BAB III
LANDASAN TEORI

3.1 Stroke Non Hemoragik (SNH)

3.1.1. Pengertian Stroke Non Hemoragik

Stroke adalah penyakit gangguan fungsional otak sebagian maupun global akut dengan gejala dan tanda sesuai bagian otak yang terkena, yang sebelumnya tanpa peringatan dan dapat sembuh sempurna, sembuh dengan cacat, atau kematian, yang diakibatkan oleh gangguan aliran darah ke otak karena pendarahan ataupun nun pendarahan (Junaidi, 2004). Dari beberapa definisi *stroke*, dapat diambil kesimpulan bahwa *stroke* merupakan suatu serangan mendadak yang terjadi di otak dan dapat mengakibatkan kerusakan pada sebagian atau secara keseluruhan dari otak yang disebabkan oleh gangguan peredaran pada pembuluh darah yang mensuplai darah ke otak. Secara garis besar berdasarkan kelainan patologis yang terjadi, *stroke* diklasifikasikan menjadi 2 yaitu, stroke *Iskemik (stroke non hemoragik)* dan *stroke Hemoragik* (Wahjoepramono, 2005).

Stroke iskemik (stroke non hemoragik) disebabkan karena aliran darah menjadi menurun atau bahkan terhenti sama sekali pada area tertentu di otak, misalnya karena terjadi emboli atau trombosis (Wahjoepramono, 2005). Hal tersebut dapat menyebabkan terhambatnya aliran darah menuju otak yang mengakibatkan sel saraf dan sel lainnya mengalami gangguan karena terhentinya suplai oksigen dan glukosa yang dibawa oleh darah. Penurunan atau terhentinya aliran darah ini dapat menyebabkan *neuron* berhenti berfungsi. Gangguan suplai darah tersebut berlangsung hingga melewati batas toleransi sel, maka akan terjadi kematian sel. Apabila aliran darah dapat diperbaiki segera, maka kerusakan yang terjadi dapat sangat minimal. Berdasarkan perjalanan klinisnya, *stroke iskemik* dikelompokkan menjadi 4, yaitu (Junaidi, 2004) :

1. *Transient Ischemic Attack* (TIA): serangan *stroke* sementara yang berlangsung kurang dari 24 jam.
2. *Reversible Ischemic Neurologic Deficit* (RIND): gejala neurologis yang akan menghilang antara > 24 jam sampai dengan 21 hari.
3. *Progressing* stroke atau Stroke Nievolution: kelainan atau defisit neurologis yang berlangsung secara bertahap dari yang ringan sampai yang berat.
4. *Completed Stroke* : kelainan neurologis yang sudah menetap dan tidak berkembang lagi.

3.1.2. Tanda dan Gejala Stroke Non Hemoragik

Menurut Masayu Prakasita (2015) menjelaskan tanda dan gejala yang timbul dapat berbagai macam tergantung dari berat ringannya lesi dan juga topisnya. Namun ada beberapa tanda dan gejala yang umum dijumpai pada penderita stroke non hemoragik yaitu:

1. Gangguan Motorik
    a. Tonus abnormal (hipotonus/ hipertonus)
    b. Penurunan kekuatan otot
    c. Gangguan gerak volunter
    d. Gangguan keseimbangan
    e. Gangguan koordinasi
    f. Gangguan ketahanan
2. Gangguan Sensorik
    a. Gangguan propioseptik
    b. Gangguan kinestetik
    c. Gangguan diskriminatif
3. Gangguan Kognitif, Memori dan Atensi
    a. Gangguan atensi
    b. Gangguan memori
    c. Gangguan inisiatif
    d. Gangguan daya perencanaan

e. Gangguan cara menyelesaikan suatu masalah

4. Gangguan Kemampuan Fungsional
   a. Gangguan dalam beraktifitas sehari-hari seperti mandi, makan, ke toilet, dan berpakaian.

3.1.3 Faktor Risiko Stroke Non Hemoragik

Menurut Bustami (2007) faktor resiko penyakit stroke non hemoragik terbagi 2 faktor yaitu faktor risiko *internal* dan faktor risiko *eksternal* sebagai berikut:

1. Faktor risiko *internal*, yang tidak dapat dikontrol/diubah/dimodifikasi yang meliputi usia, jenis kelamin, ras, dan genetik.
2. Faktor risiko *eksternal*, yang dapat dikontrol/diubah/dimodifikasi yang meliputi hipertensi, merokok, diabetes melitus, fibrilasi atrium, kelainan jantung, hiperlipidemia, terapi pengganti hormon, anemia sel sabit, nutrisi, obesitas, dan aktifitas fisik.

