

ANALISIS *SURVIVAL* DALAM MEMODELKAN SISWA PUTUS SEKOLAH

Rahmat Hidayat¹, Yuli Hastuti²

Program Studi Matematika, Fakultas Sains

Universitas Cokroaminoto Palopo

Email: dayatmath@gmail.com¹

Abstract. One of hurdles faced in increasing quantity and quality of education in Indonesia is the number of children who cannot access education in the school has increased overtime. The purpose of this study was to make a data model related the number of children who cannot access education by using survival analysis. This study was conducted in the Palopo city, South Sulawesi. The result of this study showed that the low of parent's education level, the number of family member were factors which has caused plenty of children in Palopo cannot access education properly.

Abstrak. Salah satu kendala dalam meningkatkan kuantitas dan kualitas pendidikan di Indonesia adalah masih tingginya angka anak putus sekolah usia wajib belajar. Metode yang dapat digunakan untuk memodelkan data tersebut adalah Analisis *Survival*. Penelitian ini memodelkan data usia anak putus sekolah di Kota Palopo dengan menggunakan *Bayesian Analisis Survival*. Hasil penelitian menunjukkan bahwa, tingkat pendidikan orangtua yang rendah dan jumlah anggota rumah mengakibatkan tingginya resiko putus sekolah siswa di Kota Palopo.

Kata Kunci: Analisis *Survival*, Angka Putus Sekolah, dan Pemodelan.

Salah satu upaya pemerintah dalam meningkatkan kuantitas dan kualitas pendidikan adalah menyelenggarakan Program Wajib Belajar (Wajar) 6 tahun mulai tahun 1984, Program Wajib Belajar 9 tahun mulai tahun 1994 dan rencananya Program Wajib Belajar 12 tahun dimulai pada tahun 2013. Agar program pendidikan dasar tersebut dapat mencapai sasaran, maka dilaksanakan di tiap-tiap daerah. Arah dan tujuan utama perluasan pendidikan untuk meningkatkan angka partisipasi kasar (APK) dan menekan angka putus sekolah (APtS). Namun APK dan APtS hanya memberikan gambaran secara umum tentang besarnya peluang peserta didik yang sedang atau telah menerima pendidikan pada jenjang tertentu, sehingga kita akan mengalami kesulitan untuk mengetahui seberapa besar peserta didik dapat melanjutkan pendidikan atau seberapa besar peserta didik akan putus sekolah. Jangka waktu peserta didik kelas I sampai dengan kelas VI salah satu ukuran terlaksananya program wajib belajar merupakan contoh dari data *survival*.

Berbagai penelitian seperti Bradby *et al* (2000) di Texas, Schwartz, W (1999) di San Francisco, Calderon, D (1996) di kota Beirut,

Tripoli, Bequ'a, Saida dan Tyre, A.A. Ketut Oka (2000) di Bali serta Sugeng Arianto (2001) di Jambi, Syarifuddin A (2012) dan Yensy NA menyatakan beberapa faktor yang mempengaruhi anak putus sekolah, yaitu: status ekonomi, jenis pendidikan siswa (umum atau kejuruan), kehamilan, kemiskinan, ketidaknyamanan, kenakalan siswa, penyakit, minat, tradisi/adat istiadat, pendidikan orangtua, pekerjaan orangtua, usia orangtua, jumlah tanggungan keluarga, kondisi tempat tinggal serta perhatian orang tua. Oleh karena itu, pemodelan data anak putus sekolah wajib belajar Sembilan Tahun di Kota Palopo dalam penelitian ini menggunakan metode *Survival*. Model *Survival* untuk data tersebut dapat digunakan untuk mengidentifikasi faktor-faktor yang mempengaruhi anak putus sekolah.

ANALISIS *SURVIVAL*

Data *survival* dicatat sebagai data tentang jangka waktu terjadinya suatu kejadian mulai awal sampai akhir (Lee ET dan Wang JW, 2003). Hal yang menarik dalam analisis *survival* adalah adanya titik kejadian (*event point*) dalam kelompok atau kelompok-kelompok individu yang disebut gagal

(*failure*), dan waktu bertahannya disebut waktu hidup (*life time*). Ada tiga syarat yang harus dipenuhi dalam menentukan waktu gagal secara teliti, yaitu ada waktu permulaan, satuan pengukuran yang jelas, dan definisi gagal harus jelas.

