

BAB I PENDAHULUAN

1.1 Latar Belakang

Analisis survival atau analisis ketahanan hidup adalah kumpulan prosedur statistik yang digunakan untuk menganalisis data berkaitan dengan waktu yaitu dimulai dari *start point* sampai *end point*. Waktu survival dapat berupa tahun, bulan, hari, jam atau bahkan menit yang diukur dari awal pengamatan hingga terjadinya suatu *event* atau kejadian. Kejadian yang diamati dapat berupa kematian, penyembuhan, pemulihan atau kondisi apapun yang mungkin terjadi pada seseorang (Kleinbaum & Klein, 2012). Variabel dependen umumnya mengacu pada lama waktu bertahan hidup hingga terjadinya suatu kejadian tertentu dan variabel independen merupakan faktor-faktor yang mempengaruhi waktu kejadian tersebut. Salah satu keunggulan analisis survival adalah kemampuannya untuk memasukkan data tersensor sehingga memungkinkan prediksi waktu kejadian yang lebih akurat dan menjadikannya sebagai model yang lebih disukai dalam analisis ketahanan hidup (Thiruvengadam et al., 2021). Terdapat tiga jenis data tersensor, yaitu sensor kanan, sensor kiri dan sensor interval. Data tersensor yang digunakan pada penelitian ini adalah sensor interval yang digunakan ketika waktu pasti terjadinya kejadian (*event*) tidak diketahui secara tepat namun diketahui terjadi dalam suatu rentang waktu tertentu (Singh & Totawattage, 2013). Pada kasus kematian bayi dikatakan sensor interval apabila waktu kematian bayi tidak dapat diamati secara tepat namun diketahui terjadi antara dua waktu pengamatan.

Salah satu metode yang paling umum digunakan adalah *Cox Proportional Hazard* karena dapat memperkirakan hazard ratio tanpa perlu mengetahui bentuk fungsi hazard dasar (*baseline hazard function*) (Thiruvengadam et al., 2021). Perkembangan *Cox Proportional Hazard* juga menyertakan hubungan antara variabel respon dan variabel prediktor dengan fungsi model yakni *Multivariate Adaptive Regression Splines* (Saputro et al., 2017). Penggunaan metode MARS didasarkan pada hubungan yang masih belum jelas antara variabel respon dan variabel prediktor, sehingga pendekatan ini lebih fleksibel dibandingkan metode parametrik yang mengharuskan asumsi distribusi tertentu (Hasyim & Prastyo, 2018). Kriner (2007) menyatakan bahwa dalam pemodelan MARS, variabel respon dapat berupa residual yang diperoleh dari model *Cox Proportional Hazard*. Pemodelan survival dengan pendekatan MARS dapat dilakukan dengan menggunakan residual dari model Cox sebagai variabel respon. Salah satu residual yang dapat digunakan sebagai variabel respon adalah martingale residual. Hal ini memungkinkan MARS untuk menangkap pola hubungan non-linear serta interaksi antara kovariat yang mungkin tidak dapat ditangkap oleh model Cox standar. Penggunaan MARS dalam analisis survival dapat meningkatkan akurasi prediksi serta memberikan pemahaman yang lebih mendalam terhadap faktor-faktor yang mempengaruhi waktu.



Pemodelan MARS telah banyak digunakan dalam analisis data untuk menangani hubungan yang bersifat linear dan non linear penelitian sebelumnya dari Nair & Yadav (2024) tentang analisis regresi Cox untuk menganalisis faktor-faktor yang mempengaruhi kualitas hidup pasien Covid-19 di wilayah Afrika serta penggunaan model regresi Cox dengan diagnostik residual

Schoenfeld. Penelitian selanjutnya Dauda et al. (2015) juga melakukan penelitian dengan mengembangkan dan mengevaluasi analisis survival menggunakan pendekatan Multivariate Adaptive Regression Splines (MARS) dengan residual Cox-Snell sebagai variabel respons dan membandingkannya dengan pendekatan yang menggunakan residual martingale serta model *Cox Proportional Hazard* pada data *Pneumocystis Cariniip Neumonia* (PCP). Penelitian Khoirunnisa et al., (2024) yang memodelkan ketahanan hidup pasien Covid-19 menggunakan regresi Cox dengan pendekatan MARS sehingga mendapatkan variabel yang paling signifikan dan paling berpengaruh yaitu status gagal napas.

Kasus kematian bayi adalah salah satu data survival yang dapat dianalisis menggunakan pendekatan survival dengan MARS menggunakan residual martingale sebagai variabel respon. Pendekatan ini memungkinkan untuk mengidentifikasi faktor-faktor yang memengaruhi lamanya waktu bertahan hidup bayi hingga meninggal terutama ketika hubungan antara kovariat dan waktu kejadian bersifat kompleks dan non-linear. Model MARS yang memanfaatkan residual martingale dari model *Cox Proportional Hazard* mampu menangkap pola interaksi dan ketidakaturan data yang sulit ditangani oleh model parametrik konvensional sehingga menghasilkan model yang lebih fleksibel dan akurat dalam memprediksi waktu kejadian.