3.2 Analisis Deskriptif

Penelitian deskriptif merupakan metode penelitian yang digunakan untuk menemukan pengetahuan yang luas terhadap objek penelitian pada suatu masa tertentu (Syah, 2010). Menjelaskan bahwa statistik deskriptif adalah bagian dari statistika yang mempelajari cara pengumpulan data dan penyajian data sehingga mudah dipahami. Statistika deskriptif hanya berhubungan dengan hal menguraikan atau memberikan keteranganketerangan mengenai suatu data atau keadaan. Statistika deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu data sehingga memberikan informasi yang berguna (Walpole, 1995). Pengukuran ini bertujuan dalam memberikan gambaran tentang data yang diperoleh, baik dari sampel maupun populasi.

Analisis deskriptif adalah statistik yang digunakan untuk menganalisa data dengan cara mendeskripsikan atau menggambarkan data yang telah terkumpul sebagaimana adanya tanpa bermaksud membuat kesimpulan yang berlaku untuk umum atau generalisasi (Sugiyono, 2006).

3.3 Analisis Survival

Analisis *survival* merupakan alat penting dalam statistika yang digunakan untuk menganalisis data waktu antar kejadian (*time to event data*) atau menganalisis data yang berhubungan dengan waktu, mulai dari *time origin* sampai terjadinya satu peristiwa khusus. Kejadian khusus (*failure event)* tersebut dapat berupa kegagalan, kematian, kambuhnya suatu penyakit, respon dari suatu percobaan, atau peristiwa lain yang dipilih sesuai dengan kepentingan peneliti. Peristiwa khusus tersebut tidak selalu berupa kejadian negatif, tetapi juga dapat berupa kejadian positif seperti kelahiran, kelulusan sekolah, kesembuhan dari suatu penyakit (Kleinbaum & Klein, (2005), dalam Nurjanah (2015)).

Analisis *survival* banyak diterapkan dalam bidang biologi, kedokteran, kesehatan umum seperti daya hidup pasien kanker paru-paru, sosiologi, teknik, seperti menganalisis masa hidup lampu pijar, ekonomi, demografi, dan epidemilogi (Collett, 2003).

Dalam analisis *Survival* biasanya variabel waktu disebut juga sebagai waktu *Survival* karena mengindikasikan bahwa seorang individu telah bertahan *(survived)* selama periode pengamatan. Waktu *Survival* adalah waktu yang diperoleh dari suatu pengamatan terhadap objek yang dicatat dari awal sampai terjadinya *event* (Collet, (2003), dalam Nurfain (2017)).

Ada tiga faktor yang harus diperhatikan dalam menentukan waktu *Survival* menurut (Cox, 1972), yakni:

1. Waktu awal *(time origin/starting point)* suatu kejadian,
2. *Event* dari keseluruhan kejadian harus jelas,
3. Skala pengukuran sebagai bagian dari waktu harus jelas

Data *survival* terdapat kemungkinan ditemukan data tersensor, dan tersensor merupakan data yang tidak dapat teramati secara utuh, karena adanya individu yang hilang ataupun dengan alasan yang lain, sehingga tidak dapat terambil datanya sampai akhir penelitian. Pada akhir pengamatan individu tersebut belum mengalami *failure*

event. Berbeda ketika sampai akhir pengamatan individu tersebut telah mengalami *failure event*, maka individu tersebut tidak tersensor (Collet, 2003).

Analisis *survival* terdapat empat jenis penyensoran. (Klein *and* Moeschberger (2003), dalam Iskandar (2014)).

a. Penyensoran kanan (*right censoring*)

Penyensoran ini terjadi jika objek pengamatan atau individu yang diamati masih tetap hidup pada saat waktu yang telah ditentukan. Dengan kata lain individu tersebut belum mengalami kejadian sampai akhir periode pengamatan, sedangkan waktu awal dari objek pengamatan dapat diamati secara penuh. Contoh, seorang pasien kanker diamati dari awal perawatan sampai akhir perawatan ternyata pasien tersebut masih hidup. Kemudian pasien melanjutkan perawatan di luar negeri sehingga tidak bisa diamati lagi (*lost to follow up*). Pasien ini memiliki waktu *survival* setidaknya beberapa waktu. Sehingga waktu pengamatan individu tersebut dikatakan penyensoran kanan.

