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005. Nilai $f(t)$ menunjukkan peluang kelulusan mahasiswa pada waktu ke-t. Berdasarkan tabel tersebut dapat diketahui bahwa peluang tertinggi untuk mahasiswa lulus adalah pada saat $t = 63$, sedangkan peluang terendahnya adalah pada saat $t = 51$.

2. Fungsi Kepadatan Kumulatif

Fungsi peluang kumulatif adalah rumusan matematis untuk menentukan peluang kumulatif kelulusan seorang mahasiswa untuk yang telah lulus pada waktu ke-*t*. Adapun fungsi peluang kumulatif distribusi Log Logistik berdasarkan persamaan (8) adalah sebagai berikut:

$$F(t) = 1 - \frac{1}{1 + (0{,}016t)^{8,364}}$$

Untuk nilai peluang kumulatif distribusi data lamanya masa studi mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005 hingga 14 Desember 2013 dapat dilihat pada Tabel 4. Tabel 4 menunjukkan peluang kelulusan kumulatif mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005. Nilai $F(t)$ menunjukkan peluang kelulusan kumulatif mahasiswa pada waktu ke-t. Misalkan pada saat $t = 54$, peluang kelulusan mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005 untuk waktu kurang dari atau sama dengan 54 bulan atau 4 tahun 6 bulan adalah sebesar 0,2206.

3. Fungsi *Survival*

Funsi *survival* merupakan nilai fungsi peluang seorang mahasiswa dapat bertahan hingga waktu tertentu. Adapun fungsi *survival* berdasarkan persamaan (9) adalah sebagai berikut:

$$S(t) = \frac{1}{(1 + (0{,}016t)^{8,364})}$$

Untuk nilai fungsi *survival* data lamanya masa studi mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005 hingga 14 Desember 2013 dapat dilihat pada Tabel 5.

Tabel 3. Perhitungan Kepadatan Peluang

Waktu (bulan)	$f(t)$
51	0,0213
53	0,0255
54	0,0275
58	0,0336
63	0,0342
65	0,0323
70	0,0245
77	0,0136
83	0,0076

Tabel 4. Perhitungan Peluang Kumulatif Distribusi

Waktu (bulan)	F(t)
51	0,1473
53	0,1941
54	0,2206
58	0,3441
63	0,5172
65	0,5839
70	0,7268
77	0,8583
83	0,9205

Tabel 5 menunjukkan peluang kumulatif mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005 yang masih menjalani studi. Nilai $S(t)$ menunjukkan peluang kumulatif mahasiswa yang masih menjalani studi kurang dari atau sama dengan t. Misalkan pada saat $t = 58$, peluang kumulatif mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005 masih menjalani studi (belum lulus) untuk waktu kurang dari atau sama dengan 58 bulan atau 4 tahun 10 bulan adalah sebesar 0,6559.

Tabel 5. Perhitungan Nilai *Survival*

Waktu (bulan)	$S(t)$
51	0,8527
53	0,8059
54	0,7794
58	0,6559
63	0,4828
65	0,4161
70	0,2732
77	0,1417
83	0,0795

4. Fungsi *Hazard*

Fungsi *hazard* merupakan nilai fungsi rata-rata banyaknya mahasiswa program studi statistika FMIPA Unmul pada waktu tertentu atau sebagai fungsi laju rata-rata seorang mahasiswa Program Studi Statistika FMIPA UNMUL pada saat t bulan. Adapun fungsi risiko (*hazard*) berdasarkan persamaan(10) adalah sebagai berikut:

$$h(t) = \frac{8{,}634\, t^{7{,}634}\, 0{,}016^{8{,}634}}{1 + (0{,}016t)^{8{,}634}}$$

Untuk nilai fungsi *hazard* data data lamanya masa studi mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005 hingga 14 Desember 2013 dapat dilihat pada Tabel 6.

