

IMPLEMENTASI *COX PROPORTIONAL HAZARD* MODEL PARAMETRIK PADA ANALISIS SURVIVAL (Studi Kasus: Mahasiswa Universitas Internasional Batam)

Yayuk Setyaning Astutik¹, Dian Tresnawan²

¹Universitas Internasional Batam, yayuk@uib.ac.id

²Universitas Internasional Batam, dian.tresnawan@uib.ac.id

Abstract There are several factors that affect the study period of students at the International University of Batam (UIB) such as GPA, Parents Earnings, School Origin, and others. Therefore, it is necessary to do research to find out the factors that significantly affect the study period of students. The method used in this survival analysis is the cox proportional hazard regression parametric model that can explain the influence of independent factors in an event. From the process of survival analysis with regression cox proportional hazard parametric model obtained a significant factor influencing the study period of students of Faculty of Engineering, International University of Batam who entered in the year 2013 is GPA with model $h_i(t) = \exp(-2.089X_6(1) - 1.155X_6(2) + 0X_6)h_0(t)$ and obtained the highest probabilities of students who graduated in the 8th semester with GPA > 3,723 is 0,73; and the highest probabilities of students who graduate more than 8 semesters with GPA > 3,723 has a probability 0,916.

Keywords : *Survival, Cox Proportional Hazard, Study Period, Student, Parametric Model.*

Abstrak Ada beberapa faktor yang mempengaruhi masa studi mahasiswa di Universitas Internasional Batam (UIB) seperti IPK, Penghasilan Orang Tua, Asal Sekolah, dan lain-lain. Oleh karena itu, perlu dilakukan penelitian untuk mengetahui faktor-faktor apa saja yang signifikan mempengaruhi masa studi mahasiswa. Metode yang digunakan dalam analisis survival ini adalah *regresi cox proportional hazard* model parametrik yang dapat menjelaskan pengaruh faktor independen dalam suatu kejadian. Dari proses analisis survival dengan metode *regresi cox proportional hazard* model parametrik diperoleh faktor yang signifikan mempengaruhi lama studi mahasiswa Fakultas Teknik Universitas Internasional Batam angkatan 2013 adalah IPK. Dari proses analisis survival dengan *regresi cox proportional hazard* model parametrik diperoleh faktor yang signifikan mempengaruhi masa studi mahasiswa Fakultas Teknik Universitas Internasional Batam yang masuk pada tahun 2013 adalah IPK dengan model $h_i(t) = \exp(-2.089X_6(1) - 1.155X_6(2) + 0X_6)h_0(t)$ serta diperoleh peluang tertinggi mahasiswa yang lulus pada semester 8 dengan IPK > 3,723 sebesar 0,73; dan peluang tertinggi mahasiswa yang lulus lebih dari 8 semester dengan IPK > 3,723 memiliki peluang sebesar 0,916.

Kata kunci: *Survival, Cox Proportional Hazard, Masa Studi, Mahasiswa, Model Parametrik*

1 Pendahuluan

Pada dasarnya setiap Perguruan Tinggi berusaha semaksimal mungkin untuk meningkatkan mutu kelulusan para mahasiswanya, baik secara kuantitas maupun kualitas. Secara kuantitas diharapkan jumlah mahasiswa yang lulus sama dengan yang terdaftar, artinya tidak ada mahasiswa yang DO atau mengundurkan diri. Sedangkan

secara kualitas diharapkan para mahasiswa dapat lulus dengan IPK yang tinggi dan masa studi tepat waktu.

Proses seleksi penerimaan mahasiswa baru untuk tingkat sarjana dilakukan 10 kali gelombang dalam satu tahun. Setiap tahun jumlah mahasiswa yang masuk selalu mengalami kenaikan untuk diseluruh program studi. Dari tahun 2000 sampai dengan 2016 Universitas Internasional Batam telah melaksanakan wisuda sebanyak 12 kali dengan total wisudawan 3434 dari jenjang D-3, S-1 dan S-2.

Kualitas lulusan dari Perguruan Tinggi dapat dipengaruhi oleh beberapa faktor, baik internal maupun eksternal. Faktor internal merupakan faktor yang berasal dari diri lulusan tersebut, seperti kecerdasan, kemampuan belajar, tingkat kesibukan mahasiswa, dan lain-lain. Sedangkan faktor eksternal merupakan faktor yang berasal dari luar diri mahasiswa, seperti kondisi lingkungan, besar dukungan orang tua, prasarana dan sarana yang dimiliki, dan lain-lain.