b. Penyensoran kiri (*left censoring*)

Penyensoran kiri terjadi jika semua informasi yang diinginkan diketahui dari seseorang individu telah diperoleh pada awal pengamatan. Dengan kata lain pada saat waktu awal pengamatan individu tidak teramati pada awal pengamatan sementara kejadian dapat diamati secara penuh sebelum penelitian berakhir. Contoh, dalam sebuah penelitian yang bertujuan untuk menentukan sebaran pengguna ganja di kalangan anak laki-laki di sebuah sekolahan. Dengan mengajukan pertanyaan "kapan pertama kali anda menggunakan ganja". Ternyata terdapat beberapa anak menjawab "saya pernah menggunakannya, tetapi saya tidak tahu tepatnya kapan pertama kali mengunakannya", pada kasus ini anak tersebut mengalami penyensoran kiri.

c. Penyensoran selang (*interval censoring*)

Penyensoran selang terjadi jika informasi yang dibutuhkan telah dapat diketahui pada kejadian peristiwa di dalam selang pengamatan atau penyensoran yang waktu

daya tahannya berada dalam suatu selang tertentu. Contoh, beberapa tikus yang diberikan karsinogen pada makanannya, dilakukan studi selama 10 bulan kepada 10 tikus dan penelitian dilakukan setiap akhir tahun, jika 2 dari 8 tikus tewas karena kanker pada bulan ke-5 dan ke-7, maka dua tikus tersebut mengalami penyensoran selang.

d. Penyensoran acak (*random censoring*)

Penyensoran acak terjadi jika individu yang diamati meninggal atau mengalami kejadian karena sebab yang lain, bukan disebabkan dari tujuan utama penelitian. Contoh, 10 tikus yang diberikan zat karsinogen pada makanannya. Pada saat pengamatan ada 1 dari 10 tikus tersebut meninggal karena terjepit (tewas bukan karena penelitian utama) bukan karena terkena kanker, maka tikus tersebut mengalami pensensoran acak.

Menurut Klein dan Kleinbaum (2005) pada data survival . terdapat tiga alasan terjadinya suatu penyensoran pada data survival, yaitu:

a. Subyek pengamatan yang diamati tidak mengalami suatu *event* sampai penelitian berakhir (*loss to follow-up*).
b. Subyek pengamatan hilang selama penelitian.
c. Subyek pengamatan ditarik dari penelitian karena meninggal dimana meninggal merupakan suatu peristiwa yang tidak diperhatikan oleh peneliti atau alasan yang lain, misalnya reaksi obat yang buruk atau resiko yang lain.

3.4 Dasar Teori Analisis Survival

Untuk mendeskripsikan suatu variabel random T yang menunjukkan waktu *survival* setiap individu dalam bentuk eksplisit yang berupa model matematika, digunakan fungsi variabel random tersebut, misalnya berupa fungsi distribusi dan fungsi probabilitas. Menurut Lee dan Wang (2003), terdapat tiga cara menentukan distribusi dari T yaitu fungsi kepadatan peluang, fungsi survival dan fungsi hazard.

3.4.1. Fungsi Kepadatan Peluang (PDF)

Fungsi kepadatan peluang merupakan peluang suatu individu mengalami *event*, gagal atau mati dalam interval waktu *t* sampai (t+ Δ*t*), fungsi kepadatan peluang dinotasikan dengan *f(t).* Fungsi ini dirumuskan sebagai berikut:

$$f(t) = \lim_{\Delta t \to 0} \left[\frac{P(t<T<(t+\Delta t))}{\Delta t} \right] = \lim_{\Delta t \to 0} \left[\frac{F(t+\Delta t - F<(t)}{\Delta t} \right] \quad (3.1)$$

dengan,

f(t) : peluang individu mengalami *event*,

F(t) : peluang variabel individu akan lebih kecil dengan nilai T,

T : variabel random yang menunjukkan waktu *survival*,

t : waktu yang ditentukan,

Δ*t* : interval waktu individu sampai mengalami *event*

Jika T merupakan variabel random non negatif dalam interval [0,∞), *F(t)* merupakan fungsi distribusi kumulatif (cdf) dari *T*. Didefinisikan sebagai peluang suatu individu mengalami event sampai dengan waktu *t* yang dapat dituliskan sebagai berikut.