Tabel 6 menunjukkan risiko kelulusan mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005. Nilai $h(t)$ menunjukkan risiko kelulusan mahasiswa pada waktu ke-t. Berdasarkan tabel tersebut dapat diketahui bahwa risiko tertinggi untuk mahasiswa lulus adalah pada saat $t = 77$, sedangkan risiko terendahnya adalah pada saat $t = 51$.

5. Fungsi Kumulatif *Hazard*

Fungsi kumulatif *hazard* aitu nilai fungsi kumulatif rata-rata banyaknya mahasiswa program studi statistika FMIPA Unmul hingga Maret 2014. Adapun fungsi kumulatif risiko (*hazard*) berdasarkan persamaan (11) adalah sebagai berikut:

$$H(t) = \ln(1 + (0{,}016t)^{8{,}364})$$

Untuk nilai fungsi kumulatif *hazard* data lamanya masa studi mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005 hingga 14 Desember 2013 dapat dilihat pada Tabel 7.

Tabel 6. Perhitungan Tingkat Risiko (*Hazard*)

Waktu (bulan)	$h(t)$
51	0,0249
53	0,0316
54	0,0353
58	0,0512
63	0,0709
65	0,0776
70	0,0896
77	0,0962
83	0,0958

Tabel 7. Perhitungan Tingkat Risiko Kumulatif (*Cumulative Hazard*)

Waktu (bulan)	$H(t)$
51	0,1594
53	0,2158
54	0,2492
58	0,4217
63	0,7281
65	0,8767
70	1,2976
77	1,9542
83	2,5321

Tabel 7 menunjukkan risiko kelulusan kumulatif mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005. Nilai $H(t)$ menunjukkan nilai risiko seorang mahasiswa untuk lulus kurang dari atau sama dengan t. Misalkan pada saat $t = 58$ maka tingkat risiko kelulusan mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005 untuk waktu kurang dari atau sama dengan 58 bulan adalah sebesar 0,4217.

Analisis *Survival* Menggunakan Nelson-Aalen Estimator

1. Estimasi Fungsi Kumulatif *Hazard*

Fungsi kumulatif *hazard* adalah nilai fungsi kumulatif rata-rata banyaknya mahasiswa Program Studi Statistika FMIPA UNMUL hingga 14 Desember 2013. Perhitungan dilakukan dengan menggunakan persamaan (12) dengan hasil pada Tabel 8 atau dengan bantuan software R 3.0.1. Berdasarkan pada Tabel 8 dapat dijelaskan bahwa pada saat $t = 54$, banyaknya

jumlah individu yang berisiko dalam hal ini adalah jumlah mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005 yang berisiko (lulus) adalah sebanyak 32 orang dan jumlah individu yang mengalami *event* (lulus) adalah sebanyak 7 orang. Misalkan pada saat $t = 58$ maka tingkat risiko kelulusan mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005 untuk waktu kurang dari atau sama dengan 58 bulan adalah sebesar 0,7892.

Tabel 8. Estimasi Fungsi *Cumulative Hazard* Metode Nelson-Aalen

Waktu	**Jumlah Mahasiswa yang Berisiko Lulus (n_j)**	**Jumlah Mahasiswa yang Lulus (d_j)**	***Cumulative hazard*** ($\hat{H}(t)$)
51	32	7	0,2188
53	25	6	0,4588
54	19	1	0,5114
58	18	5	0,7892
63	13	1	0,8661
65	12	2	1,0328
70	10	1	1,1328
77	9	2	1,3550
83	7	7	2,3550

2. Estimasi Fungsi *Survival*

Fungsi *survival* adalah nilai fungsi peluang seorang mahasiswa S1 Program Studi Statistika FMIPA UNMUL dapat bertahan hingga waktu tertentu. Penghitungan estimasi tingkat kelulusan mahasiswa S1 Program Studi Statistika FMIPA UNMUL adalah dengan menggunakan metode Nelson-Aalen. Untuk nilai fungsi data lamanya masa studi mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005 hingga 14 Desember 2013 dapat dilihat pada Tabel 9.

