

**MODEL REGRESI COX PROPORTIONAL HAZARDS
PADA DATA LAMA STUDI MAHASISWA
(Studi Kasus Di Fakultas Sains dan Matematika
Universitas Diponegoro Semarang Mahasiswa Angkatan 2009)**

Landong P Hutahaean¹, Moch. Abdul Mukid², Triastuti Wuryandari³

¹Alumni Jurusan Statistika FSM UNDIP

^{2,3}Staff Pengajar Jurusan Statistika FSM UNDIP

Abstract

High education has important role to increase the intellectual life of the nation and the development of natural sciences and technology by producing the quality graduates. The quality graduates just need 48 month to finish the study. There are many factors that will affect time of study students as Grade Point Average(GPA), Bustle student level, etc. Hence, need to know what factors affecting time of study students. One method that can be used is Survival analysis. Survival Analysis is analysis of survival data from the beginning of time research until certain events occurred. One of the methods of survival analysis is Cox Proportional Hazards Regression. Cox Proportional Hazards Regression is a regression which used data of intervals of time an event. The case which is discussed in this research is factors that affect time of study students of Faculty of Science and Mathematics started 2009 Diponegoro of University with the second type of censoring. From the research give conclusion that factors affecting time of study students is Department, GPA, and Organization.

Keywords : Time of study, Survival Analysis, Cox Proportional Hazards Regression.

I. PENDAHULUAN

Lama studi merupakan waktu yang dibutuhkan seorang mahasiswa untuk menyelesaikan pendidikan sesuai dengan jenjang masing-masing, biasanya untuk jenjang S1 adalah 4 tahun. Diduga ada beberapa faktor yang mempengaruhi lama studi seperti Indeks prestasi kumulatif, tingkat kesibukan mahasiswa dan lain-lain. Oleh karena itu peneliti tertarik untuk menyelidiki faktor-faktor yang diduga mempengaruhi lama studi. Analisis ini disebut dengan analisis ketahanan hidup. Menurut Collett (2004), analisis ketahanan hidup menggambarkan analisis data waktu tahan hidup dari awal waktu penelitian sampai kejadian tertentu terjadi. Regresi *Cox Proportional Hazards* merupakan metode regresi dalam analisis ketahanan hidup yang digunakan untuk mengetahui hubungan antara variabel dependen dengan variabel independen. Data yang digunakan pada regresi *Cox Proportional Hazards* berupa data waktu tahan hidup dari suatu individu. Kasus yang diangkat pada penelitian ini adalah model regresi *Cox Proportional Hazards*, estimasi parameter, perhitungan rasio kegagalan dan estimasi peluang mahasiswa S1 Fakultas Sains dan Matematika Universitas Diponegoro Semarang angkatan 2009 yang dapat lulus tepat waktu. Permasalahan dibatasi dengan menggunakan data lama studi mahasiswa S1 Fakultas Sains dan Matematika

Universitas Diponegoro angkatan 2009 dan metode yang digunakan dalam mengestimasi parameter adalah Maximum Likelihood Estimation.

II. TINJAUAN PUSTAKA

2.1. Analisis Ketahanan Hidup

Analisis ketahanan hidup merupakan alat statistik yang tujuan utamanya adalah menganalisis data yang selalu positif dalam skala pengukuran dengan jarak interval data awal dan akhir (Lee, 1992). Menurut Cox dalam Collet (2004), ada tiga hal yang harus diperhatikan dalam menentukan waktu ketahanan hidup (biasanya disebut dengan T) secara tepat, yaitu sebagai berikut :

- a. Waktu awal tidak ambigu yang berarti tidak ada dua pengertian atau lebih.
- b. Definisi terjadinya kegagalan secara keseluruhan harus jelas.
- c. Skala waktu sebagai satuan pengukuran harus jelas.

2.2. Tipe Penyensoran Data

Analisis ketahanan hidup sangat mempertimbangkan penyensoran. Data dikatakan tersensor apabila waktu hidup obyek penelitian tidak diketahui. Hal ini disebabkan oleh kejadian tak terduga yang mengakibatkan objek keluar dari penelitian (Collet, 2004). Dalam analisis ketahanan hidup dikenal 3 jenis penyensoran data, yakni:

- Data Sensor Tipe I

Data sensor tipe I adalah percobaan yang dilakukan selama waktu T yang telah ditentukan dan akan berakhir setelah mencapai waktu T , berakhirnya waktu T menyatakan waktu tersensor. Pada penyensoran tipe I semua unit uji n masuk pada waktu yang bersamaan dan jika tidak terdapat individu yang hilang secara tiba-tiba maka waktu tahan hidup observasi tersensor sama dengan waktu selama observasi berlangsung.