$$F(t) = P(T \leq t) = \int_0^t f(t)dt \quad (3.2)$$

Berdasarkan persamaan (3.2) diatas, dapat diperoleh:

$$f(t) = \frac{d(F(t))}{dt} = F'(t) \quad (3.3)$$

3.4.2. Fungsi *Survival*

Fungsi *Survival* $S(t)$ adalah probabilitas satu individu hidup *(survive)* lebih lama daripada *t* , yaitu sebagai berikut.

$$S(t) = P(T \geq t) \quad = \int_t^{\infty} f(x)dx \quad (3.4)$$

Sesuai dengan definisi fungsi distribusi kumulatif F(t) dari T, fungsi *survival* dapat dinyatakan dengan,

$$S(t) = 1 - P(T \geq t)$$
$$= 1 - F(t)$$

Menurut Klein dan Kleinbaum (2012) dalam Nurfain (2017), secara teori fungsi *Survival* dapat diplot sebagai kurva *Survival* yang menggambarkan peluang ketahanan suatu individu pada waktu *t* dalam interval 0 sampai ∞. Fungsi *Survival* mempunyai beberapa karakteristik, yaitu:

a. Fungsi *Survival* merupakan fungsi monoton tak naik.
b. Pada saat t=0,$S(t)$=S(0)=1

   Pada awal dimulainya penelitian, karena belum ada individu yang mengalami *event* maka probabilitas *Survival* pada saat t=0 adalah 1.
c. Pada saat t=∞,$S(t)$=0

   Secara teori, apabila periode penelitian meningkat tanpa batas, maka diakhir waktu tidak ada seorang individu yang akan bertahan hidup, sehingga kurva *Survival* akan bergerak menuju nol.

3.4.3. Fungsi *Hazard*

Menurut Lawles (2007), misalkan T variabel *random* tidak negatif pada interval [0, ∞) yang menunjukkan waktu individu sampai mengalami kejadian pada suatu populasi, maka tingkat bahwa individu mengalami kejadian pada interval (t + Δt) dinyatakan dengan fungsi *hazard* $h(t)$.

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t | T \ge t)}{\Delta t}$$

$$= \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t, T \ge t)}{\Delta t . P(T \ge t)}$$

$$= \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t)}{\Delta t . S(t)}$$

$$= \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t . S(t)}$$

$$= \frac{1}{S(t)} \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t}$$

$$= \frac{F'(t)}{S(t)} = \frac{f(t)}{S(t)} \tag{3.5}$$

3.5 Regresi Cox

Pada awalnya regresi Cox diperkenalkan oleh Cox. Seperti metode regresi lainnya, regresi Cox digunakan untuk mengetahui efek dari beberapa variabel prediktor terhadap variabel respon. Variabel respon dalam regresi Cox adalah waktu *Survival* suatu objek terhadap suatu peristiwa tertentu. Regresi Cox termasuk regresi semiparametrik dimana dalam pemodelannya terdapat komponen parametrik dan non-parametrik. Regresi ini tidak memiliki asumsi mengenai sifat dan bentuk sesuai dengan distribusi seperti asumsi pada regresi yang lain. Maka hal tersebut membuat regresi Cox baik digunakan apabila distribusi dari waktu *survival* tidak diketahui secara pasti sehingga hasil estimasi parameter regresi masih dapat dipercaya (Lee, 2003).

Nurfain menguraikan dalam Kleinbaum dan Klein (2012) bahwa metode Regresi Cox tanpa perlu diketahui $h_0(t)$ dapat mengestimasi $h_0(t)$, $h_0(t,x)$, dan fungsi *survival* meskipun $h_0(t)$ tidak spesifik, hasil dari cox model hampir sama dengan hasil model parametrik. Penaksiran fungsi hazard dapat dipakai untuk menghitung risiko relatif terjadinya kejadian.

3.6 Regresi Cox Proportional Hazard

Lee dan Wang (2003) menjelaskan bahwa model *Cox Proportional Hazard* disebut juga dengan model *Cox* karena asumsi *Proportional Hazard*nya yaitu fungsi *hazard* dari individu yang berbeda adalah *proportional* atau rasio dari fungsi *hazard* dua individu yang berbeda adalah konstan. Iskandar (2015), menjelaskan bahwa persamaan regresi cox merupakan model berdistribusi semiparametrikkarena dalam persamaan cox tidak memerlukan informasi mengenai distribusi khusus yang mendasari waktu survival dan untuk mengestimasi parameter regresi cox tanpa harus menentukan fungsi hazard dasar.