Faktor-faktor internal dan eksternal diduga berpengaruh terhadap kemampuan seorang mahasiswa khususnya pada penelitian ini Program Sarjana (S1) yang dapat menyelesaikan studi tidak lebih dari 4 tahun. Oleh karena itu, penelitian ini dilakukan untuk mengetahui faktor-faktor yang mempengaruhi lama studi mahasiswa Fakultas Teknik Universitas Internasional Batam khususnya angkatan 2013. Alasan peneliti mengambil sampel mahasiswa Fakultas Teknik dikarenakan dapat dilihat dilapangan bahwa jumlah kelulusan mahasiswa Universitas Internasional Batam tiap tahunnya tidak seimbang dengan jumlah yang masuk, dan meskipun banyak mahasiswa yang telah menyelesaikan studinya tepat waktu sampai 8 semester bahkan 7 semester, namun ada sebagian juga mahasiswa yang belum menyelesaikan studinya sampai semester 9.

2 Kajian Teori

2.1 Penelitian Terdahulu

Estimasi parameter pada variabel dependen-variabel prediktor menggunakan analisis Regresi Cox diperoleh hasil mengenai rata-rata waktu hidup, bentuk fungsi kegagalan dan ketahanan hidup serta peluang keberhasilan dan kegagalannya dari Pasien Kanker Paru di RSUP Dr. Kariadi Semarang [6].

Regresi kegagalan proporsional dari Cox adalah model dalam analisis ketahanan hidup yang digunakan untuk mengetahui besarnya hubungan antara satu atau lebih variabel independen dan dependennya. Model ini tidak memerlukan informasi tentang distribusi yang mendasari, tetapi fungsi kegagalan dari individu yang berbeda diasumsikan proporsional dan hasilnya faktor Usia adalah faktor yang mempengaruhi laju kesembuhan pasien rawat inap yang menderita demam tifoid [5].

Metode yang cukup baik untuk digunakan dalam analisis survival adalah Cox dan Weibull Proportional Hazard. Hasil analisis menunjukkan bahwa semua peubah penjelas yang berpengaruh nyata pada model Cox juga berpengaruh terhadap model Weibull. Model Weibull merupakan model yang terbaik untuk menganalisa daya tahan pemakai pil KB untuk setiap provinsi karena memiliki nilai AIC yang lebih kecil dibandingkan model Cox dengan laju pertumbuhan penduduk yang berbeda [1].

Model Cox Proportional hazard Model pada kejadian bersma merupakan modifikasi dari model Cox ketika ada dua atau lebih individu yang mengalami kejadian bersama (*ties*). Dengan melakukan modifikasi estimasi parameter dengan pendekatan Breslow dapat diaplikasikan pada kasus transportasi dan lalu lintas. Hasilnya diperoleh

Faktor yang mempengaruhi kecelakaan lalu lintas di USA adalah Umur Pengemudi dan Penggunaan Sabuk Pengaman [2].

Selain itu, analisis survival juga bisa digunakan untuk mengetahui probabilitas *student dropout*. Tingkat *Student Dropout* dikaitkan dengan Fakultas atau matakuliah yang diambil, jenis kelamin dan nilai yang diperoleh. Hasilnya selama tahun akademik pertama sebesar 26% meninggalkan Universitas dan *Student Dropout* tertinggi ada di Fakultas Science Engineering tahun 2012-2014 di Latvia University of Agriculture [4].

2.2 Analisis Survival

Analisis survival telah menjadi alat penting untuk menganalisis data waktu antar kejadian (*time to event data*) atau menganalisis data yang berhubungan dengan waktu, mulai dari *time origin* sampai terjadinya suatu peristiwa khusus. Waktu awal (*time origin*) yaitu pada saat terjadinya kejadian awal, seperti waktu seseorang menderita kanker, waktu pemberian perlakuan, waktu awal studi dan lain-lain. Waktu kegagalan (*failure time*) adalah waktu pada saat terjadinya kejadian akhir seperti kematian, *dropout*, kelulusan dan lainnya. Waktu survival yang dicatat adalah satu dari kriteria berikut (1) Selisih waktu mulai dilakukan pengamatan sampai waktu terjadinya kegagalan atau disebut data tidak tersensor (*uncensored data*) dan (2) Jika waktu akhir tidak diketahui, maka selisih waktu mulai dilakukannya pengamatan sampai waktu akhir penelitian, disebut data tersensor (*censored data*).