- Data Sensor Tipe II

Data sensor tipe II adalah data waktu tahan hidup yang diperoleh setelah observasi mengalami kegagalan sebanyak r dari n individu yang ada. Pada data sensor tipe II, semua unit uji n masuk pada waktu yang bersamaan dan jika terdapat individu yang hilang secara tiba-tiba maka waktu tahan hidup observasi tersensor adalah waktu tahan hidup observasi tidak tersensor terbesar. Observasi akan berakhir jika telah ditemukan sejumlah r kegagalan.

- Data Sensor Tipe III

Data sensor tipe III adalah penelitian yang dilakukan untuk individu yang masuk dalam percobaan pada waktu yang berlainan. Percobaan akan berhenti setelah waktu yang telah ditentukan oleh peneliti berakhir

2.3. Fungsi Ketahanan Hidup

Fungsi ketahanan hidupnya dinotasikan dengan $S(t)$ yang menunjukkan probabilitas suatu individu bertahan hidup lebih dari waktu t , dimana $t > 0$.

$S(t)$ dapat didefinisikan sebagai berikut:

$$S(t) = P(\text{individu bertahan hidup lebih dari waktu } t)$$

$$S(t) = P(T > t)$$

$$\begin{aligned} S(t) &= 1 - P(\text{individu gagal atau mati sebelum waktu } t) \\ &= 1 - P(T \leq t) \end{aligned}$$

Fungsi ketahanan hidup atau $S(t)$ memiliki sifat sebagai berikut:

$$S(t) = \begin{cases} 1 & \text{untuk } t = 0 \\ 0 & \text{untuk } t = \infty \end{cases}$$

yang artinya, peluang individu dapat bertahan hidup pada waktu nol adalah 1 dan pada saat waktu tak terbatas seorang individu dapat bertahan hidup adalah nol.

2.4. Fungsi Kepadatan Peluang

Jika T merupakan variabel random dari waktu hidup suatu individu dalam interval $[0, \infty)$, maka fungsi kepadatan peluangnya adalah $f(t)$ dan fungsi distribusi kumulatifnya adalah $F(t)$. Waktu tahan hidup T memiliki fungsi kepadatan peluang yang didefinisikan sebagai peluang individu gagal pada interval waktu t sampai $t + \Delta t$ atau peluang kegagalan dalam interval per satuan waktu. Hal ini dapat dinyatakan sebagai

$$\begin{aligned} f(t) &= \frac{\lim_{\Delta t \rightarrow 0} P[\text{objek gagal pada interval } (t, t + \Delta t)]}{\Delta t} \\ &= \frac{\lim_{\Delta t \rightarrow 0} P(t < T < t + \Delta t)}{\Delta t} \end{aligned}$$

sehingga diperoleh

$$F(t) = P(T \leq t) = \int_0^t f(x) dx = 1$$

$$f(t) = \frac{dF(t)}{dt}$$

2.5. Fungsi Kegagalan

Fungsi kegagalan (*hazards function*) didefinisikan sebagai tingkat kegagalan bersyarat yaitu limit dari peluang suatu individu gagal bertahan dalam interval waktu yang sangat pendek dari t sampai $t + \Delta t$, jika individu tersebut telah bertahan hingga waktu t . Fungsi kegagalan $h(t)$ dari waktu tahan hidup T memiliki fungsi yang didefinisikan sebagai berikut :

$$\begin{aligned} h(t) &= \frac{\lim_{\Delta t \rightarrow 0} P \left\{ \begin{array}{l} \text{seorang individu gagal pada interval waktu } (t, t + \Delta t) \\ \text{jika diketahui individu tersebut telah bertahan hingga } t \end{array} \right\}}{\Delta t} \\ h(t) &= \frac{\lim_{\Delta t \rightarrow 0} P(t < T < t + \Delta t | T > t)}{\Delta t} \end{aligned}$$