Tabel 9. Estimasi Fungsi *Survival* Nelson-Aalen Estimator

Waktu	**Jumlah Mahasiswa yang Berisiko Lulus (n_j)**	**Jumlah Mahasiswa yang Lulus (d_j)**	**Fungsi *Survival*** ($\hat{S}(t)$)
51	32	7	0,8035
53	25	6	0,6321
54	19	1	0,5997
58	18	5	0,4542
63	13	1	0,4206
65	12	2	0,3560
70	10	1	0,3221
77	9	2	0,2580
83	7	7	0,0949

Nilai $S(t)$ menunjukan peluang kumulatif mahasiswa yang masih menjalani studi. Misalkan pada saat t = 63 maka peluang kumulatif mahasiswa Program Studi Statistika FMIPA UNMUL menjalani studi untuk waktu kurang dari atau sama dengan 63 bulan atau 5 tahun 3 bulan adalah sebesar 0,4206.

Perbandingan Distribusi Log Logistik dan Nelson-Aalen Estimator

Gambar 2.dan 3. menunjukkan nilai estimasi dari fungsi *survival* dan kumulatif *hazard* untuk distribusi Log Logistik dan Nelson-Aalen estimator. Berdasarkan pada kedua grafik tersebut dapat disimpulkan bahwa t kecil yaitu $t = 51$ hingga $t = 63$, nilai fungsi *survival* dan kumultif *hazard* dari kedua metode tersebut menunjukan nilai estimasi yang hampir sama. Namun, pada saat waktu studi mengalami peningkatan (lebih besar dari 63), perbedaan nilai estimasi dari kedua metode tersebut juga semakin membesar.

Selanjutnya dilakukan perhitungan variansi dari distribusi Log Logistik dan Nelson-Aalen estimator. Kemudian dilakukan perbandingan nilai variansi dari kedua metode tersebut untuk mengetahui metode mana yang lebih efisien untuk data lama studi mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005.

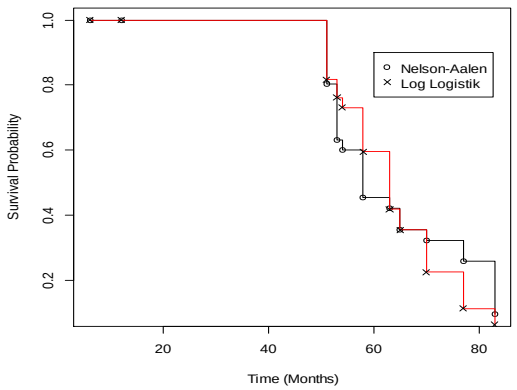
1. Variansi Distribusi Log Logistik
   Dengan menggunakan persamaan (13) maka diperoleh variansi dari distribusi Log Logistik yaitu :
   $$Var = \frac{1}{0{,}016^2}\frac{2\left(\frac{\pi}{8{,}634}\right)}{\sin\left(\frac{2\pi}{8{,}634}\right)} - \frac{1}{0{,}016^2}\frac{\left(\frac{\pi^2}{8{,}634^2}\right)}{\sin^2\left(\frac{\pi}{8{,}634}\right)}$$
   $$= 190{,}288$$
2. Variansi Nelson-Aalen Estimator
   Dengan menggunakan persamaan (14) maka diperoleh variansi dari Nelson-Aalen yaitu:
   $$Var\left(\hat{H}(t)\right) = \frac{\frac{1}{32}\left[\frac{32-7}{32}\right]}{32-1} + \frac{\frac{1}{25}\left[\frac{25-6}{25}\right]}{25-1}$$
   $$+\frac{\frac{1}{19}\left[\frac{19-1}{19}\right]}{19-1} + \cdots + \frac{\frac{1}{7}\left[\frac{7-7}{7}\right]}{7-1} = 0{,}0402$$

Dengan menggunakan persamaan (16) maka diperoleh nilai efisien relatif yaitu :