Klein dan Kleinbaum (2005) menyatakan persamaan C*ox Proportional Hazard* dapat dituliskan sebagai berikut:

$$h(t,X) = h_0(t)\exp\left(x_1\beta_1 + x_2\beta_2 + \cdots + x_p\beta_p\right)$$

$$= h_0(t)\exp\left(\sum_i^p x_i\beta_i\right) \quad (3.6)$$

dengan,

$h(t,X)$: resiko kematian individu pada waktu t dengan karakteristik X

$h_0(t)$: fungsi *hazard baseline*

β_1 : koefisien regresi, dengan i=1,2,3,…, p

x_i : nilai variabel untuk individu ke-i, dengan i=1,2,3,…, p

3.7 Estimasi Parameter

Parameter regresi pada persamaan regresi cox *proportional hazard* dapat diketahui dengan menggunakan metode *Maximum Partial Likelihood Estimator (MPLE)*. Pendugaan nilai β dengan metode MPLE adalah nilai ketika fungsi *partial likelihood* maksimum. Secara umum, fungsi *partial likelihood* dapat dinyatakan sebagai berikut :

$$L(\beta) = \prod_{\boldsymbol{k}\in\boldsymbol{D}} \frac{\boldsymbol{exp}(\boldsymbol{x_k\beta})}{\sum_{\boldsymbol{j}\in\boldsymbol{R_k}}\boldsymbol{exp}(\boldsymbol{x_j\beta})} \quad (3.7)$$

dengan,

x : vektor kovariat atau variabel penjelas,

β : parameter regresi yang akan diestimasi,

D : himpunan indeks j dari semua waktu kejadian (semua t_k yang mendapatkan *event*),

R_k : himpunan risiko semua individu yang belum mendapatkan kejadian pada saat tertentu.

Dari persamaan (3.9) ditransformasikan ke dalam bentuk ln untuk mempermudah pencarian penduga maksimum L(β), sehingga persamaannya menjadi sebagai berikut:

$$lnL(\beta) = \ln\left(\Pi_{k\in D}\frac{\exp(x_k\beta)}{\sum_{j\in R_k}\exp(x_j\beta)}\right)$$

$$= \sum_{j=1}^{p}\left(\ln\frac{\exp(x_k\beta)}{\sum_{j\in R_k}\exp(x_j\beta)}\right)$$

$$= \sum_{j=1}^{p}\left[x_k\beta - \sum_{j=1}^{p}\ln\left(\sum_{j\in R_k}\exp(x_j\beta)\right)\right] \quad (3.8)$$

Nilai penduga β dapat diperoleh dengan memaksimumkan fungsi *log partial likelihood*nya, dengan mencari solusi dari persamaan berikut:

$$\frac{\partial \ln(\beta)}{\partial\beta} = 0 \quad (3.9)$$

Persamaan tersebut dapat diselesaikan secara numerik atau secara komputasi dengan menggunakan bantuan *software*.

3.8 Ties dalam Partial Likelihood

Dalam analisis *survival* sangat sering ditemukan kejadian bersama atau disebut dengan *ties*. Kajadian *ties* merupakan kejadian dimana dua individu atau lebih yang mengalami kejadian pada waktu yang bersamaan. Terjadinya *ties* pada data *survival* akan menimbulkan permasalahan dalam *partial likelihood* yaitu saat menentukan anggota dari suatu himpunan risikonya. Sebagai contoh, dapat dilihat pada tabel di bawah ini :

Tabel 3.1. Contoh Data *Ties* Analisis *Survival*

i	1	2	3	4
ti	5	5	7+	10

Dari tabel 3.1 misalkan, i adalah individu ke-i dan t_i merupakan waktu kejadian untuk individu ke-i. Jika $t_1 < t_2 < t_3 < t_4$ merupakan waktu yang diamati yang telah diurutkan. Pada waktu t=5, terdapat dua individu yang mengalami kejadian dan tidak diketahui individu mana yang mengalami kejadian lebih dahulu. Kejadian bersama tersebut dapat menimbulkan masalah pada estimasi paremeter yang berhubungan

dengan penentuan anggota dari himpunan risiko. Terdapat beberapa metode alternatif untuk menangani *ties* menurut Klein dan Mochberger (2003) meliputi :