Analisis *survival* merupakan suatu metode statistik yang berkaitan dengan waktu, sekumpulan prosedur statistik untuk keperluan analisis data dimana data yang digunakan berupa data waktu sampai terjadinya *event* tertentu (Clark et al, 2003). Oleh sebab itu, dalam analisis *survival* dibutuhkan hal berikut :

1. waktu asal yang terdefinisi dengan baik (yaitu ketika suatu objek masuk dalam studi/pengamatan)
2. skala waktu pengukuran jelas
3. waktu yang juga terdefinisi dengan baik.

Fungsi Survival

Misalkan T adalah variabel random yang menggambarkan *survival time* dan t menyatakan beberapa nilai tertentu yang diperhatikan untuk variabel T . Maka, fungsi *survival* dapat didefinisikan sebagai probabilitas suatu objek bertahan melebihi suatu waktu tertentu t (Collet, 2003), atau dengan kata lain, probabilitas bahwa variabel random T lebih besar dari waktu yang ditentukan, yaitu $t(t>0)$ secara matematis dapat dinyatakan sebagai :

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

Jika T adalah variabel random kontinu, maka fungsi *survival* merupakan komplemen dari fungsi distribusi kumulatif, dimana fungsi distributif kumulatif adalah probabilitas bahwa variabel random T kurang dari atau sama dengan waktu t atau secara matematis dinyatakan dengan:

$$F(t) = P(T \leq t)$$

Fungsi *survival* yang merupakan komplemen dari fungsi distribusi kumulatif yaitu:

$$\begin{aligned} S(t) &= P(T > t) \\ &= 1 - P(T \leq t) \\ &= 1 - F(t) \end{aligned}$$

Fungsi *survival* digunakan untuk menyatakan probabilitas suatu individu bertahan dari waktu mula-mula sampai waktu t . Waktu *survival* dilambangkan dengan T yang merupakan variabel random dan mempunyai fungsi distribusi peluang $f(t)$. Fungsi *survival* $S(t)$, didefinisikan sebagai probabilitas bahwa waktu *survival* lebih besar atau sama dengan t sehingga:

$$S(t : \mu, \sigma) = 1 - \Phi\left(\frac{\ln(t) - \mu}{\sigma}\right)$$

Fungsi Kegagalan (fungsi hazard)

Fungsi hazard yaitu fungsi yang menyatakan peluang seseorang mengalami risiko atau kejadian seperti kegagalan atau meninggal pada waktu t dengan syarat bahwa seseorang itu telah bertahan hingga waktu t (Hidayat R et al, 2014), fungsinya diberikan Cox pada tahun 1972:

$$h(t) = \lim_{\delta t \rightarrow 0} \frac{P(t \leq T \leq t + \delta t | T \geq t)}{\delta t}$$

Model ini melibatkan komponen yang disebut fungsi *baseline hazard* yang melibatkan waktu t tetapi tidak melibatkan variabel prediktor. Selain itu, model ini juga melibatkan komponen lain yaitu pernyataan eksponensial terhadap jumlah perkalian efek dari masing-masing prediktor terhadap prediktor tersebut komponen ini menjamin taksiran hazard yang dihasilkan selalu *non-negatif*. Hal ini sesuai yang diharapkan karena variabel denpennya adalah waktu sampai terjadi kejadian sehingga tidak mungkin negatif.

METODE PENELITIAN

Sampel penelitian ini adalah siswa SD dan SMP Kota Palopo yang putus sekolah yang tercatat di kelurahan Kota Palopo pada bulan penelitian yaitu Mei 2016.

Tabel 1. Variabel penelitian

Variabel	Keterangan
Jangka waktu sekolah	Skala numerik
Jenis Kelamin	1: Laki-laki

	0: Perempuan
Pendidikan kepala keluarga	3: Tamat PT 2: Tamat SMA 1: Tamat SLTP 0: Tidak Tamat SLTP
Jumlah anggota rumah tangga	Skala numerik

Langkah-langkah pemodelan data anak putus sekolah di Kota Palopo menggunakan Analisis *Survival* sebagai berikut:

1. Mengumpulkan data
2. Membuat *row* data berdasarkan variabel
3. Mengidentifikasi data tersensor atau data tidak tersensor
4. Menentukan model *Survival* yang terbaik
5. Menentukan distribusi prior untuk setiap parameter dalam model *Survival*
6. Mengestimasi *Fungsi Hazard Proportional*
7. Melakukan uji parsial terhadap estimator parameter yang diperoleh dari langkah (6)
8. Menginterpretasi model untuk mengetahui faktor-faktor yang mempengaruhi anak putus sekolah di Kota Palopo.