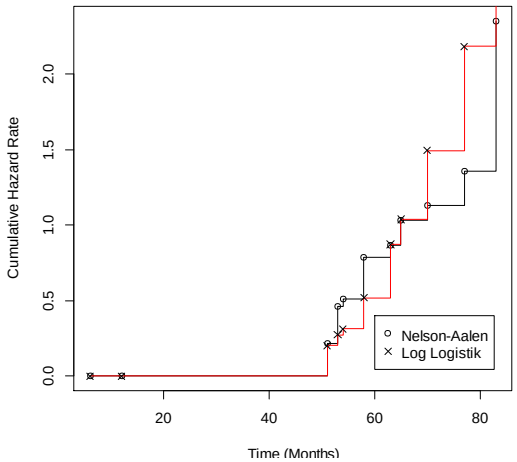
$$re(\hat{\theta}, \hat{\theta}^*) = \frac{Var(Nelson - Aalen)}{Var(log\ logistik)}$$

$$re(\hat{\theta}, \hat{\theta}^*) = \frac{0{,}0402}{190{,}288} = 0{,}0002$$

Berdasarkan perhitungan di atas, dapat dilihat bahwa nilai efisien relatif kurang dari 1, yang mengindikasikan bahwa Nelson-Aalen Estimator merupakan metode yang lebih efisien dari distribusi Log Logistik. Sehingga, dapat disimpulkan dalam penelitian ini, metode non parametrik merupakan metode yang lebih efisien dibandingkan metode parametrik pada data mahasiswa S1 Program Studi Statistika angkatan 2005 FMIPA UNMUL.



Gambar 2. Perbandingan Fungsi *Survival* Distribusi Log Logistik dan Nelson-Aalen Estimator



Gambar 3. Perbandingan Fungsi Kumulatif *Hazard* Distribusi Log Logistik dan Nelson-Aalen Estimator

Kesimpulan

Adapun kesimpulan yang dapat diambil dari hasil analisis adalah sebagai berikut :

1. Dari 38 orang mahasiswa S1 Program Studi Statistika angkatan 2005 FMIPA UNMUL yang teramati sampai dengan 14 Desember 2013, dapat diketahui bahwa sebesar 84,2% telah mencapai kelulusan.
2. Berdasarkan grafik fungsi *survival* untuk lama masa studi mahasiswa S1 Program Studi Statistika FMIPA UNMUL yang teramati sampai dengan 14 Desember 2013 dapat diketahui bahwa peluang tertinggi mahasiswa Program Studi Statistika FMIPA UNMUL angkatan 2005 menyelesaikan studi adalah pada saat $t = 61$. Sedangkan berdasarkan grafik fungsi *survival* untuk lama masa studi mahasiswa S1 Program Studi Statistika angkatan 2005 FMIPA UNMUL yang teramati sampai dengan 14 Desember 2013 dapat diketahui bahwa semakin lama masa studi yang ditempuh maka peluang mahasiswa tersebut masih menempuh studi (belum lulus) akan semakin kecil.
3. Nilai efisiensi relatif yang diperoleh adalah kurang dari 1, yang mengindikasikan bahwa metode Nelson-Aalen Estimator merupakan metode yang lebih efisien dari distribusi Log Logistik untuk studi kasus dalam penelitian ini.

Daftar Pustaka

Collet, D. 1994. *Modelling Survival Data In Medical Research*. London: Chapman and Hall,.

Cox, D. R. dan Oakes, D. 1984. *Analysis of Survival Data*. London: Chapman and Hall.

Herrhyanto N dan T., Gantini 2009. *Pengantar Statistika Matematis*. Bandung : Yrama Widya.

Lawless, J. F. 1994. *Statistical Models and Methods for Lifetime Data*. Canada: John Willey and Sons.

Lee, E. T. dan Wang, John Wenyu. 2003. *Statistical Methods for Survival Data Analysis*. , Belmont, CA: Lifetime Learning Publications.