

Analisis Regresi COX Proportional Hazard pada Pemodelan Waktu Tunggu Mendapatkan Pekerjaan

Hendra Dukalang^{1*}

¹ Jurusan Ekonomi Syariah, Fakultas Ekonomi dan Bisnis Islam, IAIN Sultan Amai Gorontalo, Jl. Sultan Amai, Kec. Ponelo, Kabupaten Gorontalo 96210, Indonesia

* Penulis Koresponden. Email: hendra.statistika@iaingorontalo.ac.id

ABSTRAK

Analisis Survival merupakan suatu metode statistika yang memperhatikan waktu hingga terjadinya suatu kejadian (*event*). Tujuan dari analisis survival adalah menaksir probabilitas kelangsungan hidup, kematian, dan peristiwa lainnya sampai periode tertentu. Salah satu Analisis Survival yang paling populer digunakan adalah regresi Cox proportional hazard. Regresi Cox proportional hazard merupakan salah satu regresi semiparametrik yang bertujuan mengetahui kombinasi faktor-faktor yang berpengaruh terhadap variable respon yang berupa waktu survival. Dalam penelitian ini waktu survival yang digunakan adalah selang waktu tunggu sarjana dalam mendapatkan pekerjaan pertamanya. Data yang digunakan adalah data mengenai faktor-faktor yang mempengaruhi sarjana matematika tahun 2016 dan 2017 dalam memperoleh pekerjaan pertamanya. Berdasarkan hasil analisis dan pengujian, diperoleh variable yang berpengaruh adalah asal sekolah SMA dan SMK, informasi lowongan pekerjaan, kesesuaian pekerjaan dengan bidang ilmu dan pengalaman kerja dari sarjana tersebut.

Kata Kunci:

Analisis Survival; Asumsi Proportional Hazard; Regresi Cox Proportional Hazard

Diterima:
13-12-2018

Disetujui:
28-1-2019

Online:
28-1-2019

ABSTRACT

Survival analysis is a statistical method that considers the time until the occurrence of an event (*event*). The purpose of survival analysis is estimating the probability of survival, death, and other events up to a certain period. One of the most popular Survival Analysis used is Cox proportional hazard regression. Cox proportional hazard regression is one of the semiparametric regression which aims to determine the combination of factors that influence the response variable in the form of survival time. In this study the survival time used was a time delay for scholars to get their first job. The data used are data on the factors that influence mathematics undergraduates in 2016 and 2017 in obtaining their first job. Based on the results of analysis and testing, it was found that the influential variables were the origin of high school and vocational school, job vacancy information, suitability of work with the field of science and work experience of the bachelor.

Keywords:

Survival Analysis; Proportional Hazard Assumptions; Cox Proportional Hazard Regression

Received:
2018-12-13

Accepted:
2019-1-28

Online:
2019-1-28

1. Pendahuluan

Analisis survival adalah suatu metode yang berhubungan dengan waktu, mulai dari time origin atau start point sampai dengan terjadinya suatu kejadian khusus atau end point. Dalam analisis survival terdapat beberapa metode yang dapat digunakan untuk menganalisis data survival yakni metode parametrik, nonparametrik, dan semiparametrik. Metode semiparametrik adalah gabungan dari model parametrik dan nonparametrik. Dalam menyelesaikan metode semiparametrik dapat menggunakan metode Regresi Cox karena fungsi baseline hazard mengikuti model-model nonparametrik sedangkan variabel bebasnya mengikuti model parametrik.

Regresi Cox dikenal juga dengan istilah Regresi Cox Proportional Hazard karena asumsi proportional hazard. Apabila tidak memenuhi asumsi propotional hazard maka ada 3 pilihan yang dapat digunakan untuk mengatasinya yaitu mengeluarkan variabel yang tidak memenuhi asumsi dari model, menggunakan model cox stratified dan dengan model cox extended.