Peningkatan status kesehatan ibu dan anak merupakan salah satu dari lima indikator prioritas pembangunan kesehatan 2020-2024. Angka Kematian Bayi (AKB) di Indonesia tercatat mencapai 16,85 per 1.000 kelahiran hidup. Berdasarkan data BPS (2021) di Provinsi Sulawesi Selatan pada tahun 2020 tercatat sebesar 18,20 per 1.000 kelahiran hidup. Angka ini menunjukkan penurunan signifikan dibandingkan tahun 2010 yang mencapai 45 kematian per 1.000 kelahiran hidup. Penurunan ini mencerminkan perbaikan akses dan kualitas pelayanan kesehatan ibu dan anak di wilayah tersebut. Namun, meskipun terjadi tren penurunan, kasus kematian bayi harus tetap menjadi perhatian serius karena setiap kematian bayi mencerminkan potensi masalah dalam layanan kesehatan dasar, gizi, sanitasi, serta pengetahuan ibu terhadap perawatan bayi baru lahir. AKB dihitung berdasarkan jumlah kematian bayi di bawah usia satu tahun per 1.000 kelahiran hidup dalam satu tahun. Angka ini merupakan salah satu indikator penting dalam menilai derajat kesehatan masyarakat (Irkan et al., 2022).

Penyebab kematian bayi ada 2 yaitu langsung (endogen) dan tidak langsung (eksogen). Kematian bayi endogen atau kematian neonatal (0-28 hari) disebabkan oleh faktor-faktor yang dibawa anak sejak lahir yang diperoleh dari orang tuanya pada saat konsepsi. Kematian bayi yang berasal dari kondisi bayinya sendiri yaitu BBLR, bayi prematur, dan kelainan koagenital. Kematian bayi yang dibawa sejak lahir adalah asfiksia. Kematian bayi eksogen atau kematian post-neonatal (29 hari-330 hari) disebabkan oleh faktor-faktor yang berkaitan dengan pengaruh lingkungan luar (Susanty et al., 2016). Berdasarkan latar belakang diatas maka penelitian ini berjudul Analisis Survival Dengan Pendekatan Multivariate Adaptive Regression Splines (MARS) menggunakan Residual Martingale Pada Kasus Kematian Bayi.



salah

alah berdasarkan latar belakang yang telah diuraikan:
 model survival dengan pendekatan MARS menggunakan
 gale?
 g mempengaruhi kematian bayi di RS Labuang Baji Makassar?

1.3 Batasan Masalah

1. Data yang digunakan merupakan data sekunder kematian bayi di RS Labuang Baji Makassar tahun 2021-2024.
2. *Failure event* dalam penelitian ini yaitu ketika bayi meninggal di antara usia 29 hari sampai 330 hari.
3. Sensor yang digunakan adalah sensor interval artinya waktu meninggal tidak diketahui secara tepat namun terjadi dalam suatu rentang waktu tertentu.
4. Penelitian ini tidak membandingkan metode MARS dengan metode analisis survival lainnya melainkan hanya fokus pada pendekatan MARS dengan evaluasi residual martingale.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan maka tujuan dari penelitian ini adalah sebagai berikut:

1. Mengestimasi model survival dengan pendekatan MARS menggunakan residual martingale
2. Mengidentifikasi faktor-faktor yang mempengaruhi kematian bayi di RS Labuang Baji Makasar.

1.5 Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat sebagai berikut:

1. Menambah wawasan dan literatur mengenai penggunaan analisis survival dengan pendekatan MARS menggunakan residual martingale.
2. Memberikan informasi mengenai faktor-faktor yang mempengaruhi kematian bayi di RS Labuang Baji Makassar, sehingga dapat dijadikan dasar dalam upaya pencegahan, peningkatan kualitas layanan, dan pengambilan keputusan di bidang kesehatan ibu dan anak.
3. Memberikan implikasi praktis terhadap kebijakan rumah sakit khususnya dalam mengidentifikasi variabel risiko utama yang berkontribusi terhadap kematian bayi sehingga dapat digunakan sebagai indikator prioritas dalam pemantauan klinis.

1.6 Teori

Sub-bab teori ini membahas mengenai beberapa teori yang digunakan untuk membahas pokok pembahasan terkait analisis survival dengan pendekatan Multivariate Adaptive Regression Spline (MARS) menggunakan residual martingale.

