

# BAB I PENDAHULUAN

## A. Latar Belakang Masalah

Analisis survival merupakan prosedur statistika yang digunakan untuk menganalisis data dimana variabel yang diperhatikan adalah jangka waktu dari awal pengamatan sampai suatu *event* terjadi dengan melihat variabel-variabel yang mempengaruhi *event* tersebut. Jangka waktu dari awal pengamatan sampai terjadinya suatu *event* disebut dengan waktu survival yang biasanya diukur dalam hari, minggu, bulan maupun tahun. *Event* yang diperhatikan dalam analisis survival dapat berupa kematian, munculnya suatu penyakit, kambuhnya suatu penyakit, perceraian atau kejadian-kejadian yang lain (Klein dan Kleinbaum, 2005).

Tujuan dari analisis survival adalah (1) memperkirakan fungsi survival dan fungsi *hazard*, (2) membandingkan fungsi survival dan fungsi *hazard*, dan (3) melihat hubungan antara suatu variabel terhadap waktu survival. Contoh-contoh analisis survival antara lain adalah mengetahui respon seseorang terhadap pengobatan yang diberikan, misalnya pengaruh kemoterapi yang dijalani terhadap pasien yang mengalami leukemia. Contoh yang lain adalah meneliti kekambuhan suatu penyakit pada seseorang, meneliti berapa lama pasien bertahan hidup setelah menerima transplantasi hati, melepaskan tahanan yang dibebaskan secara bersyarat untuk beberapa minggu untuk mengetahui apakah dia akan ditangkap lagi, dan lain sebagainya.

Perbedaan analisis survival dengan analisis statistika yang lain terletak pada konsep penyensoran. Data tersensor adalah data yang tidak bisa diamati secara keseluruhan karena terdapat subyek pengamatan yang hilang atau tidak lengkap. Ada berbagai alasan mengapa sensor mungkin terjadi, yaitu : (1) subyek pengamatan tidak mengalami *event* apapun sebelum penelitian berakhir (2) subyek pengamatan menghilang selama penelitian terjadi dan (3) subyek pengamatan ditarik dari penelitian karena meninggal dimana meninggal merupakan suatu peristiwa yang tidak diperhatikan oleh peneliti.

Terdapat beberapa metode yang digunakan untuk menganalisis data survival. Metode-metode tersebut antara lain metode parametrik, metode nonparametrik serta metode semiparametrik. Metode parametrik mengasumsikan bahwa distribusi yang mendasari waktu survival mengikuti suatu distribusi tertentu misalnya *Weibull*, *gamma*, eksponensial dan lain sebagainya. Metode nonparametrik digunakan apabila data yang digunakan tidak mengikuti suatu distribusi tertentu yang sudah ada. Metode-metode yang termasuk dalam metode nonparametrik antara lain adalah metode Kaplan-Meier (*Product Limit*) serta *Nelson-Aalen*. Selain kedua metode tersebut, analisis regresi dapat digunakan untuk menyelesaikan persoalan analisis survival. Analisis regresi merupakan suatu analisis statistika yang memanfaatkan hubungan antara dua atau lebih peubah kuantitatif sehingga salah satu peubah dapat diramalkan dari peubah lainnya. Salah satu analisis regresi yang terkenal untuk menganalisa data survival adalah regresi Cox. Regresi Cox termasuk ke dalam metode semiparametrik yang mana fungsi

*baseline hazard* mengikuti model nonparametrik sedangkan variabel-variabel independennya mengikuti model parametrik. Tujuan dari metode regresi Cox adalah untuk mengetahui hubungan antara waktu survival dengan variabel-variabel yang diduga mempengaruhi waktu survival. Seperti pada regresi linier biasa, regresi Cox memiliki variabel terikat (variabel dependen) yaitu waktu survival dan variabel bebas (variabel independen) yaitu variabel yang diduga mempengaruhi waktu survival. Regresi Cox dikenal juga dengan istilah regresi Cox *proportional hazard* karena asumsi *proporsional hazard*. Asumsi *proporsional hazard* merupakan asumsi terpenting yang dipenuhi dalam regresi Cox yang berarti seiring berjalannya waktu tingkat kematian (*event*) adalah konstan.

Pada analisis survival, dimungkinkan terdapat dua individu atau lebih mengalami *event* (kejadian) pada waktu yang sama. Hal ini disebut dengan kejadian bersama yang seringkali dihindari karena dapat mengakibatkan permasalahan pada saat pembentukan *partial likelihood*. Disebut dengan “*partial*” karena fungsi *likelihood* yang digunakan hanya sebagian saja yaitu pada data yang tersensor. Pada umumnya terdapat tiga metode untuk menyelesaikan kejadian bersama di model *Cox proportional hazard*, yaitu metode *Exact*, metode *Breslow*, dan metode *Efron*. Metode *Exact* merupakan metode yang paling akurat akan tetapi memiliki perhitungan yang cukup rumit dan tidak praktis untuk data yang besar. Sedangkan metode *Breslow* dan *Efron* merupakan metode yang lebih sederhana dan perhitungannya

cepat, akan tetapi metode *Efron* lebih akurat daripada metode *Breslow* (Allison, 2010).