Dalam penelitian ini model tersebut diterapkan pada data selang waktu tunggu sarjana dalam mendapatkan pekerjaan pertama. Proses menunggu yang dilakukan oleh sarjana dari setelah lulus sampai akhirnya mendapatkan pekerjaan pertama termasuk dalam waktu survival. Waktu menunggu masing-masing sarjana untuk mendapatkan pekerjaan pertama-pun berbedabeda bahkan ada yang belum mendapatkan pekerjaan hingga akhir waktu penelitian (tersensor). Berdasarkan hasil survey peneliti dari 91 lulusan sarjana Pendidikan Matematika Universitas Negeri Gorontalo tahun 2016 dan 2017, terdapat 85 sarjana yang sudah mendapatkan pekerjaan pertama setelah lulus, sedangkan 6 sarjana belum mendapatkan pekerjaan (tersensor). Hal ini tentunya dipengaruhi oleh berbagai faktor. Dalam penelitian [1] beberapa faktor yang mempengaruhi waktu tunggu sarjana dalam mendapatkan pekerjaan pertama adalah pengalaman organisasi. Selain itu dalam penelitian [2], faktor yang berpengaruh signifikan terhadap selang waktu tunggu sarjana adalah IPK . Sedangkan dalam penelitian [3] yang berpengaruh adalah jumlah anggota keluarga, status SLTA, jenis pekerjaan yang dipilih dan motivasi dalam mencari kerja. Serta dalam penelitian [4] faktor yang mempengaruhi lama mencari kerja bagi tenaga terdidik adalah umur, tingkat pendidikan, pendapatan dan pengalaman kerja. Sehingga model Regresi Cox sangat tepat digunakan jika terdapat data yang tersensor yang akan digunakan untuk menganalisa hubungan dan pengaruh berbagai faktor terhadap variabel responnya yang terkait dengan waktu bertahan (survival time) individu terhadap suatu kejadian.

Berdasarkan uraian tersebut maka penulis memilih judul "Pemodelan Selang Waktu Tunggu Sarjana dalam Mendapatkan Pekerjaan Pertama Menggunakan Regresi Cox Proportional Hazard" untuk studi kasus Sarjana Pendidikan Matematika Universitas Negeri Gorontalo tahun 2016 & 2017.

Penelitian ini bertujuan memodelkan selang waktu tunggu sarjana pendidikan matematika dalam mendapatkan pekerjaan pertama dengan menggunakan regresi Cox Proportional Hazard serta menentukan faktor-faktor yang mempengaruhi waktu tunggu tersebut.

2. Metode

Data dalam penelitian ini menggunakan data primer dan data sekunder. Data primer diperoleh dengan menyebarkan kuisioner kepada sarjana pendidikan UNG tahun 2016

dan 2017. Sedangkan data sekunder adalah IPK yang diperoleh dari operator jurusan Matematika UNG. Sampel yang digunakan dalam penelitian adalah 91 sarjana pendidikan matematika UNG tahun 2016 dan 2017 dengan berbagai faktor yang diteliti.

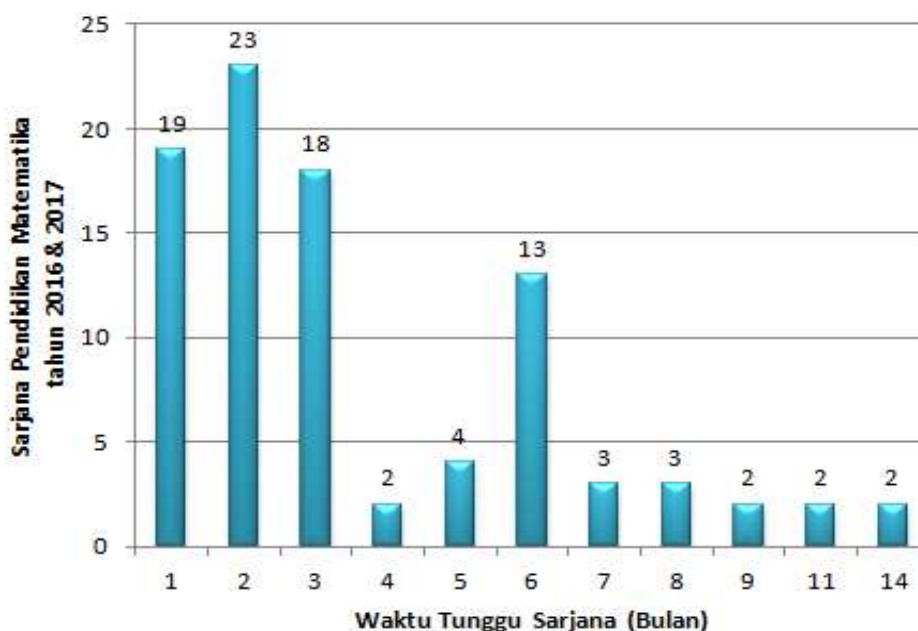
Variabel bebas dalam penelitian ini meliputi asal SLTA (X_1), IPK (X_2), informasi lowongan pekerjaan (X_3), kesesuaian pekerjaan dengan bidang ilmu (X_4), dan pengalaman kerja (X_5). Sedangkan variabel terikat dalam penelitian ini adalah selang waktu sarjana dalam mendapatkan pekerjaan pertama.