1.6.1 Analisis Survival

Survival berasal dari kata *survive* yang berarti ketahanan atau kelangsungan hidup (Kleinbaum & Klein, 2012). Analisis survival merupakan salah satu metode statistika yang dapat digunakan untuk menganalisis data yang berhubungan dengan waktu. Waktu dapat berupa tahun, bulan, hari, jam, atau bahkan menit yang diukur sejak pengamatan dimulai hingga muncul kejadian. Kejadian yang dimaksud berupa perkembangan suatu penyakit, respon penyakit terhadap



perikan, kambuhnya suatu penyakit, kematian atau kejadian peneliti (Gayatri, 2014). Analisis Survival memerlukan data waktu survival. Menurut Collett (2003) Hal yang harus ditentukan waktu survival, yaitu :

1. *starting point* (titik awal) adalah waktu awal dimana pengamatan untuk menganalisis suatu kejadian.

2. *end point* (kejadian akhir) adalah kejadian yang menjadi

- c) *Measurement scale for the passage of the time* (skala ukuran untuk berlalunya waktu) adalah batas dari waktu kejadian dari awal sampai akhir kejadian. Skala diukur dalam hari, minggu, bulan atau tahun

1.6.2 Penyensoran Data

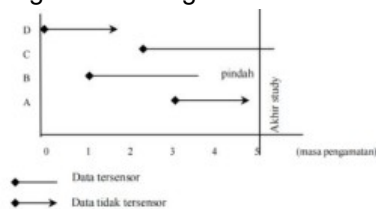
Penyensoran data terjadi jika dalam analisis survival terdapat kesulitan pengamatan data yaitu adanya kemungkinan pengamatan beberapa individu yang tidak dapat di observasi berawal dari *start point* hingga *end point*. Penyensoran data merupakan langkah yang diperlukan untuk menangani ketidaklengkapan dalam suatu data pengamatan. Data dikategorikan sebagai tersensor ketika waktu terjadinya kegagalan tidak dapat diketahui secara pasti hingga batas waktu tertentu. Dalam kondisi ini data yang digunakan adalah selisih antara waktu dimulainya pengamatan hingga waktu terakhir penelitian (Lee & Wang, 2003).

Menurut Kleinbaum & Klein (2012) terdapat tiga alasan terjadinya suatu penyensoran, yaitu:

- Subjek pengamatan yang diamati tidak mengalami suatu *event* sampai pengamatan berakhir (*loss to follow-up*).
- Subjek menghilang selama masa pengamatan, misalnya subjek pindah atau menolak untuk diamati.
- Subjek terpaksa diberhentikan dari penelitian karena meninggal sebelum pengamatan berakhir atau alasan lain.

Menurut Collet (2003) dalam analisis survival terdapat tiga jenis penyensoran, yaitu:

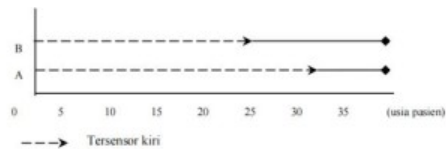
- Sensor Kanan (*Right Censoring*) : Penyensoran ini terjadi ketika objek pengamatan belum mengalami kejadian hingga akhir periode penelitian sementara waktu awal pengamatan dapat diamati secara lengkap. Sebagai contoh seorang pasien kanker yang diamati sejak awal perawatan hingga akhir penelitian masih bertahan hidup. Namun, jika pasien tersebut melanjutkan pengobatan di luar negeri sehingga tidak dapat lagi diikuti perkembangannya (*lost to follow up*) maka waktu kelangsungan hidupnya diketahui setidaknya hingga batas tertentu. Oleh karena itu, waktu pengamatannya dikategorikan sebagai tersensor kanan.



Gambar 1. Ilustrasi Sensor Kanan

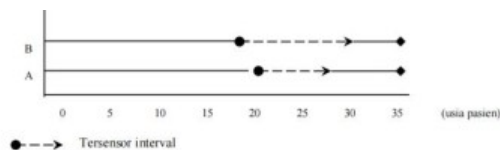
- Sensor Kiri (*Left Censoring*) : Penyensoran ini terjadi ketika waktu awal suatu kejadian tidak dapat diketahui secara pasti pada awal pengamatan tetapi it dapat diamati sebelum penelitian selesai. Sebagai contoh ang pasien kanker seorang individu dinyatakan positif kanker ma. Namun, tidak ada catatan yang menunjukkan kapan kit tersebut mulai berkembang. Dalam kasus ini individu ip mengalami penyensoran kiri.





Gambar 2. Ilustrasi Sensor Kiri

- c) **Sensor Interval (*Interval Censoring*)** : Penyensoran ini terjadi ketika waktu terjadinya suatu kejadian diketahui berada dalam rentang waktu tertentu. Misalnya, dalam catatan medis seorang pasien diketahui dalam kondisi sehat pada usia 45 tahun dan belum terdiagnosis kanker. Namun, saat menjalani tes pertama pada usia 50 tahun, pasien didiagnosis menderita kanker. Dengan demikian waktu terjadinya penyakit tersebut berada di antara usia 45 hingga 50 tahun sehingga dikategorikan sebagai penyensoran interval.