Pada penelitian ini akan digunakan model *Cox proportional hazard* dengan menggunakan pendekatan *Efron* pada kasus kecelakaan lalu lintas. Kecelakaan merupakan suatu kejadian yang memakan banyak korban jiwa dan termasuk ke dalam sepuluh penyebab utama kematian. Kecelakaan dapat terjadi dimana saja, kapan saja dan menimpa siapa saja. Tidak terkecuali di suatu negara maju seperti Amerika Serikat. Amerika merupakan negara terpadat ketiga di dunia yang memiliki peraturan berlalu lintas yang sangat lengkap dan fasilitas yang cukup memadai, akan tetapi angka kematian yang disebabkan oleh kecelakaan lalu lintas masih cukup tinggi. Karena tingginya angka kematian yang disebabkan oleh kecelakaan lalu lintas, Amerika menduduki peringkat pertama di dunia sebagai negara dengan banyaknya korban meninggal yang disebabkan oleh kecelakaan. Setiap tahun di Amerika terjadi kecelakaan lalu lintas lebih dari sepuluh juta kasus kecelakaan dan terdapat korban meninggal lebih dari 50.000 orang yang disebabkan oleh kecelakaan lalu lintas.

Terdapat tiga faktor utama yang menjadi penyebab kecelakaan lalu lintas, yaitu faktor pengemudi (*road user*), faktor kendaraan (*vehicle*) dan faktor lingkungan jalan (*road environment*). Kecelakaan tidak terjadi hanya disebabkan oleh satu faktor, namun kecelakaan terjadi karena interaksi dengan faktor yang lain. Manusia merupakan faktor terbesar penyebab terjadinya kecelakaan. Adapun faktor-faktor yang sering menjadi penyebab

terjadinya kecelakaan antara lain lengah, mengantuk, mengonsumsi alkohol dan obat-obat terlarang, melanggar rambu-rambu lalu lintas, tidak memakai helm, tidak memakai sabuk pengaman.

Penelitian sebelumnya yang berkaitan dengan kecelakaan dilakukan oleh Haryono Sukarto (1993) meneliti tentang interaksi faktor-faktor penyebab kecelakaan lalu lintas di jalan tol sekitar Jakarta yang menggunakan metode analisis komponen utama (*principle component analysis*) memperlihatkan bahwa faktor penyebab kecelakaan yang paling besar pengaruhnya adalah faktor pengemudi. Penelitian lain oleh Jeffry Rio H. Manurung (2012) membahas tentang hubungan faktor-faktor penyebab dan akibat kecelakaan lalu lintas pada pengendara sepeda motor di kota Medan menggunakan metode *cross sectional* menunjukkan bahwa kecelakaan lalu lintas disebabkan oleh interaksi antara faktor manusia, lingkungan fisik dan kendaraan, dimana faktor manusia berperan sangat penting dalam terjadinya kecelakaan lalu lintas. Penelitian lain yang dilakukan oleh Ahmad Wahidin (2008) membahas tentang pengaruh penggunaan sabuk pengaman (*safety belt*) terhadap tingkat fatalitas kecelakaan dan tingkat keparahan kecelakaan menggunakan metode pendekatan analisis statistik (analisis univariat, bivariat dan multivariat). Hasil dari penelitian tersebut adalah terdapat pengaruh penggunaan sabuk keselamatan terhadap tingkat fatalitas kecelakaan dan tingkat keparahan kecelakaan. Penelitian lain dilakukan oleh M. Bayu Iskandar (2014) membahas penyebab terjadinya kecelakaan lalu lintas di Amerika Serikat dengan model *Cox proportional hazard* menggunakan

pendekatan *Breslow* memperlihatkan bahwa usia dan penggunaan sabuk pengaman menjadi faktor penyebab terjadinya kecelakaan lalu lintas di Amerika Serikat. Berdasarkan uraian tersebut, penulis ingin membahas penyebab kecelakaan lalu lintas yang terjadi di Amerika Serikat dengan analisis survival menggunakan model *Cox proportional hazard* dengan pendekatan *Efron*.

## **B. Rumusan Masalah**

Berdasarkan uraian latar belakang maka rumusan masalah yang akan dibahas dalam skripsi ini sebagai berikut :

1. Bagaimana menyusun model *Cox proportional hazard* dengan kejadian bersama menggunakan pendekatan *Efron*?
2. Bagaimana penerapan model *Cox proportional hazard* pada kasus kecelakaan di Amerika Serikat?

## **C. Batasan Masalah**

Di dalam penelitian ini hanya akan digunakan model *Cox proportional hazard* dengan pendekatan *Efron* pada kasus kecelakaan lalu lintas yang terjadi di Amerika Serikat. Data kecelakaan yang diambil adalah data kecelakaan lalu lintas di Amerika Serikat pada tahun 2012 dari website *National Highway Traffic Safety Administration* (NHTSA.gov). Untuk mempermudah perhitungannya akan digunakan *software R 3.1.2* yang mana R adalah sebuah *software* matematika untuk mengolah data statistik.

#### **D. Tujuan Penulisan**

Tujuan dari penulisan skripsi ini adalah sebagai berikut :

1. Menjelaskan pemodelan *Cox proportional hazard* pada kasus kejadian bersama.
2. Mengetahui hasil penerapan model *Cox proportional hazard* pada kasus kecelakaan lalu lintas di Amerika Serikat.

#### **E. Manfaat Penulisan**

Penulisan skripsi ini diharapkan dapat memberikan manfaat sebagai berikut :

1. Bagi Mahasiswa

Pengembangan ilmu pengetahuan yang dipelajari diperkuliahan dan menambah wawasan tentang analisis survival khususnya model *Cox proportional hazard*.

2. Bagi Penulis

Menambah pengetahuan tentang analisis survival khususnya model *Cox proportional hazard* dengan pendekatan *Efron*.

3. Bagi perpustakaan Jurusan Pendidikan Matematika UNY

Menambah referensi mengenai model *cox proportional hazard* bagi mahasiswa Jurusan Pendidikan Matematika.

4. Bagi Pembaca

Penelitian ini dapat dijadikan sebagai referensi pembaca untuk mengembangkan model *Cox proportional hazard* pada kasus-kasus yang lainnya.