Dalam menganalisis data menggunakan Software SAS. Adapun langkah-langkah analisis dalam penelitian ini adalah sebagai berikut:

- a. Analisis deskriptif hasil survey
- b. Pengujian asumsi proportional dengan menggunakan uji GOF
- c. Membentuk model regresi cox proportional hazard
- d. Interpretasi nilai hazard ratio

3. Hasil dan Pembahasan

Berdasarkan hasil survey, sarjana pendidikan matematika yang sudah mendapatkan pekerjaan (tidak tersensor) sebanyak 85 sarjana (93%) sedangkan yang belum mendapatkan pekerjaan sampai dengan akhir waktu penelitian yakni Maret 2018 (tersensor) sebanyak 6 sarjana (7%). Untuk waktu tunggu sarjana pendidikan matematika dalam mendapatkan pekerjaan pertama berkisar antara 1-14 bulan dengan waktu tunggu paling banyak dalam mendapatkan pekerjaan adalah 2 bulan sesuai dengan gambar berikut ini:



Gambar 1. Sebaran waktu tunggu sarjana pendidikan matematika

Selanjutnya statistik deskriptif rata-rata waktu tunggu sarjana pendidikan matematika berdasarkan variabel.

Tabel 1. Statistik deskriptif rata-rata waktu tunggu sarjana berdasarkan variabel

Variabel	Kategori	N	Rata-rata waktu tunggu sarjana
Asal SLTA	SMA	60	3.3500
	SMK	19	2.3616
	MA	6	4.500
IPK	< 3.00	6	5.3333
	3.00-3.50	66	3.0909
	>3.50	13	3.2308
Info low. kerja	Iklan	2	5.500
	Internet	5	2.600
	Koneksi	70	3.3268
	Info low. Kema	1	6
Kesesuaian Pekerjaan	Lainnya	7	2.1429
	Sesuai	78	3.0897
	Tidak sesuai	7	5.2587
Pengalaman Kerja	Pernah bekerja	55	2.9818
	Tidak pernah berkerja	30	3.800

3.1. Pengujian Asumsi Proportional Hazard

Berikut ini merupakan hasil pengujian asumsi PH dengan menggunakan uji GOF:

Tabel 1. Hasil Pengujian Asumsi menggunakan uji GOF

Variabel	Korelasi	P-value	Keputusan
Asal SLTA	-0.11309	0.3028	Gagal tolak H_0
IPK	0.15119	0.1672	Gagal tolak H_0
Info low. pekerjaan	0.05330	0.6280	Gagal tolak H_0
Kesesuaian pekerjaan	-0.05870	0.5936	Gagal tolak H_0
Pengalaman kerja	-0.08212	0.4550	Gagal tolak H_0

Berdasarkan Tabel 1, dapat dilihat bahwa nilai p-value semua variabel lebih kecil alpha (0.05) maka dapat menghasilkan keputusan gagal tolak H_0 sehingga tidak terdapat korelasi yang besar antara variabel asal SLTA, IPK, informasi lowongan pekerjaan, kesesuaian pekerjaan, pengalaman kerja dengan waktu survival. Dengan demikian dapat disimpulkan bahwa semua variable memenuhi asumsi proportional hazard, sehingga dapat dilanjutkan dengan memodelkan menggunakan regresi cox proportional hazard.

3.2. Pemodelan Regresi Cox Proportional Hazard

Berikut ini merupakan hasil uji signifikansi parameter secara keseluruhan model regresi cox proportional hazard

Tabel 2. Hasil uji signifikansi parameter secara keseluruhan model regresi cox proportional hazard

Uji	Chi Square	Df	P-value	Keputusan
Likelihood Ratio	21.8075	10	0.0161	Tolak H_0

Berdasarkan Tabel 2, dapat dilihat bahwa nilai Likelihood ratio untuk model regresi cox propoertional hazard adalah sebesar 21.807 dengan df 10 diperoleh nilai p-value sebesar 0.0161 lebih kecil dari alpha 0.05, maka dapat menghasilkan keputusan tolak H_0 yang artinya minimal terdapat 1 variabel yang berpengaruh signifikan terhadap selang waktu tunggu sarjana dalam mendapatkan pekerjaan pertama. Sehingga model regresi cox proportional hazard ini dapat digunakan dalam memodelkan selang waktu tunggu sarjana dalam mendapatkan pekerjaan pertama.