HASIL DAN PEMBAHASAN

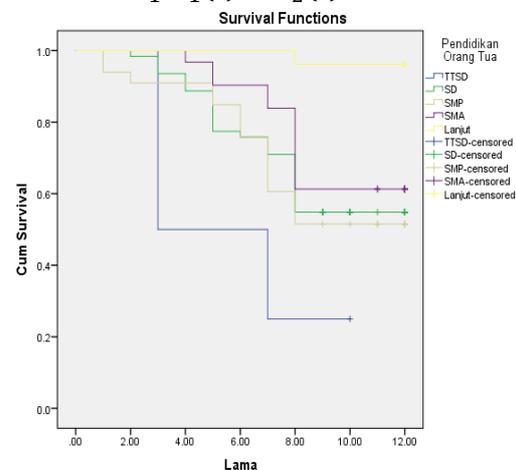
Membandingkan dua kelompok data survival dengan metode *Kaplan-Meier*

Hasil uji logrank berdasarkan variabel jenis kelamin memuat informasi bahwa dengan taraf nyata $\alpha=0,05$ diperoleh $p\text{-value}=0,001 < \alpha=0,05$ signifikan untuk menolak H_0 . Jadi dapat disimpulkan bahwa terdapat perbedaan yang nyata antara tingkat survival putus sekolah antara siswa yang berjenis kelamin laki-laki dan siswa. Pada Fungsi survival terlihat bahwa siswa perempuan kelanjutan studinya lebih baik dibandingkan siswa laki-laki.

Pada gambar 1 menunjukkan bahwa terdapat perbedaan keberlanjutan studi siswa yang orang tuanya tidak tamat SD, hanya

tamat SMP, hanya tamat SMA dan lulus D3 ataupun S1. $H_0: S_1(t) = S_2(t)$

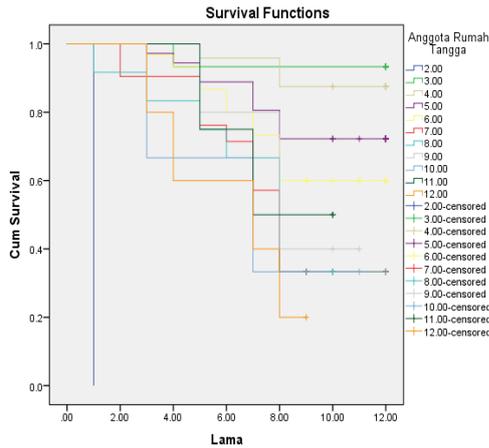
$$H_1: S_1(t) \neq S_2(t)$$



Gambar 1. *Kaplan Meier* pendidikan kepala keluarga

Hasil uji *Log Rank* pada taraf nyata $\alpha=0,05$ diperoleh $p\text{-value}=0,000 < \alpha=0,05$ signifikan untuk menolak H_0 . Jadi dapat disimpulkan bahwa terdapat perbedaan yang nyata antara tingkat pendidikan orang tua siswa. Pada gambar terlihat bahwa siswa yang orang tuanya lulus D3 atau S1 cenderung memiliki keberlanjutan studi yang tinggi dibandingkan dengan yang tidak tamat SD, hanya tamat SMP, dan hanya tamat SMA. Setelah itu siswa yang orang tuanya hanya lulus SMA, siswa yang orang tuanya hanya lulus SD, siswa yang orang tuanya hanya lulus SMP dan terakhir adalah siswa yang orang tuanya tidak lulus SD.

Hasil analisis *Kaplan Meier* pada Gambar 2. menunjukkan bahwa terdapat perbedaan keberlanjutan studi siswa yang dipengaruhi oleh jumlah ART.