Fungsi kumulatif dari fungsi kegagalan didefinisikan sebagai berikut:

$$H(t) = \int_0^t h(x) dx$$

Dari ketiga fungsi tersebut, yakni fungsi ketahanan hidup, fungsi kepadatan peluang dan fungsi kegagalan terdapat hubungan satu sama lainnya, yakni:

- $h(t) = \frac{f(t)}{1-F(t)} = \frac{f(t)}{S(t)}$
- $f(t) = \frac{dF(t)}{dt} = \frac{d(1-S(t))}{dt} = -S'(t)$
- $S(t) = \frac{f(t)}{h(t)} = 1 - F(t)$
- $h(t) = \frac{f(t)}{S(t)} = -\frac{S'(t)}{S(t)} = -\frac{d(\ln S(t))}{dt}$

2.6. Regresi Cox Proportional Hazards

Menurut Collett (2004), regresi *Cox Proportional Hazards* atau lebih dikenal sebagai model regresi Cox digunakan untuk mengetahui hubungan antara variabel dependen dengan variabel independen, dimana data yang digunakan pada regresi *Cox Proportional Hazards* berupa data waktu tahan hidup dari suatu individu. Menurut Collet (2004), model regresi *Cox Proportional Hazards* adalah sebagai berikut:

$$h_i(t|X) = h_0(t) \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi}) = h_0(t) e^{\sum_{j=1}^p \beta_j x_{ji}} \quad (1)$$

dengan

$h_i(t|X)$: Fungsi kegagalan individu $ke-i$

$h_0(t)$: Fungsi kegagalan dasar

x_{ji} : Nilai variabel $ke-j$ dari individu $ke-i$, dengan $j=1,2,\dots,p$ dan $i=1,2,\dots,n$

β_j : Koefisien regresi $ke-j$, dengan $j=1,2,\dots,p$

Fungsi kegagalan dalam model ini dapat mengambil bentuk apapun, tetapi fungsi kegagalan dari individu yang berbeda diasumsikan harus proporsional setiap waktu. Menurut Kleinbaum dan Klein (2005), apabila plot $\log\{-\log[S(t,x)]\}$ terhadap waktu ketahanan hidup antar kategori dalam satu variabel independen terlihat sejajar atau berhimpit maka asumsi *proportional hazards* terpenuhi dan variabel penjelas dapat dimasukkan ke dalam model.

2.7. Metode Maximum Likelihood Estimation (MLE)

Misalkan X_1, X_2, \dots, X_n adalah sampel acak dari populasi dengan fungsi kepadatan peluang $f(x_1, x_2, \dots, x_n | \theta)$ dimana x_1, x_2, \dots, x_n adalah himpunan data yang diamati, dan θ merupakan parameter yang dapat berbentuk tunggal namun juga dapat berbentuk vektor yaitu $\theta_j = \theta_1, \theta_2, \dots, \theta_k$ untuk $j=1,2,\dots,k$ dan fungsi likelihoodnya adalah $L(\theta_1, \theta_2, \dots, \theta_k | \tilde{X}) = \prod_{i=1}^n f(x_i | \theta_1, \theta_2, \dots, \theta_k)$. Untuk estimasi parameter menggunakan MLE terlebih dahulu mencari fungsi log dari fungsi likelihoodnya, dengan fungsi log likelihoodnya adalah

$$l = \log L(\theta_1, \theta_2, \dots, \theta_k | \tilde{X}) = \log(\prod_{i=1}^n f(x_i | \theta_1, \theta_2, \dots, \theta_k))$$

Turunan pertama dari fungsi likelihood terhadap $\theta_1, \theta_2, \dots, \theta_k$ merupakan calon estimator bagi $\theta_1, \theta_2, \dots, \theta_k$ dan hasil dari turunan pertama diperoleh dengan cara memaksimumkan atau menyamadengankan nol turunan fungsi log likelihood, yaitu

$$\frac{\partial \log(L(\theta_1, \theta_2, \dots, \theta_k | \tilde{X}))}{\partial \theta_j} = 0 \quad \text{untuk } j = 1, 2, \dots, k$$