2.3 Penyensoran

Menurut Johnson dalam [3] ada tiga yakni:

1. Penyensoran Tipe I

Sensor tipe I adalah tipe penyensoran dimana percobaan akan diberhentikan setelah mencapai waktu T yang telah ditentukan untuk mengakhiri semua n individu yang masuk pada waktu yang sama. Pada umumnya untuk setiap T_i diberikan waktu sensor L_i . Semua komponen dikatakan terobservasi jika $T_i < L_i$ diperoleh variabel waktu dan variabel yang menunjukkan semua komponen telah mati yaitu t_i dan δ_i dengan demikian, fungsi kepadatan peluangnya adalah $f(t_i)^{\delta_i} S(L_i)^{1-\delta_i}$.

2. Penyensoran Tipe II

Pada penyensoran tipe II, pengamatan diakhiri dengan sejumlah kegagalan yang telah ditetapkan diperoleh, atau dapat dikatakan banyaknya kegagalan adalah tetap dan waktu pengamatan adalah acak. Secara formal, data terdiri dari r terkecil waktu hidup $T_{(1)} \leq T_{(2)} \leq \dots \leq T_{(r)}$ dari sampel acak n waktu hidup T_1, T_2, \dots, T_n dari distribusi hidup. Jika T_1, T_2, \dots, T_n *i.i.d* dan memiliki distribusi kontinu dengan *f.k.p* $f(t)$ dan fungsi survival $S(t)$, maka hasil umum pada urutan statistik *f.k.p* bersama dari $T_{(1)}, T_{(2)}, \dots, T_{(r)}$ adalah

$$\frac{n!}{(n-r)!} f(t_{(1)}) \dots f(t_{(r)}) [S(t_{(r)})]^{n-r}$$

3. Penyensoran Maju (*Progressive Censoring*)

Pada penyensoran maju, suatu jumlah ditentukan dari unit-unit bertahan dikeluarkan dari penelitian berdasarkan kejadian dari tiap kegagalan terurut. Secara konseptual, hal ini sama dengan suatu praktek yang dikenal sebagai *sudden-death testing*, dimana tes secara serempak memuat beberapa pengetesan dan apabila terjadi kegagalan pertama keseluruhan pengetesan dianggap gagal.

2.4 Fungsi Survival

Fungsi *survival* adalah probabilitas bahwa suatu individu akan tetap hidup sampai waktu t ($t > 0$). Jika u variabel *random* yang menotasikan waktu bertahan hidup dari seorang individu, maka $S(t)$ adalah probabilitas bahwa T lebih besar dari t . Dalam statistik fungsi kumulatif distribusi $F(t)$ dengan $t > 0$ didefinisikan:

$$F(t) = \Pr(T \leq t) = \int_{-\infty}^t f(u) du \quad (1)$$

Fungsi *hazard* dari waktu ketahanan hidup T merupakan *Probabilitas Bersyarat* bahwa individu akan mati pada interval $[t, t + \Delta t]$ jika diketahui tetap hidup sampai saat t adalah :

$$\frac{F(t + \Delta t) - F(t)}{S(t)} \quad (2)$$

Probabilitas kepadatan adalah estimasi probabilitas gagal pada rentang interval, dihitung pada masing-masing waktu yaitu :

$$F_i = \frac{(P_i - P_{i+1})}{h_i} \quad (3)$$

Langkah untuk mengklasifikasikan observasi data ke *Life Table* yaitu dengan mengestimasi *survival function* langsung dari waktu *continuous survival* atau gagal. Dengan mengkalikan probabilitas *survival* pada interval-interval akan kita dapatkan *survival function* sebagai berikut :

$$S(t) = \prod_{j=1}^t \left[\frac{(n-j)}{(n-j+1)} \right]^{o_{(j)}} \quad (4)$$

Metode yang digunakan untuk melihat *survival* di dalam suatu grup adalah *Log-rank test*. Caranya adalah dengan membandingkan estimasi *hazard function* dari grup yang diobservasi dalam waktu tertentu. *Log-rank test* ini dapat dinotasikan sebagai berikut :

$$E_j = O_j \frac{N_{1j}}{N_j} \quad \text{dan} \quad V_j = \frac{O_j \left(\frac{N_{1j}}{N_j} \right) \left(1 - \frac{N_{1j}}{N_j} \right) (N_j - O_j)}{N_j - 1} \quad (5)$$

Log-rank statistik membandingkan setiap O_{1j} dengan ekspektasi E_j dalam *null hypothesis* sehingga dapat dinyatakan sebagai berikut :

$$Z = \frac{\sum_{j=1}^J (O_{1j} - E_j)}{\sqrt{\sum_{j=1}^J V_j}} \quad (6)$$

dengan *log-rank test* akan didapat *hazard* dan *ratio* grup di dalam masing-masing *covariate* dan akan diketahui grup mana yang mempunyai *hazard* dan risiko yang terbesar dan terkecil.