Gambar 3. Ilustrasi Sensor Interval

1.6.3 Konsep Dasar Analisis Survival

Analisis survival terdiri dari tiga konsep dasar yang sering digunakan yaitu, fungsi kepadatan peluang (*probability density function*), fungsi survival (*survival function*), dan fungsi hazard (*hazard function*).

1.6.3.1 Fungsi Kepadatan Peluang

Fungsi kepadatan peluang dinotasikan dengan $f(t)$ yang merupakan peluang individu mengalami *event*. Fungsi ini dirumuskan sebagai berikut:

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T(t + \Delta t))}{\Delta t} \quad (1)$$

Dimana :

T : Variabel acak kontinu yang menyatakan waktu terjadinya *event*

$P(t \leq T(t + \Delta t))$: Peluang *event* dalam interval waktu t sampai $(t + \Delta t)$

1.6.3.2 Fungsi Survival

Fungsi ini dinotasikan dengan $S(t)$ dan didefinisikan sebagai probabilitas suatu objek belum mengalami *event* hingga waktu t dengan $t > 0$. Fungsi survival dihitung dengan cara mengurangkan peluang kejadian yang sudah terjadi dari total peluang. Fungsi ini dirumuskan dengan:

$$\begin{aligned} S(t) &= P(T > t) \\ &= 1 - P(T \leq t) \\ &= 1 - F(t) \end{aligned} \quad (2)$$



uang *event* belum terjadi hingga waktu t
uang *event* telah terjadi sebelum atau pada waktu t
ngsi distribusi kumulatif

rd

h fungsi yang menyatakan besarnya risiko atau laju terjadinya
li antara individu yang masih bertahan hingga waktu tersebut.
azard menunjukkan kemungkinan individu mengalami *event*

dalam interval waktu yang sangat kecil setelah t dengan kondisi bahwa *event* belum terjadi sebelum t . Fungsi ini dinotasikan $h(t)$ dan dirumuskan sebagai:

$$h(t) = \lim_{\Delta t \rightarrow \infty} \left\{ \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \right\} \quad (3)$$

Dimana :

- T : Variabel acak kontinu yang menyatakan waktu terjadinya *event*
 $t \leq T < t + \Delta t$: Individu mengalami *event* pada interval sangat kecil antara waktu t hingga Δt
 $T \geq t$: Individu belum mengalami *event* hingga waktu t

Maka dapat dituliskan :

$$\begin{aligned} h(t) &= \lim_{\Delta t \rightarrow \infty} \left\{ \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \right\} \\ &= \lim_{\Delta t \rightarrow \infty} \left\{ \frac{P(t \leq T < t + \Delta t)}{S(t) \Delta t} \right\} \\ &= \frac{1}{S(t)} \lim_{\Delta t \rightarrow \infty} \left\{ \frac{P(t \leq T < t + \Delta t)}{\Delta t} \right\} \\ &= \frac{1}{S(t)} (f(t)) \\ h(t) &= \frac{f(t)}{S(t)} \end{aligned} \quad (4)$$

Dimana $f(t)$ merupakan *probability density function* (PDF) dari distribusi data survival yang di estimasi dan diketahui bahwa :

$$\int_0^t f(u) du = 1 - S(t) \quad (5)$$

1.6.4 Kurva Survival

Kurva survival atau yang sering dikenal kurva kaplan meier dalam analisis survival digunakan untuk menaksir fungsi survival yaitu peluang individu bertahan hidup atau belum mengalami suatu peristiwa hingga waktu tertentu t (Kleinbaum & Klein, 2012). Kurva kaplan meier menggambarkan hubungan antara estimasi fungsi survival pada setiap titik waktu dengan lamanya waktu ketahanan hidup individu yang diamati. Kurva ini memiliki dua sumbu utama, yaitu sumbu vertikal yang merepresentasikan estimasi fungsi survival dan sumbu horizontal yang menunjukkan waktu survival. Secara teori, kurva survival berbentuk halus (*smooth curves*), namun dalam praktiknya, kurva ini berbentuk fungsi tangga (*step function*) karena perubahan probabilitas survival terjadi pada setiap kejadian (*event*) yang diamati (Suhartini et al., 2018).



a Survival

Analisis distribusi data survival dilakukan untuk mengetahui variabel dependen. Proses ini dilakukan dengan statistik uji. Uji ini dinilai memiliki kekuatan dan tingkat ketepatan yang sama dengan metode uji distribusi lainnya. Seperti yang dikemukakan oleh (4) uji AD berfungsi sebagai uji kenormalan maupun uji (goodness of fit). Uji *Anderson Darling* dapat diterapkan pada berbagai jenis distribusi normal, lognormal, eksponensial, weibull, dan

logistik. Persamaan uji *Anderson Darling* ditunjukkan sebagai berikut (Fa'rifah & Purhadi, 2012) :