Calon estimator bagi $\theta_1, \theta_2, \dots, \theta_k$ selanjutnya diuji untuk mengetahui bahwa calon estimator bagi $\theta_1, \theta_2, \dots, \theta_k$ benar-benar memaksimumkan fungsi likelihood. Untuk mengetahui bahwa hasil dari turunan pertama memaksimumkan fungsi likelihood harus ditunjukkan bahwa

$$\left(\frac{\partial^2 \log(L(\theta_1, \theta_2, \dots, \theta_k | \tilde{X}))}{\partial \theta_j^2} \Big|_{\theta_j = \hat{\theta}_j} \right) < 0$$

Jika terdapat data yang terdiri dari n waktu tahan hidup yang dinotasikan t_1, t_2, \dots, t_n dan δ_i adalah nilai indikator kejadian sebagai berikut:

$$\delta_i = \begin{cases} 0, & \text{individu yang tersensor} \\ 1, & \text{individu tidak tersensor} \end{cases} \text{ dengan } t_i, i=1,2,\dots,n$$

maka fungsi likelihood persamaan (1) dapat dinyatakan dalam

$$L(\boldsymbol{\beta}) = \prod_{i=1}^n \left[\frac{\exp(\boldsymbol{\beta}' \mathbf{x}_i)}{\sum_{l \in R(t_i)} \exp(\boldsymbol{\beta}' \mathbf{x}_l)} \right]^{\delta_i}$$

dan fungsi log likelihood yang bersesuaian adalah

$$\log L(\boldsymbol{\beta}) = \sum_{i=1}^n \delta_i [\boldsymbol{\beta}' \mathbf{x}_i - \log[\sum_{l \in R(t_i)} \exp(\boldsymbol{\beta}' \mathbf{x}_l)]]$$

dengan

\mathbf{x}_i : Vektor variabel dari individu yang gagal pada saat waktu $ke-i$.

$R(t_i)$: Himpunan individu yang masih hidup pada waktu $ke-i$.

\mathbf{x}_l : Vektor variabel individu yang masih hidup dan merupakan elemen dari $R(t_i)$

2.8 Rasio Kegagalan

Menurut Kleinbaum dan Klein (2011), rasio kegagalan adalah kegagalan untuk satu kelompok individu dibagi dengan kegagalan untuk kelompok individu lainnya. Rasio kegagalan dapat dinyatakan ke dalam bentuk seperti dibawah ini :

$$\widehat{HR} = \frac{\widehat{n}_0(t) e^{\sum_{j=1}^p \beta_j x_j^*}}{\widehat{n}_0(t) e^{\sum_{j=1}^p \beta_j x_j}} = \exp^{\sum_{j=1}^p \hat{\beta}_j (x_j^* - x_j)}$$

Dimana HR merupakan rasio kegagalan, $\mathbf{x}^* = (x_1^*, x_2^*, \dots, x_p^*)$ sebagai nilai variabel independen untuk satu kelompok individu, dan $\mathbf{x} = (x_1, x_2, \dots, x_p)$ menunjukkan nilai variabel independen untuk satu kelompok individu lain dan β adalah parameter regresi.

2.9 Taksiran Peluang

Fungsi kegagalan individu dapat ditaksir ketika nilai dari $h_0(t)$ telah diketahui, dengan

$$h_0(t) = 1 - \xi_m$$

dengan ξ_m didefinisikan sebagai berikut :

$$\xi_m = \left[\frac{-d_m}{\sum_{l \in R(t(m))} \exp(\boldsymbol{\beta}' \mathbf{x}_l)} \right]$$

digunakan jika hanya ada satu individu yang gagal pada waktu ke $t(m)$.

$$\xi_m = \left[1 - \frac{\exp(\boldsymbol{\beta}' \mathbf{x}(m))}{\sum_{l \in R(t(m))} \exp(\boldsymbol{\beta}' \mathbf{x}_l)} \right]^{\exp(-\boldsymbol{\beta}' \mathbf{x}(m))}$$

digunakan jika dua atau individu yang gagal pada waktu ke $t(m)$. Sehingga diperoleh persamaan untuk mencari nilai fungsi kegagalan dasar kumulatif dan nilai fungsi ketahanan hidup dasar adalah.