2.5 Fungsi Kegagalan (Hazard Function)

Fungsi kegagalan didefinisikan sebagai laju kegagalan dari suatu individu pada selang waktu yang pendek $[t, \Delta t]$ untuk mampu bertahan setelah melewati waktu yang ditetapkan yaitu t . Misalkan T variabel *random non negative* pada interval $[0, \infty]$ yang menunjukkan waktu individu sampai mengalami kejadian pada suatu populasi, maka

peluang bahwa individu mengalami kejadian pada interval $(t, t + \Delta t)$. Berikut ini persamaan fungsi hazard $h(t)$ adalah:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = \frac{F'(t)}{S(t)} = \frac{f(t)}{S(t)} \quad (7)$$

2.6 Maximum Likelihood Estimation (MLE)

Misalkan X_1, X_2, \dots, X_n adalah sampel *random* dari populasi dengan fungsi kepadatan peluang $f(x|\theta_1, \theta_2, \dots, \theta_m)$, *fungsi likelihood* didefinisikan dengan :

$$L(\theta_1, \theta_2, \dots, \theta_m | X) = \prod_{i=1}^n f(x_i | \theta_1, \theta_2, \dots, \theta_m) \quad (8)$$

Bila fungsi *likelihood* dideferensialkan dalam $\theta_1, \theta_2, \dots, \theta_m$, maka calon estimator *maximum likelihood* yang mungkin adalah harga $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m$ sedemikian sehingga :

$$\frac{\partial}{\partial \theta_k} L(\theta_1, \theta_2, \dots, \theta_m | X) = 0 \quad (9)$$

Untuk membuktikan bahwa $\hat{\theta}_k$ benar-benar memaksimumkan fungsi *likelihood*, $L(\theta_1, \theta_2, \dots, \theta_m | X)$ harus ditunjukkan bahwa

$$\frac{\partial^2}{\partial \theta_k^2} L(\theta_1, \theta_2, \dots, \theta_m | X) |_{\theta_k} < 0 \quad (10)$$

Differensial digunakan pada *logaritma natural* dari $L(\theta_1, \theta_2, \dots, \theta_m | X)$, yaitu $l = \log L(\theta_1, \theta_2, \dots, \theta_m | X)$. Hal ini dimungkinkan karena fungsi *log* naik tegas pada $(0, \infty)$, yang berarti bahwa $L(\theta_1, \theta_2, \dots, \theta_m | X)$ dan $l = \log L(\theta_1, \theta_2, \dots, \theta_m | X)$ mempunyai ekstrim yang sama.

2.7 Model Cox Proportional Hazard

Salah satu model regresi yang sering digunakan adalah regresi *Cox Proportional Hazard*. Bentuk model *Cox Proportional Hazard* adalah :

$$h(t, X) = h_0(t) \exp \sum_{y=1}^m \beta_y X_y, \quad X = (X_1, X_2, \dots, X_m) \quad (11)$$

Model ini menyatakan *hazard rate* dari satu individu pada waktu t dengan diketahui *covariate* X . Fungsi *baseline hazard* adalah *hazard rate* saat $X = 0 \cdot h_0(t)$ merupakan fungsi yang tidak diketahui karena distribusi dari *survival time* (T) tidak diketahui. Fungsi ini hanya bergantung waktu t dan tidak mengandung X . Kuantitas ini hanya bergantung pada X yang disebut *time independent covariate*. Hal ini dikarenakan X tidak bergantung pada waktu. Jika X bergantung pada waktu, maka X disebut *time independent covariate* yaitu variabel yang nilainya tidak akan berubah sepanjang waktu.

Sifat dari model *Cox Proportional Hazard* adalah meskipun kita tidak mengetahui bentuk $h_0(t)$, akan tetapi kita tetap dapat menaksir koefisien regresi (β). Dalam menaksir β untuk mengetahui efek dari *covariate*-nya. Besarnya efek ini dapat dihitung tanpa menaksir fungsi *baseline hazard*. Jadi, dengan asumsi yang terbatas, kita dapat mengetahui informasi penting yang diperoleh dari data *survival* melalui nilai *hazard ratio* dan *survival experience*.