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) [\ln F(X_i) + \ln (1 - F(X_{n+1-i}))] \quad (6)$$

Dimana :

F : fungsi distribusi kumulatif distribusi tertentu

X_i : data waktu survival

n : banyaknya data

Menurut Fa'rifah & Purhadi (2012) apabila nilai statistik uji AD yang dihitung lebih besar daripada nilai kritis pada tabel maka hipotesis nol (H_0) ditolak. Sebaliknya, jika nilai AD yang diperoleh lebih kecil dari nilai kritis tabel maka (H_0) diterima. Selain itu bisa juga menggunakan pendekatan jika suatu data dianggap mengikuti distribusi tertentu memiliki nilai AD paling kecil dari nilai distribusi lainnya karena menunjukkan tingkat kesesuaian yang paling baik dengan data.

1.6.6 Regresi Cox Proportional Hazard

Regresi cox adalah salah satu metode pada analisis survival yang bertujuan menganalisa data waktu kejadian untuk mengetahui hubungan waktu kejadian dengan salah satu variabel bebasnya. Metode ini pertama kali diperkenalkan oleh D. R. Cox pada tahun 1972 dengan respon yang digunakan adalah data yang diperoleh dari perhitungan waktu suatu peristiwa yang disebut waktu survival. Misalnya, data waktu pasien menderita penyakit tertentu, dimana perhitungannya dimulai dari pasien didiagnosis menderita penyakit hingga terjadi kejadian khusus (Putu & Ernawatiningsih, 2012).

Regresi Cox termasuk dalam kategori model semiparametrik karena tidak memerlukan asumsi distribusi tertentu untuk waktu survival. Regresi cox memiliki asumsi yang disebut *Proportional Hazard*. Asumsi ini menyatakan bahwa rasio hazard antar dua individu dengan nilai kovariat yang berbeda harus tetap konstan sepanjang waktu. Parameter regresi dalam model ini dapat diestimasi tanpa harus menentukan bentuk spesifik dari fungsi baseline hazard (Lee & Wang, 2003). Model ini menggambarkan hubungan antara variabel independen dan variabel dependen yaitu waktu survival melalui fungsi hazardnya (Soraya et al., 2018). Regresi cox secara umum lebih sering digunakan pada bidang kesehatan, namun seiring berkembangnya waktu, regresi cox dapat diterapkan pada bidang-bidang lain. Secara umum menunjukkan bahwa persamaan *Cox Proportional Hazard* dapat dituliskan sebagai berikut (Kleinbaum, D.G. and Klein, 2012).

Persamaan model *Cox Proportional Hazard* dapat dituliskan sebagai berikut :

Dimana :

$$h(t, x) = h_0(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n) \quad (7)$$

$h(t, x)$: fungsi hazard individu pada waktu t dengan kovariat x

$h_0(t)$: baseline hazard pada waktu t

$X_1, X_2, \dots, X_n$: variabel kovariat/prediktor

$\beta_1, \beta_2, \dots, \beta_n$: koefisien regresi yang menunjukkan pengaruh masing-masing variabel terhadap hazard

Model regresi *Cox Proportional Hazard* harus memenuhi asumsi *Proportional Hazard* yang merupakan salah satu syarat utama agar model valid dan dapat diandalkan. Asumsi ini menyatakan bahwa pengaruh kelompok bersifat konstan terhadap waktu artinya pengaruh



kovariat terhadap risiko kejadian tidak berubah sepanjang waktu pengamatan. Apabila asumsi ini tidak terpenuhi maka model tersebut dikategorikan sebagai nonproportional hazard dan hasil analisis dapat menjadi bias (Grambsch & Therneau, 1994). Menurut Tanjung et al (2024) untuk menguji apakah asumsi proportional hazard terpenuhi salah satu metode yang dapat digunakan metode *Goodness Of Fit* dengan pendekatan *residual schoenfeld*. *Residual schoenfeld* merepresentasikan perbedaan antara kovariat yang diamati pada waktu kejadian dengan nilai kovariat yang diharapkan berdasarkan model. Jika residual ini berkorelasi dengan waktu. Berikut langkah-langkah pengujian asumsi *proportional hazard*: (Sumbernya)