$$H_0(t) = -\sum_{m=1}^r \log \xi_m$$

$$S_0(t) = \prod_{m=1}^r \xi_m$$

$$S(t) = [S_0(t)]^{\exp(\boldsymbol{\beta}' \mathbf{x}(m))}$$

III. METODOLOGI PENELITIAN

3.1 Jenis dan Sumber Data

Populasi dalam penelitian ini adalah seluruh mahasiswa S1 Fakultas Sains dan Matematika Universitas Diponegoro angkatan 2009 yang berjumlah 383 mahasiswa, sedangkan sampelnya adalah 100 mahasiswa S1 Fakultas Sains dan Matematika Universitas Diponegoro angkatan 2009 yang diperoleh dengan menggunakan bilangan acak yang diperoleh secara komputerisasi dengan bantuan Microsoft Excel. Jenis data yang digunakan dalam penelitian ini adalah data sekunder dan primer. Lokasi dari penelitian ini adalah di Fakultas Sains dan Matematika Universitas Diponegoro Semarang, Jawa Tengah. Data sekunder yang dimaksud dalam penelitian ini adalah data yang diperoleh dari buku “*Wisuda ke-130, Wisuda ke-131 dan Wisuda ke-132*”. Untuk data primer diperoleh dengan cara melakukan wawancara secara langsung, kusioner dan melalui media elektronik (*Telephone* dan *e-mail*).

3.2 Variabel Penelitian

Variabel dependen pada penelitian ini adalah lama studi mahasiswa angkatan 2009 Fakultas Sains dan Matematika Universitas Diponegoro. Dikatakan bertahan hidup (tersensor) apabila mahasiswa tersebut belum lulus lebih dari waktu 48 bulan dan dikatakan mengalami kegagalan (tidak tersensor) apabila mahasiswa tersebut telah lulus sampai dengan waktu 48 bulan. Sedangkan variabel independennya adalah variabel yang dikategori menjadi dua atau lebih kategori yakni Jenis kelamin (laki-laki=1 dan perempuan=0), Jurusan (Matematika=0, Biologi=1, Kimia=2, Fisika=3, Statistika=4 dan Informatika=5), IPK ($IPK < 3=0$; $3 \leq IPK \leq 3,5=1$; $IPK > 3,5=2$), Pekerjaan orang tua (PNS=0 atau tidak PNS=1), Asal (dari daerah pulau Jawa=0 atau luar pulau Jawa=1), Jalur masuk ke Universitas (Mandiri=0, PSSB=1, SNMPTN=2), Beasiswa (Pernah=1 atau tidak menerima beasiswa=0), Organisasi (Pernah=1 atau tidak mengikuti organisasi=0), *Part time* (pernah=1 atau tidak kerja sampingan=0).

IV. PEMBAHASAN

Analisis Deskriptif

Jumlah ukuran sampel penelitian ini adalah 100 mahasiswa dengan persentase 56 (56%) diantaranya adalah mahasiswa yang lulus sampai dengan waktu 48 bulan (status tidak tersensor), sedangkan sisanya yakni 44 (44%) merupakan mahasiswa belum lulus lebih dari waktu 48 bulan (status tersensor).

Model Regresi Cox Proportional Hazards

Untuk asumsi Kegagalan proporsional, seluruh variabel independen pada penelitian yakni Jenis kelamin(JK), Jurusan(J), IPK, Pekerjaan orang tua(POT), Asal(A), Jalur masuk(JM), Beasiswa(B), Organisasi(O), *Part time*(Pt).

Berdasarkan Pengolahan data yang dilakukan dengan bantuan SPSS 17, maka diperoleh parameter (β) dari setiap variabel. Selanjutnya ialah melakukan pengujian parameter baik secara serentak maupun parsial. Hasil yang diperoleh dari pengujian secara parsial adalah secara bersama-sama variabel IPK, Jurusan dan Organisasi memberikan pengaruh terhadap lama studi mahasiswa, Hal tersebut dikarenakan $\chi^2_{LR} = 37,474 > \chi^2_{8;0,05} = 15,507$ dan nilai sig (0,000) < α (0,05).

Sedangkan untuk hasil pengujian secara Parsial dapat dilihat pada Tabel 1, dimana variabel yang memberikan pengaruh terhadap lama studi mahasiswa ialah IPK, Jurusan dan Organisasi hal tersebut dikarenakan nilai dari sig < $\alpha = 0,05$ dan nilai $\chi^2_W > \chi^2_{1;0,05} = 3,84$.