Gambar 2. Kaplan Meier jumlah ART

Hasil uji logrank berdasarkan jumlah ART pada Tabel 4. Dieproleh p -value sebesar $0,000 < \alpha=0,05$ signifikan untuk menolak H_0 . Jadi dapat disimpulkan bahwa terdapat perbedaan yang nyata antara jumlah ART. Pada Gambar 3 terlihat bahwa siswa yang memiliki jumlah ART 3 memiliki keberlanjutan studi yang paling tinggi, kemudian berturut-turut siswa yang memiliki jumlah ART empat orang, jumlah ART lima orang, jumlah ART enam orang, jumlah ART 11 orang, jumlah ART sembilan orang. Sedangkan untuk siswa dengan jumlah ART 7, 8, dan 10 orang memiliki keberlanjutan studi yang cenderung sama, dan kelompok yang paling rentan untuk putus sekolah adalah siswa dengan jumlah ART 12 orang.

Membandingkan dua kelompok data survival dengan metode Cox Proporsional Hazard

Jika data survival yang akan dibandingkan lebih dari dua kelompok individu, misalnya mengetahui perbedaan karakteristik jenis kelamin, tingkat pendidikan orang tua, jumlah ART dan lain-lain maka metode Kaplan-Meier menjadi tidak praktis. Hal ini dikarenakan dalam metode Kaplan-Meier setiap dua kelompok populasi harus diuji secara tersendiri, sehingga jika ada beberapa kelompok maka harus dilakukan uji berulang-ulang. Jika responden mempunyai beberapa karakteristik, metode Cox hazard

proporsional dapat menerangkan pengaruh karakteristik-karakteristik tersebut terhadap peubah respon secara simultan.

Tabel berikut memuat hasil pendugaan parameter dengan menggunakan metode Cox proporsional hazard.

Tabel 2. Estimasi parameter survival

Variabel	Penduga parameter	Hazard ratio
Jenis Kelamin	0,081	1,085
Lulus SD	2,869	17,616
Lulus SMP	2,344	10,425
Lulus SMA	2,569	13,047
Lulus D3/S1	2,289	9,866
Jumlah ART	0,180	1,198

Nilai hazard ratio memungkinkan kita untuk membandingkan antara beberapa grup dalam analisis survival. Dari nilai hazard ratio terlihat bahwa risiko putus sekolah siswa laki-laki lebih tinggi yaitu sebesar 1,085 kali dibandingkan siswa perempuan, serta risiko putus sekolah siswa dengan riwayat pendidikan orang tua yang hanya lulus SD, lulus SMA, lulus SMP dan lulus D3 atau S1 masing-masing 17.616, 13.047, 10.425, dan 9.866. Setiap penambahan satu jumlah keluarga maka akan meningkatkan risiko putus sekolah sebesar 19.8%.

KESIMPULAN

Berdasarkan model survival data putus sekolah Kota Palopo diperoleh informasi bahwa siswa laki-laki memiliki risiko putus sekolah yang lebih tinggi dibandingkan siswa perempuan, serta siswa yang orang tuanya memiliki riwayat pendidikan tinggi cenderung memiliki keberlanjutan studi yang tinggi, artinya semakin rendah riwayat pendidikan orang tua siswa, maka risiko putus sekolah siswa tersebut juga tinggi, dan semakin banyak jumlah anggota keluarga maka risiko putus sekolah siswa tersebut juga semakin tinggi.

DAFTAR PUSTAKA

- Bradburn M, Clark TG, Love C. 2003. Multivariate data analysis, an introduction to concepts and methods. *British Journal of Cancer*. 89(3): 431-436.
- Cox DR. 1972. Regression models and life tables (with discussion). *J R Statisc Sic B* 34(2): 187-220.
- Hoon TS. 2008. Using Kaplan-Meier and Cox Regresion in Survival Analysis. *Journal ESTEEM* 4(2): 3-14.
- Iriawan, N. 2012. *Pemodelan dan Analisis Data-Driven Volume 1*, ITS Press, Surabaya.
- Kaplan EL, Meier P. 1958. *Non Proportional Estimation from Incomplet Observation*. Journal of The Americal Statistical Association (53): 457-481.
- Kleinbaum DG, Klein M. 2012. *Survival Analysis A Self-Learning Text. 3thed*. New York: Springer.
- Love C, Altman DG, Bradburn M. 2003. Multivariate data analysis. *British Journal of Cancer* 89(3): 437-443.