1. Mendapatkan residual schoenfeld dari setiap variabel independen pada model cox PH dan pada setiap objek yang mengalami event.
2. Membuat variabel rank waktu survival yang telah diurutkan berdasarkan waktu survival mulai dari individu yang mengalami event pertama kali.
3. Menguji korelasi antara variabel residual schoenfeld dan rank waktu survival, dengan hipotesis sebagai berikut.
- Hipotesis :
 $H_0 : \rho = 0$ (Tidak ada korelasi antara residual dan waktu, asumsi *proportional hazard* terpenuhi)
 $H_1 : \rho \neq 0$ (Ada korelasi antara residual dan waktu, asumsi *proportional hazard* tidak terpenuhi)
- Statistik uji :

$$t = \frac{r_{RT, PR_i} \sqrt{n-2}}{\sqrt{1 - r_{RT, PR_i}^2}} \quad (8)$$

- Kriteria Keputusan :
Terima H_0 jika $|t_{hit}| > t_{\frac{\alpha}{2}, n-2}$
Atau $pvalue > 0,05$ (maka Asumsi *proportional hazard* terpenuhi)

1.6.7 Hazard Ratio

Hazard ratio merupakan inti dari interpretasi dalam regresi cox yang secara matematis didefinisikan sebagai perbandingan antara fungsi hazard dari dua individu dengan kovariat berbeda. Secara matematis hazard ratio dapat dinyatakan sebagai berikut (Kleinbaum & Klein, 2012) :

$$HR = \frac{h(t|X_i)}{h(t|X_j)} = \frac{h_0(t) \cdot e^{\beta_1 X_{i1} + \dots + \beta_p X_{ip}}}{h_0(t) \cdot e^{\beta_1 X_{j1} + \dots + \beta_p X_{jp}}} \quad (9)$$

Karena fungsi baseline hazard $h_0(t)$ sama untuk kedua individu maka persamaan (9) akan menghasilkan :

$$HR = e^{\beta_1 (X_{i1} - X_{j1}) + \dots + \beta_p (X_{ip} - X_{jp})} \quad (10)$$



lalu HR bergantung pada jenis skala pengukuran variabel unakan (Gayatri, 2014). Untuk variabel dikotomi (biner) hazard r risiko mengalami event antara dua kelompok kategori yang bila suatu variabel dikodekan sebagai 1 (kelompok eksperimen risiko) dan 0 (kelompok referensi atau tanpa faktor risiko) maka

nilai $HR = \exp(\beta)$ diinterpretasikan sebagai besarnya risiko relatif individu pada kelompok eksperimen terhadap kelompok referensi. Interpretasi hazard ratio adalah sebagai berikut :

- $HR > 1$: individu dengan nilai 1 memiliki risiko kematian lebih tinggi dibandingkan nilai 0.
- $HR < 1$: individu dengan nilai 1 memiliki risiko kematian lebih rendah (efek protektif) dibandingkan nilai 0.
- $HR = 1$: tidak ada perbedaan risiko antara kelompok 1 dan 0.

Sementara itu, untuk variabel kontinu hazard ratio mengukur proporsi perubahan risiko event untuk setiap kenaikan satu unit pada variabel prediktor. Interpretasinya adalah besarnya perubahan relatif risiko yang terkait dengan setiap pertambahan satu satuan dalam variabel kontinu. Interpretasi hazard ratio untuk variabel kontinu adalah sebagai berikut :

- $HR > 1$: setiap kenaikan 1 satuan pada variabel risiko kematian meningkat sebesar $(HR - 1) \times 100\%$.
- $HR < 1$: setiap kenaikan 1 satuan pada variabel risiko kematian menurun sebesar $(1 - HR) \times 100\%$.
- $HR = 1$: tidak ada perubahan risiko seiring kenaikan variabel.

1.6.8 Residual Martingale

Model yang telah disesuaikan perlu dilakukan penilaian terhadap kecukupan model tersebut. Prosedur diagnostik untuk pemeriksaan model merupakan bagian penting dalam proses pemodelan, dan banyak prosedur tersebut didasarkan pada analisis residual. Dalam analisis survival, khususnya dalam membangun model *Cox Proportional Hazards* berbagai jenis residual dapat digunakan untuk berbagai tujuan. Salah satu jenis residual yang penting adalah residual martingale yang berfungsi untuk memeriksa kesesuaian model secara keseluruhan pada model regresi cox (Fitrianto & Jiin, 2013). Persamaan residual dirumuskan sebagai berikut:

$$M_i(t) = \delta_i - \hat{H}_i(t) \quad (11)$$

Dimana :

$M_i(t)$: Residual martingale subyek pengamatan ke-i

δ_i : 0, untuk data tersensor
1, untuk data tidak tersensor

$\hat{H}_i(t)$: Hazard Kumulatif subyek pengamatan ke-i pada waktu ke-t

Residual martingale dapat menjadi gambaran mengenai perbedaan hasil pengamatan δ_i dengan angka prediksi pada kejadian $\hat{H}_i(t)$. Ketika perbedaan antara hasil pengamatan dengan angka prediksi untuk subjek ke-i cukup besar berarti menunjukkan bahwa subyek ke-i tidak sesuai dengan model dan mengakibatkan suatu nilai yang besar pada $\hat{M}_i(t)$. Karena interval nilai dari $\hat{H}_i(t)$ adalah $(0, \infty)$ dan δ_i hanya bernilai 0 atau 1 maka dapat diambil kesimpulan bahwa $M_i(t)$ bernilai $(-\infty, 1)$ dan bernilai negatif pada data tersensor).