3 Metodologi Penelitian

3.1 Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yaitu data lama studi mahasiswa (dalam semester) program sarjana Universitas Internasional Batam angkatan 2013. Data sekunder tersebut diperoleh dari Biro Administrasi Akademik Kemahasiswaan, Biro Humas dan Biro *Quality Assurance Center* (QAC) Universitas Internasional Batam.

3.2 Populasi dan Sampel

Populasi dalam penelitian ini adalah seluruh mahasiswa program sarjana Universitas Internasional Batam angkatan 2013. Sampel dalam penelitian ini adalah mahasiswa Fakultas Teknik angkatan 2013 yang memiliki data yang lengkap dengan tujuan penelitian yang ingin dicapai.

3.3 Metode Pengumpulan Data

Metode pengumpulan data yang dilakukan dalam penelitian ini adalah metode dokumentasi yaitu data yang dianalisis diperoleh dari Biro Administrasi Akademik Kemahasiswaan, Biro Humas dan Biro *Quality Assurance Center* (QAC) Universitas Internasional Batam yaitu berupa data lama studi mahasiswa program sarjana Fakultas Teknik angkatan 2013 baik yang sudah dinyatakan lulus maupun yang belum sampai Agustus 2017, Indeks Prestasi Kumulatif (IPK), Jenis Kelamin, Asal Daerah Mahasiswa, Penghasilan Orang Tua, Pekerjaan Orang Tua, dan Status Sekolah Asal.

4 Hasil dan Pembahasan

4.1 Model Regresi *Cox Proportional Hazard*

Pengujian secara serentak digunakan untuk menentukan apakah peubah-peubah atau variabel penjelas pada model regresi *hazard proportional* berpengaruh secara bersama-sama terhadap respon. Dengan software R, diperoleh model baru *cox* regresi sebagai berikut:

$$\begin{aligned} h_i(t) = \exp(& 0.043X_1 - 0.177X_2(1) + 0X_2 - 0.0093X_3 - 0.266X_4(1) \\ & - 0.0275X_4 + 0X_4 - 0.266X_5(1) - 0.275X_5(2) + 0X_5 \\ & - 2.191X_6(1) - 1.164X_6(2) + 0X_6) h_0(t) \end{aligned}$$

Untuk pengujian parsial atau *Uji Wald* digunakan untuk melihat apakah masing-masing peubah atau variabel penjelas pada model berpengaruh terhadap respon diperoleh hasil sebagai berikut:

Tabel 1. Hasil Pengujian Parsial Model Awal

Variabel	β_j	$Se(\beta_j)$	χ^2_w	Sig
Jenis Kelamin (X_1)	0.043	0.258	0.027	0.869
Asal Daerah (X_2)			0.196	0.907
Asal Daerah (X_{2_1})	-0.0177	0.436	0.164	0.685
Asal Daerah (X_{2_2})	-0.154	0.381	0.163	0.686
Status Sekolah (X_3)	-0.009	0.058	0.000	0.987
Pekerjaan Ortu (X_4)			1.136	0.567
Pekerjaan Ortu (X_{4_1})	-0.266	0.291	0.838	0.361
Pekerjaan Ortu (X_{4_2})	-0.275	0.284	0.983	0.333
IPK (X_6)			13.75	0.001
IPK (X_{6_1})	-2.191	0.636	11.85	0.01
IPK (X_{6_2})	-1.164	0.371	9.838	0.02

Dari Tabel 1 diketahui bahwa variabel Jenis Kelamin (X_1), Asal Daerah (X_2), Status Sekolah (X_3), dan Pekerjaan Ortu (X_4). Selanjutnya dilakukan analisis ulang untuk pengujian serentak pada model terbaik dan pengujian parsial pada model terbaik menggunakan variabel yang signifikan sesuai Tabel 1 yaitu variabel IPK (X_6).

4.2 Pengujian Model Baru

Pengujian secara serentak digunakan untuk menentukan apakah peubah-peubah atau variabel penjelas pada model regresi *hazard proportional* berpengaruh secara bersama-sama terhadap respon. Dengan software R, diperoleh model baru *cox* regresi sebagai berikut:

$$h_i(t) = \exp(-2.089X_6(1) - 1.155X_6(2) + 0X_6)h_0(t)$$

Untuk pengujian model parsial atau dengan Uji Wald diperoleh hasil sebagai berikut.