3.8.1 Metode *Partial Likelihood Breslow*

Metode *Breslow* merupakan metode paling sederhana dibandingkan dengan metode lainnya, karena tingkat komputasi yang tidak *intensive*. Selain itu, metode ini memberikan hasil estimasi yang baik jika data kejadian *ties* dalam ukuran kecil. Secara umum bentuk persamaan *breslow* adalah sebagai berikut:

$$L(\beta_{\text{breslow}}) = \prod_{k \in D} \frac{\exp(S_k\, \beta)}{\left[\sum_{j \in R_k} \exp\,(S_k\, \beta)\right]^{d_k}} \tag{3.10}$$

dengan,

x : vektor kovariat (variabel penjelas),

β : parameter regresi yang akan di estimasi,

D : himpunan indeks k dari semua waktu kejadian (semua t_k yang mendapatkan *event*),

R_k : himpunan resiko (*risk set*) semua individu yang belum mendapatkan kejadian pada saat tertentu,

S_k : jumlah kovariat x pada kasus kejadian bersama,

d_k : banyaknya kasus kejadian bersama (*ties*) pada waktu $t_{k..}$

3.8.2 Metode Partial Likelihood Efron

Dalam pendekatan ini memang sedikit lebih *intensive* pada tingkat komputasinya dibandingkan dengan metode *breslow*. Namun jika dibandingkan dengan metode sebelumnya juga pendekatan metode ini mampu menghasilkan nilai estimasi yang baik ketika data kejadian *ties* dalam ukuran besar. Klein dan Moeschberger (2003), menyatakan pendekatan *efron* secara umum memiliki persamaan *partial likelihood* sebagai berikut :

$$L(\beta_{efron}) = \prod_{k\in D} \frac{\exp(S_k\,\beta)}{\prod_{j=1}^{d_k}\left[\sum_{i\in R_k}\exp\,(x_i\beta) - \frac{j-1}{d_k}\sum_{i\in D_k}\exp\,(x_i\beta)\right]} \quad (3.11)$$

dengan,

x : vektor kovariat (variabel penjelas),

β : parameter regresi yang akan di estimasi,

D : himpunan indeks k dari semua waktu kejadian (semua t_j yang mendapatkan *event*),

R_k : himpunan resiko (*risk set*) semua individu yang belum mendapatkan kejadian pada saat tertentu,

S_k : jumlah kovariat x pada kasus kejadian bersama,

d_k : banyaknya kasus kejadian bersama (*ties*) pada waktu t_k,

D_k : himpunan individu yang mendapat kejadian pada waktu t_k.

3.8.3 Metode Partial Likelihood Exact

Collet (2003) menyatakan metode *partial likelihood exact* sebagai alternatif kasus kejadian bersama. Metode ini memang memiliki tingkat komputasi yang sangat *intensive* namun mampu mengahsilkan estimasi parameter yang memiliki bias mendekati 0 meskipun data kejadian ties dalam ukuran yang sangat besar. Cox (1972) menyatakan persamaan umum dari metode *exact* adalah sebagai berikut :

$$L(\beta_{exact}) = \prod_{k\in D} \frac{\exp(S_k\,\beta)}{\left[\sum_{l\in R_{(t_k d_k)}}\exp\,(S_l\,\beta)\right]} \quad (3.12)$$

dengan,

β : parameter regresi yang akan di estimasi,

D : himpunan indeks k dari semua waktu kejadian (semua t_k yang mendapatkan *event*),

R_{tk} : himpunan resiko (*risk set*) semua individu yang belum mendapatkan kejadian pada saat tertentu,

S_k : jumlah kovariat x pada kasus kejadian bersama,

d_k : banyaknya kasus kejadian bersama (*ties*) pada waktu t_k,

S_l : merupakan himpunan resiko yang dipilih dari himpunan R_{tk}.