Tabel 1. Pengujian Secara Parsial

Variabel	Koefisien Regresi	Standar error	Wald	Derajat bebas	Sig.	Keputusan
J(1)	-0.292	0.393	0.551	1	0.458	Terima Ho
J(2)	-0.346	0.369	0.881	1	0.348	Terima Ho
J(3)	-0.656	0.517	1.613	1	0.204	Terima Ho
J(4)	-0.835	0.475	3.095	1	0.079	Terima Ho
J(5)	-3.077	1.039	8.773	1	0.003	Tolak Ho
O	-0.964	0.45	4.599	1	0.032	Tolak Ho
IPK(1)	2,003	1.024	3.829	1	0.05	Terima Ho
IPK(2)	2,705	1.045	6.703	1	0.01	Tolak Ho

Sehingga diperoleh model akhir dari penelitian ini adalah :

$$h_i(t|X) = h_0(t) \exp \{ -0,292J(1) - 0,346J(2) - 0,656J(3) - 0,835J(4) - 3,077J(5) - 0,964 O + 2,003IPK(1) - 2,705IPK(2) \}$$

Rasio Kegagalan

Rasio kegagalan adalah resiko kegagalan untuk satu kelompok individu dibagi dengan kegagalan untuk kelompok individu lainnya. Tabel 2 merupakan nilai dari rasio kegagalan untuk setiap variabel independen yang memberikan pengaruh terhadap lama studi mahasiswa.

Tabel 2. Perhitungan rasio kegagalan untuk variabel yang signifikan

Variabel	Koefisien Regresi	Exp(B)
J(1)	-0,292	0,747
J(2)	-0,346	0,707
J(3)	-0,656	0,519
J(4)	-0,835	0,434
J(5)	-3,077	0,046
O	-0,964	0,381
IPK(1)	2,003	7,414
IPK(2)	2,705	14,959

Rasio kegagalan digunakan untuk membandingkan kecenderungan peluang kegagalan individu dengan pada kategori setiap peubahnya. Rasio kegagalan J(5) adalah 0,046 yang memiliki arti bahwa kecenderungan peluang mahasiswa Jurusan Informatika atau J(5) memiliki peluang lulus tepat waktu sebesar 0,046 kalinya dari mahasiswa Jurusan Matematika atau J(0). Untuk rasio kegagalan IPK(2) dapat diartikan bahwa kecenderungan mahasiswa dengan IPK(2) atau $IPK > 3,5$ memiliki peluang lulus tepat waktu sebesar 14,959

kalinya dari mahasiswa dengan $IPK(0)$ atau $IPK < 3$. Untuk variabel lainnya dapat dilihat pada Tabel 2.

Taksiran peluang

Nilai dari $h_0(t)$, $S_0(t)$ yang terdapat pada Tabel 3 akan digunakan untuk estimasi peluang mahasiswa dapat lulus tepat waktu sampai dengan 48 bulan dan peluang mahasiswa belum lulus lebih dari 48 bulan .

Tabel 3. Tabel estimasi peluang dasar

m	Lama Studi	d_m	ξ_m	$h_0(t)$	$H_0(t)$	$S_0(t)$
1	40	1	0.99396	0.00604	0.00606	0.99396
2	41	1	0.99384	0.00616	0.01224	0.98784
3	43	1	0.99365	0.00635	0.01861	0.98157
4	44	6	0.96183	0.03817	0.05752	0.9441
5	45	5	0.96254	0.03746	0.0957	0.90874
6	46	18	0.85571	0.14429	0.25152	0.77762
7	47	11	0.85001	0.14999	0.41403	0.66098
8	48	12	0.74721	0.25279	0.70543	0.4939

Untuk nilai $h_0(t)$ akan bernilai sama dengan nilai $h_i(t)$ apabila seluruh variabel independen yang terdapat pada model bernilai 0. Berdasarkan $h_0(t)$ dan $S_0(t)$ yang telah diperoleh sebelumnya, maka dapat dihitung peluang mahasiswa lulus tepat waktu sampai dengan 48 bulan dan peluang mahasiswa yang belum lulus lebih dari 48 bulan.

- Untuk waktu lama studi $t_1=40$

$$S(40|J(5)) = S_0(40)^{\exp(-3,077*1)} = 0,99972$$

$$F(40|J(5)) = 1 - 0,99972 = 0,00028$$

Artinya adalah bahwa peluang mahasiswa Jurusan Informatika atau $J(5)$ masih belum lulus lebih dari waktu 40 bulan adalah sebesar 0,99972 dan peluang mahasiswa Jurusan Informatika atau $J(5)$ yang telah lulus sampai dengan waktu 40 bulan adalah sebesar 0,00028.