Residual martingale memiliki sifat-sifat berikut (Therneau &

$$E(M_i) = 0 \text{ dan } Cov(M_i, M_j) = 0 \quad (12)$$

dalam model yang sesuai nilai harapan residual martingale pengamatan adalah nol dan tidak berkorelasi satu sama lain. Oleh



karena itu, apabila hubungan antara kovariat dan log hazard bersifat linier maka nilai residual martingale akan berfluktuasi acak di sekitar nol tanpa menunjukkan pola tertentu. Sebaliknya, apabila terdapat pola sistematis (misalnya melengkung) hal tersebut mengindikasikan adanya hubungan nonlinier antara kovariat dan log hazard yang dapat diperbaiki dengan pendekatan nonparametrik seperti Multivariate Adaptive Regression Splines (MARS).

1.6.9 Multivariate Adaptive Regression Spline

MARS (*Multivariate Adaptive Regression Splines*) adalah metode regresi nonparametrik yang bertujuan untuk memprediksi nilai variabel respons berdasarkan sekumpulan variabel prediktor tanpa memerlukan asumsi mengenai hubungan fungsional dasar antara keduanya. Model ini dirancang untuk menangani data berdimensi tinggi, yaitu data dengan lebih dari tiga variabel prediktor (Hastie et al., 2008). Keunggulan MARS terletak pada kemampuannya dalam menghasilkan prediksi variabel respons yang akurat serta membangun model yang kontinu pada titik-titik perubahan (knot) berdasarkan nilai *Generalized Cross Validation* (GCV) terkecil (Friedman, 1991). GCV digunakan untuk menentukan lokasi knot yang optimal.

Menurut Nash & Bradford (2001) ada dua aspek penting dalam pemodelan MARS yaitu pemilihan knot dan basis fungsi. Jika satu garis regresi tidak cukup untuk menggambarkan pola data secara keseluruhan, beberapa garis regresi akan digunakan untuk menjelaskan hubungan antara variabel independen dan respons. Titik perubahan antara satu garis regresi ke garis lainnya disebut knot, yang menandai batas akhir dari satu region dan awal dari region lainnya. Kontinuitas basis fungsi pada setiap knot menjadi faktor penting dalam membangun model yang baik.

Friedman (1991) mengatakan bahwa ada beberapa hal yang harus diperhatikan dalam pemodelan menggunakan metode MARS, yaitu:

1. Basis fungsi (BF) menyatakan jumlah maksimum fungsi basis yang digunakan. Disarankan 2–4 kali jumlah variabel prediktor.
2. Maksimum interaksi (MI) merupakan tingkat interaksi maksimum antarfungsi basis. Nilai MI = 1 (tanpa interaksi), MI = 2 (interaksi dua variabel), MI = 3 (interaksi tiga variabel). Nilai MI > 3 cenderung menyebabkan *overfitting*.
3. Minimum observasi (MO) merupakan jumlah pengamatan paling minimal antar knot. Jumlah maksimum interaksi (MI) adalah 1, 2 dan 3 karena diatas itu nilai GCV akan semakin meningkat.

Penentuan basis fungsi (BF), maksimum interaksi, dan minimum observasi akan berpengaruh terhadap model MARS yang akan terbentuk. Menurut Friedman (1991) persamaan umum model *Multivariate Adaptive Regression Splines* (MARS) dapat dituliskan sebagai berikut:

$$\hat{f}(x) = a_0 + \sum_{m=1}^M \alpha_m \prod_{k=1}^{K_m} [s_{km}(x_{v(k,m)} - \tau_{km})] \quad (13)$$



ita dari fungsi basis BF_0
 igsi basis BF_m
 1 fungsi basis (*non constant* fungsi basis)

$v(k, m)$: label dari variabel prediktor

τ_{km} : nilai knots dari variabel prediktor $x_{v(k,m)}$

Model MARS dari persamaan (13) dapat juga dituliskan sebagai berikut :

$$f(x) = a_0 + \sum_{m=1}^M \alpha_m \prod_{k=1}^{K_m} [s_{km}(x_{v(k,m)} - \tau_{km})] + \varepsilon \quad (14)$$