3.9 Pengujian Parameter

Dalam regresi cox *proportional hazard*, diperlukan beberapa pengujian signifikansi parameter agar dapat diketahui apakah variabel bebas berpengaruh nyata terhadap persamaan cox yang terbentuk. Beberapa pengujian yang digunakan dalam pengujian siginifikansi parameter yaitu :

1. Uji Serentak

Uji ini digunakan untuk mengetahui apakah satu atau beberapa parameter βj sama dengan nol yang menyatakan bahwa satu atau beberapa parameter βj tidak berpengaruh secara nyata terhadap persamaan cox yang terbentuk. Statistik uji yang digunakan mengikuti distribusi *chi-square* dengan derajad bebas *p*. Secara keseluruhan, berikut merupakan rangkaian pengujian hipotesis *partial likelihood rasio* :

i) Hipotesis

H_0:$\beta 1=\beta 2=\cdots=\beta k=0$

H_1: minimal ada satu $\beta j \neq 0$,dengan $j=1,2,\ldots k$

ii) Tingkat Signifikansi

$\alpha=5\%=0.05$

iii) Statistik Uji

$$G = -2[lnLR - lnLf] \quad (3.13)$$

Dimana, L_R merupakan *partial likelihood* model awal, dan L_f merupakan *partial likelihood* model akhir.

iv) Daerah Penolakan

H_0 ditolak jika nilai $G \geq X^2_{(\alpha,\text{db=p})}$ atau nilai *P-value* $\leq \alpha$

Dimana, p merupakan banyaknya variabel bebas.

v) Kesimpulan

Jika H_0 ditolak, maka $\beta j \neq 0$ yang mengindikasikan bahwa baik satu ataupun beberapa variabel bebas memberikan pengaruh secara nyata terhadap waktu survival (variabel dependen).

2. Uji Parsial

Uji Parsial digunakan untuk mengetahui variabel-variabel bebas apa saja yang berpengaruh secara nyata. Statistik uji pada uji ini dinotasikan dengan *z* yang juga mengikuti distribsui *chi-square* dengan derajad bebas *p*, dan pengujian dengan β_j dimana $j=1,2,...k$, berikut serangkaian pengujian hipotesisnya :

i) Hipotesis

$H_0 : \beta_j = 0$ (Variabel bebas *j* tidak berpengaruh terhadap waktu lama waktu mencari kerja)

$H_1 : \beta_j \neq 0$ (Variabel bebas *j* berpengaruh terhadap waktu lama waktu mencari kerja)

ii) Tingkat Signifikansi

$\alpha=5\%=0.05$

iii) Statistik Uji

$$Z2=\left(\frac{\widehat{\beta}_j}{SE(\widehat{\beta}_j)}\right)^2 \quad (3.14)$$

iv) Daerah Penolakan

H_0 ditolak jika nilai $Z^2 \geq X^2_{(\alpha,db=p)}$ atau nilai *p-value* $\leq \alpha$

v) Kesimpulan

Jika H_0 ditolak, maka $\beta j \neq 0$ yang mengindikasikan bahwa kovariat yang diuji berpengaruh secara nyata terhadap waktu survival (variabel dependen).

3.10 Pemilihan Persamaan Akhir Regresi Cox

Pada pemilihan persamaan akhir regresi cox *proportional hazard* menggunakan aturan *backward elimination*. Langkah paling pertama yaitu memasukan semua variabel pada persamaan, kemudian dikeluarkan satu per satu berdasarkan nilai $-2lnL(\beta)$ terbesar.

Secara keseluruhan jika tidak ada peningkatan dari nilai $-2lnL(\beta)$, maka prosedur *backward* dihentikan.

3.11 Asumsi Proportional Hazard

Menurut Klein dan Kleinbaum (2005) menyatakan bahwa dalam melakukan suatu pengecekan asumsi *Proportional Hazard* dapat dilakukan dengan menggunakan *Goodness of Fit*. Metode penaksiran *Goodness of Fit* ini menggunakan statistik uji dalam evaluasi asumsi *Proportional Hazard* sehingga lebih objektif. Statistik uji yang digunakan dalam metode ini adalah *schoenfeld residuals*. *Schoenfeld residuals* merupakan sekumpulan nilai untuk masing-masing individu pada setiap kovariat dalam persamaan regresi Cox Proportional Hazard. Nilai *Schoenfeld residuals* dari kovariat ke$-i$ untuk individu ke$-j$ adalah sebagai berikut.

$$r_{ij} = \delta_i\{x_{ij} - \hat{a_{ij}}\} \tag{3.15}$$

$$\text{dengan, } \hat{\alpha}_{\text{ji}} = \frac{\sum_{l\in R(t_i)} x_{ji}\, exp(\widehat{\beta x l})}{\sum_{l\in R(t_i)} exp(\widehat{\beta x l})}$$

dimana,

δ_i : status individu yang bernilai 0 jika tersensor dan 1 jika tidak tersensor.

x_{ij} : nilai dari peubah penjelas ke$-i$, i = 1,2, … , p, untuk individu ke$-j$.