Selanjutnya pengujian signifikansi parameter secara parsial untuk mengetahui pengaruh masing-masing variabel terhadap selang waktu tunggu sarjana dalam mendapatkan pekerjaan pertama. Berikut ini merupakan hasil pengujian signifikansi parameter secara parsial.

Tabel 3. Hasil pengujian signifikansi parameter secara parsial model regresi cox proportional hazard

Variabel	Estimasi Parameter	Chi Square	P-Value	Keputusan
Asal SLTA (1)	0.87580	3.8070	0.0510	Tolak H_0
Asal SLTA (2)	0.98894	4.0893	0.0432	Tolak H_0
IPK (1)	0.07524	0.0202	0.8870	Gagal tolak H_0
IPK (2)	0.50392	2.2585	0.1329	Gagal tolak H_0
Info low. pekerjaan (1)	-0.03850	0.0015	0.9690	Gagal tolak H_0
Info low. pekerjaan (2)	-0.57598	0.7146	0.3979	Gagal tolak H_0
Info low. pekerjaan (3)	-0.78750	3.5335	0.0601	Tolak H_0
Info low. pekerjaan (4)	-1.28332	1.3464	0.2459	Gagal tolak H_0
Kesesuaian pekerjaan	1.36295	5.8524	0.0156	Tolak H_0
Pengalaman kerja	0.46806	3.7739	0.0521	Tolak H_0

Berdasarkan Tabel 3, secara parsial nilai p-value variabel asal SLTA (1), asal SLTA (2), informasi lowongan pekerjaan (3), kesesuaian pekerjaan dan pengalaman kerja peberturut-turut adalah 0.0510, 0.0432, 0.0601, 0.0156 dan 0.0529 lebih kecil dari alpha 0.1 maka dapat menghasilkan keputusan tolak H_0 yang artinya secara parsial asal sekolah SMA dan SMK, informasi lowongan pekerjaan melalui koneksi, kesesuaian pekerjaan dan pengalaman kerja berpengaruh terhadap lama waktu tunggu sarjana dalam mendapatkan pekerjaan pertama.

Selanjutnya berdasarkan Tabel 3, dapat dibentuk model regresi cox proportional hazard yakni sebagai berikut:

$$h(t) = h_0(t) \exp[0.87580 X_1(1) + 0.98894 X_1(2) + 0.07524 X_2(1) + 0.50392 X_2(2) - 0.03850 X_3(1) - 0.57598 X_3(2) - 0.78750 X_3(3) - 1.28332 X_3(4) + 1.36295 X_4 + 0.46806 X_5]$$

3.3. Interpretasi Nilai Hazard Ratio

Berdasarkan Tabel 3, dapat diperoleh nilai hazard ratio untuk variabel yang berpengaruh signifikan terhadap selang waktu tunggu sarjana dalam mendapatkan pekerjaan pertama yakni sebagai berikut:

1. Nilai hazard ratio asal SLTA (1) sebesar $e^{0.87580} = 2.401$ yang artinya sarjana pendidikan yang berasal dari SMA memiliki kemungkinan 2.401 kali untuk mendapatkan pekerjaan pertama dibandingkan yang sarjana berasal dari MA.
2. Nilai hazard ratio asal SLTA (2) sebesar $e^{0.98894} = 2.688$ yang artinya sarjana pendidikan yang berasal dari SMK memiliki kemungkinan 2.688 kali untuk mendapatkan pekerjaan pertama dibandingkan dengan sarjana yang berasal dari MA.
3. Nilai hazard ratio informasi lowongan pekerjaan (3) sebesar $e^{-0.78750} = 0.455$ yang artinya sarjana yang mendapatkan informasi lowongan pekerjaan melalui koneksi memiliki kemungkinan 0.455 kali untuk mendapatkan pekerjaan pertama dibandingkan dengan sarjana yang mendapatkan informasi melalui lainnya (mendatangi langsung instansi).
4. Nilai hazard ratio kesesuaian pekerjaan sebesar $e^{1.36295} = 3.908$ yang artinya sarjana yang mendapatkan pekerjaan yang sesuai dengan bidang ilmunya memiliki kemungkinan 3.908 kali untuk mendapatkan pekerjaan pertama dibandingkan dengan sarjana yang mendapatkan pekerjaan yang tidak sesuai dengan bidang ilmunya.