Tabel 2. Hasil Pengujian Parsial Model Baru

Variabel	β_j	$Se(\beta_j)$	χ^2_w	Sig
IPK X_6			16.768	0.000
IPK $X_6(1)$	-2.089	0.599	13.936	0.000
IPK $X_6(2)$	-1.155	0.334	11.971	0.001

Dari perhitungan pada Tabel 2 diketahui bahwa faktor IPK signifikan dalam model terbaik. Dengan demikian dapat diartikan bahwa faktor yang mempengaruhi lama studi mahasiswa fakultas teknik Universitas Internasional Batam 2013 adalah **IPK**. Pemodelan regresi untuk lama studi mahasiswa Fakultas Teknik di Universitas Internasional Batam adalah IPK dengan model sebagai berikut:

$$h_i(t) = \exp(-2.089X_6(1) - 1.155X_6(2) + 0X_6)h_0(t)$$

dengan $X_6(1) = IPK < 3.088$; $X_6(2) = 3.088 \leq IPK \leq 3.723$ dan $X_6(1) = IPK > 3.723$

4.3 Rasio Kegagalan

Setelah diperoleh model kegagalan proporsional selanjutnya menghitung rasio kegagalan untuk variabel yang signifikan berdasarkan kategorinya. Misal pada variabel IPK, mahasiswa yang memiliki $IPK < 3.088$ atau dikategorikan $IPK = 0$ dan responden yang memiliki $3.088 \leq IPK \leq 3.723$ atau dikategorikan $IPK^* = 1$, sehingga diperoleh nilai rasio kegagalan sebagai berikut.

$$\hat{HR} = \frac{\hat{h}(t, X^*)}{\hat{h}(t, X)} = \exp\left[\hat{\beta}(IPK^* - IPK)\right] = \exp[-2.089(1-0)] = 0.124$$

Berarti mahasiswa dengan $IPK < 3.088$ memiliki laju lama studi 0.124 kalinya mahasiswa dengan $3.088 \leq IPK \leq 3.723$ atau bisa disebut juga mahasiswa dengan $3.088 \leq IPK \leq 3.723$ memiliki laju lama studi lebih cepat dibandingkan dengan mahasiswa dengan $IPK < 3.088$. Pada Tabel 3 menunjukkan hasil perhitungan rasio kegagalan untuk semua variable

Tabel 3. Rasio Kegagalan Untuk Semua Variabel

Nama Variabel	$\hat{\beta}_j$	$\exp\left(\hat{\beta}_j\right)$
X_6	-	-
$X_6(1)$	-2.089	0.124
$X_6(2)$	-1.155	0.315

4.4 Taksiran Peluang

Sebelum menghitung taksiran peluangnya, terlebih dahulu mengestimasi parameter, *estimasi hazard* dan *survival* dasar. Pada tabel 4 menunjukkan estimasi parameter β yang signifikan, tabel 5 menunjukkan *estimasi hazard* dan *survival* dasar.

Tabel 4. Estimasi Parameter Yang Signifikan

Nama Variabel	$\hat{\beta}_j$	Exp $\exp\left(\hat{\beta}_j\right)$
X_6	-	-
$X_6(1)$	-2.089	0.124
$X_6(2)$	-1.155	0.315

Tabel 5. Estimasi Hazard dan Survival Dasar

Lama Studi (Semester)	Hazard Dasar $h_0(t)$	Survival Dasar $S_0(t)$
7	0.306	0.903
8	3.941	0.270
9	7.454	0.084

Setelah diperoleh estimasi parameter serta estimasi *hazard* dan *survival* dasar, maka dapat dilakukan perhitungan peluang mahasiswa yang melakukan studi atau belum lulus $S(t, X)$ dan peluang mahasiswa yang lulus $F(t, X)$ pada berbagai waktu yang ditunjukkan pada Tabel 6 dan Tabel 7.