$\hat{a}_{ij}$: rataan terboboti dari peubah penjelas ke$-i$ untuk individu dalam $R(t_j)$.

$R(t_j)$: himpunan individu yang beresiko mengalami kejadian pada saat t_j.

Asumsi *Proportional Hazard* terpenuhi maka *Schoenfeld residuals* untuk kovariat tersebut tidak akan berkorelasi dengan peringkat waktu ketahanan. Adapun langkah-langkah pengujiannya adalah sebagai berikut:

a. Mencari taksiran persamaan *Cox Proportional Hazard* dan menghitung *Schonefeld residuals* pada masing–masing inidvidu pada setiap kovariat.
b. Membuat peubah yang menyatakan peringkat dari waktu ketahanan.
c. Menguji korelasi antara variabel pada langkah pertama dan kedua dengan *Schoenfeld residuals*, dimana hipotesis nol adalah terdapat korelasi antara

Schoenfeld residuals dan *rank* waktu ketahanan sama dengan nol. Penolakan hipotesis nol berarti asumsi *proportional hazard* tidak terpenuhi.

Ukuran yang digunakan untuk mengecek asumsi *proportional hazard* adalah nilai p, dimana jika nilai $p < 0.05$ maka kovariat yang diuji tidak memenuhi asumsi *proportional hazard* (Kleinbaum dan Klein (2005) dalam Susenati (2015)). Berikut langkah pengujian hipotesisnya:

i) Hipotesis

$H_0 : \beta_j = 0$ (Asumsi *proportional hazard* terpenuhi).

$H_1 : \beta_j \neq 0$ (Asumsi *proportional hazard* tidak terpenuhi).

ii) Tingkat Signifikansi

α

iii) Statistik Uji

p-value

iv) Daerah Kritis

Tolak H_o jika nilai *p-value* $\leq \alpha$

v) Kesimpulan

Jika gagal menolak H_0 maka asumsi *proportional hazard* terpenuhi.

3.12 Hazard Ratio (HR)

Persamaan regresi cox $h(t, x) = h_0 t \ exp(\beta x)$ dapat diinterpretasikan sebagai *hazard ratio*. Lee dalam Lasmini (2013) menyatakan bahwa *hazard ratio* mampu menunjukkan adanya penurunan atau peningkatan resiko individu yang dikenai perlakuan tertentu. Misalkan terdapat dua individu dengan karakteristik tertentu maka dari persamaan umum cox *proportional hazard* diperoleh formula untuk menduga *hazard ration*ya sebagai berikut :

$$HR = \frac{hi(t)}{hj(t)}$$

$$= \frac{h_0 \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi})}{h_0 \exp(\beta_1 x_{1j} + \beta_2 x_{2j} + \cdots + \beta_p x_{pj})}$$

$$= exp\left\{\beta_1(x_{1i} - x_{1j}) + \cdots + \beta_p(x_{pi} - x_{pj})\right\} \quad (3.16)$$

Berkaitan dengan penggunaan variabel penjelas, dalam regresi cox secara umum cenderung menggunakan variabel kategorik. Menurut Allison dalam Lasmini (2013) menyatakan bahwa kovariat bersifat kuantitatif, akan lebih bermakna jika hazardnya dikurangi 1 dan dikalikan dengan 100% yang menyatakan perubahan presentase hazard penduga untuk penambahan satu satuan unit peubah tersebut. Terdapat 3 macam ketentuan tentang bertambahnya atau berkurangnya nilai *hazard*, yaitu sebagai berikut:

a. $\beta_j > 0$ maka setiap naiknya nilai x_j akan memperbesar nilai *hazard* atau semakin besar risiko seorang individu untuk mengalami *event*.
b. $\beta_j < 0$ maka setiap naiknya nilai x_j akan memperkecil nilai *hazard* atau semakin kecil risiko seorang individu untuk mengalami *event*.
c. $\beta_j = 0$ maka besar risiko seorang individu untuk hidup sama dengan risiko seorang individu untuk mengalami *event*.