Nilai hazard ratio pengalaman kerja sebesar $e^{0.46806} = 1.597$ yang artinya sarjana yang pernah bekerja sebelum lulus memiliki kemungkinan 1.597 kali untuk mendapatkan pekerjaan pertama dibandingkan dengan sarjana yang tidak pernah bekerja sebelum lulus.

4. Kesimpulan

Berdasarkan pengujian kesesuaian model secara keseluruhan pada model regresi cox proportional hazard diperoleh minimal terdapat 1 variabel yang berpengaruh signifikan terhadap model sehingga model ini dapat digunakan dalam pemodelan selang waktu tunggu sarjana dalam mendapatkan pekerjaan pertama. Model yang terbentuk yakni:

$$h(t) = h_0(t) \exp [0.87580 X_1(1) + 0.98894 X_1(2) + 0.07524 X_2(1) + 0.50392 X_2(2) - 0.03850 X_3(1) - 0.57498 X_3(2) - 0.78750 X_3(3) - 1.28332 X_3(4) + 1.36295 X_4 + 0.46806 X_5]$$

Berdasarkan pengujian kesesuaian model secara parsial pada model regresi cox proportional hazard diperoleh variabel yang berpengaruh signifikan terhadap selang waktu tunggu sarjana dalam mendapatkan pekerjaan adalah asal SLTA SMA dan SMK, informasi lowongan pekerjaan melalui koneksi, kesesuaian pekerjaan dengan bidang ilmu dan pengalaman kerja.

Saran

Karena keterbatasan peneliti, data tersensor pada penelitian ini hanyalah sarjana pendidikan yang belum mendapatkan pekerjaan sehingga untuk penelitian selanjutnya yang sudah mendapatkan pekerjaan sebelum lulus dapat dijadikan sebagai data tersensor. Dalam pengembangan penelitian selanjutnya sebaiknya menggunakan jumlah sampel yang lebih besar, diperjelas untuk pengalaman kerja dan juga jenis pekerjaan yang diperoleh sarjana pendidikan matematika diklasifikasikan serta menggunakan variabel lain sehingga dapat menghasilkan model terbaik yang lebih akurat.

Ucapan Terima Kasih

Terima Kasih diucapkan kepada Istri Tercinta yang telah membantu dalam melakukan penyusunan laporan dan analisis data, Terima kasih juga kepada putri tercinta yang telah menjadi penyemangat dalam melakukan setiap aktifitas sehingga tulisan ini dapat selesai tepat waktu.

Referensi

- [1] Faruk, A., Amran, A., & Nasir, N., 2014, Aplikasi Model Proportional Hazard Cox Pada Waktu Tunggu Kerja Lulusan Jurusan Matematika FMIPA Universitas Srwijaya, Jurnal Penelitian Sains, Vol 17. No. 1.
- [2] Ninik, H., 2016, Analisis Uji Ketahanan Hidup Data Waktu Tunggu Sarjana dengan Metode Kaplan Meier Berbantuan Software Medcalc, Libarary Universitas Negeri Semarang, 1 Desember 2017, <http://lib.unnes.ac.id/25013/1/4112313006.pdf>
- [3] Putranto, R.T. & Mashuri, M., 2012, Analisis Statistik Tentang Faktor-faktor Yang Mempengaruhi Waktu Tunggu Kerja Fresh Graduate di Jurusan Statistika Institut Teknologi Sepuluh November (ITS) dengan Metode Regresi Logistik Ordinal, Jurnal Sains dan Seni ITS, Vol.1 No.1
- [4] Setiawan, S.A., 2010, Pengaruh Umur, Pendidikan, Pendapatan, Pengalaman Kerja, dan Jenis Kelamin terhadap Lama Mencari Kerja Bagi Tenaga Kerja Terdidik di Kota Magelang, Skripsi Universitas Diponegoro. Semarang, 1 Desember 2017, <http://eprints.undip.ac.id/29486/1/JURNALSKRIPSI.pdf>
- [5] Kleinbaum, D.G. & Klein, M., 2012, Survival Analysis: A Self-Learning Text Third Edition, New York, Springer-Verlag
- [6] Hosmer, D.W. & Lemeshow, 2000, Applied Logistic Regression Second Edition, United State Of America, John Willey & Sons Inc
- [7] Allison, P.D., 2010, Survival Analysis Using SAS: A Partical Guide Second Edition, Cary NC, SAS Institute Inc
- [8] Collet, D., 2003, Modelling Survival Data in Medical Research Second Edition, London, Chapman and Hall