Sebelum melakukan perhitungan peluang mahasiswa yang melakukan studi $S(t, X)$ dan peluang mahasiswa yang lulus $F(t, X)$ perlu mengetahui model dari regresi *cox*, sebagai berikut:

$$h_i(t) = \exp(-2.089X_7(1) - 1.155X_7(2) + 0X_7)h_0(t)$$

dengan $X_7(1) = IPK < 3.088$; $X_7(2) = IPK < 3.088 \leq IPK \leq 3.723$; $X_7 = IPK > 3.723$. Misal pada data lama studi mahasiswa semester 7 dengan kategori $IPK < 3.088$ atau $X_7(1)$, maka dapat dihitung peluang mahasiswa yang masih melakukan studi $S(t, X)$ dan peluang mahasiswa yang lulus $F(t, X)$ sebagai berikut:

$$S(7|X_7(1)) = S_0(7)^{\exp(-2.089+1)} = 0.903^{0.124} = 0.9874$$

$$F(7|X_7(1)) = 1 - S(7|IPK(1)) = 1 - 0.9874 = 0.0126$$

Untuk perhitungan dugaan peluang mahasiswa yang masih melakukan studi $S(t, X)$ nilai *exp* yang digunakan adalah -2.089 karena menghitung dugaan peluang dengan kategori $IPK < 3.088$ atau $X_6(1)$. Artinya peluang mahasiswa yang memiliki $IPK < 3.088$ atau $IPK(1)$ masih belum lulus lebih dari 6 semester adalah 0.9874 dan

peluang mahasiswa yang memiliki $IPK < 3.088$ atau $IPK(1)$ yang telah lulus sampai dengan 7 semester adalah 0.0126.

Pada data lama studi mahasiswa semester 7 dengan kategori $3.088 \leq IPK \leq 3.723$ atau $X_7(2)$, maka dapat dihitung peluang mahasiswa yang masih melakukan studi $S(t, X)$ dan peluang mahasiswa yang lulus $F(t, X)$ sebagai berikut:

$$S(7|X_7(2)) = S_0(7)^{\exp(-1.155+1)} = 0.903^{0.315} = 0.9683$$

$$F(7|X_7(2)) = 1 - S(7|IPK(1)) = 1 - 0.9683 = 0.0317$$

Untuk perhitungan dugaan peluang mahasiswa yang masih melakukan studi $S(t, X)$ nilai *exp* yang digunakan adalah -1.155 karena kita akan menghitung dugaan peluang dengan kategori $3.088 \leq IPK \leq 3.723$ atau $X_7(2)$. Artinya peluang mahasiswa yang memiliki $3.088 \leq IPK \leq 3.723$ atau $X_7(2)$, masih belum lulus pada 8 semester adalah 0.9683 dan peluang mahasiswa yang memiliki $3.088 \leq IPK \leq 3.723$ atau $X_7(2)$, yang telah lulus sampai dengan 8 semester adalah 0.0317.

Tabel 6. Dugaan Peluang Mahasiswa Yang Melakukan Studi $S(t, X)$ Pada Berbagai Semester

IPK	Peluang Mahasiswa yang melakukan Studi (Belum Lulus) Lama Studi (Semester)		
	7	8	9
	$X_7(1)$	0.9874	0.850
$X_7(2)$	0.9683	0.662	0.4582
X_7	0.903	0.270	0.084

Pada Tabel 6 dapat dilihat bahwa jika semakin lama masa studi maka dugaan peluang mahasiswa yang masih melakukan studi $S(t, X)$ semakin kecil dan jika semakin tinggi IPK mahasiswa maka dugaan peluang mahasiswa yang masih studi $S(t, X)$ semakin kecil. Deskripsi dari Tabel 6 sebagai berikut dugaan peluang mahasiswa yang masih melakukan studi pada semester 7 dengan kategori $X_7(1)$ atau $IPK < 3.088$ adalah 0.9874, dugaan peluang mahasiswa yang masih melakukan studi pada semester 8 dengan kategori $X_7(1)$ atau $IPK < 3.088$ adalah 0.850, dugaan peluang mahasiswa yang masih melakukan studi pada semester 9 dengan kategori $X_7(1)$ atau $IPK < 3.088$ adalah 0.7355.

Dugaan peluang mahasiswa yang masih melakukan studi pada semester 7 dengan kategori $X_7(1)$ atau $IPK < 3.088$ adalah 0.9874, dugaan peluang mahasiswa yang masih melakukan studi pada semester 7 dengan kategori $X_7(2)$ atau $3.088 \leq IPK \leq 3.723$ adalah 0.9683 dan dugaan peluang mahasiswa yang masih melakukan studi pada semester 7 dengan kategori $X_7(1)$ atau $IPK > 3.723$ adalah 0.903.