- Untuk waktu lama studi $t_2=41$

$$S(41|J(4), O) = S_0(41)^{\exp((-0,964*1)+(-0,835*1))} = 0,99798$$

$$F(41|J(4), O) = 1 - 0,99798 = 0,00202$$

Artinya adalah peluang mahasiswa Jurusan Statistika dan yang berorganisasi belum lulus lebih dari waktu 41 bulan adalah sebesar 0,99798, sedangkan peluang mahasiswa jurusan Statistika dan yang berorganisasi telah lulus sampai dengan waktu 41 bulan adalah sebesar 0,00202. Berdasarkan perhitungan peluang yang telah dilakukan, maka dapat disimpulkan bahwa semakin lama waktu studi mahasiswa maka peluang mahasiswa tersebut belum lulus $S(t|X)$ semakin kecil dan peluang mahasiswa tersebut telah lulus $F(t|X)$ semakin besar.

V. KESIMPULAN

Berdasarkan hasil analisis penelitian tentang faktor-faktor yang mempengaruhi terhadap lama studi mahasiswa S1 angkatan 2009 Fakultas Sains dan Matematika Universitas Diponegoro, maka dapat ditarik kesimpulan bahwa :

1. Setelah diuji secara parsial dan serentak, didapat bahwa faktor-faktor yang mempengaruhi lama studi mahasiswa adalah J(4) atau Jurusan Statistika, J(5) atau Jurusan Informatika, O atau Organisasi, dan IPK(2) atau $IPK > 3,5$. Jadi dapat disimpulkan bahwa variabel Jurusan, Organisasi, dan IPK berpengaruh terhadap lama studi mahasiswa. Model akhir Regresi Cox *Proportional Hazards* pada penelitian ini adalah
$$h_i(t|X) = h_0(t) \exp \{ -0,292J(1) - 0,346J(2) - 0,656J(3) - 0,835J(4) - 3,077J(5) - 0,964 O + 2,003IPK(1) - 2,705IPK(2) \}$$
2. Rasio kegagalan dihitung hanya untuk variabel yang memberikan pengaruh terhadap lama studi yakni variabel Jurusan, Organisasi dan IPK dimana rasio kegagalan suatu kategori variabel dibandingkan dengan kategori dasar dari variabel tersebut.
3. Jika semakin lama masa studi seorang mahasiswa maka dugaan peluang mahasiswa tersebut belum lulus $S(t|X)$ semakin kecil dan peluang mahasiswa tersebut telah lulus $F(t|X)$ semakin besar.

DAFTAR PUSTAKA

- Bain, L. J & Engelhardt. 1992. *Introduction to Probability and Mathematical Statistics*. Duxburry Press. California.
- Collet, D. 2004. *Modelling Survival Data in Medical Research*. CRC Press.
- Kalbfleisch, J. D & Prentice, R. L. 2002. *The Statistical Analysis of Failure Time Data*. John Wiley & Sons, Inc. Canada.
- Kementerian Pendidikan dan Kebudayaan RI. 2012. *Undang-Undang Tentang Pendidikan Tinggi*. Kementerian Pendidikan dan Kebudayaan RI. Jakarta.
- Klein, John. P and Moeschberger, Melvin. L. 2003. *Survival Analysis Techniques for Censored and Truncated Data*. Springer. New York.
- Kleinbaum, D. C. & Klein, M. 2005. *Survival Analysis A Self-Learning Text*. Springer. New York.
- Kleinbaum, D. C. & Klein, M. 2011. *Survival Analysis*. Springer. New York.
- Lawless, F. 2003. *Statistical Models and Methods for Life Time Data*. John Wiley & Sons, Inc. New York.
- Lee, E. T. 1992. *Statistical Methods for Survival Data Analysis*. John Wiley & Sons, Inc. Canada.
- Lee, E. T & Wang, J. W. 2003. *Statistical Methods for Survival Data Analysis*. John Wiley & Sons, Inc. Canada.
- Scheaffer, Richard. L, dkk. *Elementary Survey Sampling*. PWS-KENT. Boston.
- Wienke, Anrdeas. 2011. *Frailty Models in Survival Analisis*. Taylor and Francis Group. New York.