Tabel 7. Dugaan Peluang Mahasiswa Yang Lulus $F(t, X)$ Pada Berbagai Semester

IPK	Peluang Mahasiswa Lulus Lama Studi (Semester)		
	7	8	9
	$X_7(1)$	0.0120	0.150
$X_7(2)$	0.0317	0.338	0.5418
X_7	0.0970	0.730	0.916

Pada Tabel 7 dapat dilihat bahwa jika semakin lama masa studi maka dugaan peluang mahasiswa yang lulus $F(t, X)$ semakin besar dan semakin tinggi IPK mahasiswa maka

dugaan peluang mahasiswa yang lulus $F(t, X)$ semakin besar. Deskripsi dari Tabel 7 sebagai berikut: dugaan peluang mahasiswa yang lulus pada semester 7 dengan kategori $X_7(1)$ atau $IPK < 3.088$ adalah 0.012, dugaan peluang mahasiswa yang lulus pada semester 8 dengan kategori $X_7(1)$ atau $IPK < 3.088$ adalah 0.150, dugaan peluang mahasiswa yang lulus pada semester 9 dengan kategori $X_7(1)$ atau $IPK < 3.088$ adalah 0.2645.

5 Kesimpulan

Berdasarkan hasil pengolahan data dan interpretasi hasil yang didapatkan, sehingga dapat ditarik kesimpulan berikut.

1. Faktor yang paling signifikan mempengaruhi lama studi mahasiswa Universitas Internasional Batam adalah faktor Indeks Prestasi Kumulatif (IPK).
2. Model data lama studi mahasiswa Universitas Internasional Batam Angkatan 2013 dengan metode regresi *cox proportional hazard* yaitu :

$$h_i(t) = \exp(-2.089X_6(1) - 1.155X_6(2) + 0X_6)h_0(t)$$

dengan $X_6(1) = IPK < 3.088$; $X_6(2) = IPK < 3.088 \leq IPK \leq 3.723$; $X_6 = IPK > 3.723$ dan $h_0(t)$ merupakan *hazard* dasar dengan nilai *hazard* dasar untuk tiap semester adalah $h_0(7) = 0.306$; $h_0(8) = 3.941$; $h_0(9) = 7.454$.

3. Dugaan peluang mahasiswa yang lulus tepat waktu yakni sampai dengan 8 semester dengan kategori IPK yaitu peluang lulus mahasiswa semester 8 dengan $IPK < 3.088$ adalah 0.150, peluang lulus mahasiswa semester 8 dengan $3.088 \leq IPK \leq 3,723$ adalah 0.338, dan peluang lulusan mahasiswa semester 8 dengan IPK adalah 0.73. Peluang mahasiswa yang lulus lebih dari 8 semester adalah sebagai berikut: peluang lulus mahasiswa semester 9 dengan $IPK < 3.088$ adalah 0.2645, peluang lulus mahasiswa semester 9 dengan $3.088 \leq IPK \leq 3.723$ adalah 0.5418, dan peluang lulus mahasiswa semester 9 dengan $IPK > 3.723$ adalah 0.916.

Daftar Pustaka

- [1] Afrizal, Moslem. 2014. *Penerapan Analisis Daya Tahan Ketidaklangsungan Pemakaian PIL KB*. Skripsi. Departemen Statistika FMIPA IPB. Bogor.
- [2] Iskandar, Bayu M. 2015. *Model Cox Proportional Hazard pada Kejadian Bersama*. Skripsi. FMIPA Universitas Negeri Yogyakarta. Yogyakarta.
- [3] Lawless, J.F. 1982. *Statistical Models and Methods for Lifetime Data*. John Wiley & Sons, Inc. USA.
- [4] Paura, Liga dan Arhipova, Iriana. 2014. Cause Analysis of Student Dropout Rate in Higher Education Study Program. *ELSEVIER Journal*. 109:1282-1286.
- [5] Rinni, dkk. 2014. Pemodelan Laju Kesembuhan Rawat Inap Thypus Abdominalis (Demam Tifoid) Menggunakan Model Regresi Hidup Pasien Kanker Paru di RSUP DR. Kariadi Semarang dengan Model Kegagalan Proporsional Dari Cox (Studi Kasus di RSUD Kota Semarang). *Jurnal Gaussian*, 3(1): 31-40.
- [6] Utami, Dyah Tri. 2015. *Analisis Data Uji Hidup Pasien Kanker Paru di RSUP DR. Kariadi Semarang dengan Model Regresi*. Skripsi. Jurusan Matematika, FMIPA Universitas Negeri Semarang. Semarang.