Friedman (1991) mengatakan bahwa pemilihan model terbaik dalam MARS didasarkan pada nilai GCV terkecil dibandingkan model lainnya. Model MARS dipilih menggunakan metode stepwise di mana tahap forward stepwise dilakukan untuk memperoleh jumlah basis fungsi maksimum dengan mempertimbangkan minimisasi *Average Sum Square Residual* (ASR). Selanjutnya, untuk memastikan model tetap sederhana sesuai dengan prinsip parsimoni dilakukan proses pemilihan model dilakukan dengan *trial and error* pada berbagai kombinasi parameter BF, MI dan MO. Setiap kombinasi dievaluasi dengan melihat nilai GCV di mana model terbaik dipilih berdasarkan nilai GCV terkecil. Nilai GCV sendiri didefinisikan oleh Friedman (1991) sebagai berikut.

$$GCV(M) = \frac{\frac{1}{n} \sum_{i=1}^n [y_i - f_M(x_i)]^2}{\left[1 - \frac{C(M)}{n}\right]^2} \quad (15)$$

Dimana $C(M) = \text{Trace}(B(B^T B)^{-1} B^T) + 1$

1.6.10 Kematian Bayi

Pembangunan berkelanjutan global melalui agenda *Sustainable Development Goals* (SDGs) menetapkan bahwa pada tahun 2030 setiap negara diharapkan mampu mengakhiri kematian yang dapat dicegah pada bayi baru lahir dan balita. Target spesifiknya adalah mencakup pengurangan angka kematian dalam 5 tahun pertama menjadi 25 kematian per 1000 kelahiran hidup dan 28 hari pertama menjadi 12 kematian per 1000 kelahiran hidup pada tahun 2030 (SDGs, 2019). Hampir 5 juta anak di seluruh dunia meninggal sebelum ulang tahun kelima mereka pada tahun 2021 dengan 27% tinggal di Asia. Angka kematian bayi tidak hanya mencerminkan kesehatan bayi yang baru lahir tetapi juga kesejahteraan umum masyarakat. Angka kematian pasca neonatal adalah jumlah kematian bayi berusia 28 hari hingga 330 hari dan dinyatakan sebagai kematian pasca neonatal per 1000 kelahiran hidup dalam setahun (Subramanian et al., 2024).

Keberhasilan Indonesia dalam menurunkan angka kematian anak sangat penting untuk mencapai target global SDGs. Untuk menurunkan risiko kematian pada anak usia di bawah 5 tahun data perlu dibedakan berdasarkan kelompok usia. Namun, kematian bayi pada usia 0–7 hari (neonatal dini) dan 8–28 hari (neonatal



abungkan dalam pelaporan. Begitu pula, kematian pada usia neonatal) dan 1–4 tahun (anak usia dini) juga sering disatukan (2024). Penggabungan ini menjadi masalah karena penyebab berbeda pada setiap kelompok usia sehingga membutuhkan berbeda pula. Misalnya, pada bayi baru lahir berusia 0–7 hari kematian disebabkan oleh kelahiran prematur atau gangguan saat penelitian Fentaw et al., (2022) kasus kematian bayi yang

berusia 29 hari - 330 hari biasanya terjadi karena beberapa faktor seperti interval kelahiran sebelumnya, jenis kelamin anak, usia ibu saat lahir, status kerja, pendidikan orang tua, status menyusui, usia ibu, kekayaan rumah tangga, jumlah anggota keluarga.

Kematian perinatal dapat dipengaruhi oleh dua kelompok faktor utama, yaitu faktor ibu (*High Risk Mother*) dan faktor bayi (*High Risk Infant*). Berikut ini adalah faktor dari ibu yang dapat meningkatkan kemungkinan kematian bayi (Mahmudah et al., 2015) :

a. Usia Ibu

Usia ibu merupakan salah satu faktor yang berpengaruh terhadap risiko kematian bayi. Ibu yang berusia < 20 tahun atau > 35 tahun memiliki risiko lebih tinggi terhadap kematian bayi akibat ketidakmatangan fisik, emosional, atau meningkatnya risiko komplikasi kehamilan dan persalinan. Namun jika usia yang berada pada rentang usia 20–35 tahun umumnya berada dalam kondisi reproduksi yang optimal, sehingga memiliki risiko kematian bayi yang lebih rendah dibandingkan kelompok usia di luar rentang tersebut.

b. Pendidikan Ibu

Ibu dengan tingkat pendidikan rendah (baik formal maupun informal) seringkali mengalami keterbatasan dalam memahami informasi kesehatan bayi termasuk pentingnya imunisasi lengkap, praktik gizi yang benar, serta tanda bahaya pada bayi. Hal ini dapat menyebabkan keterlambatan penanganan saat bayi sakit, yang merupakan penyebab utama kematian pasca-neonatal.

c. Kunjungan Antenatal (ANC)

Pemeriksaan kehamilan secara rutin (ANC) tidak hanya berperan dalam mendeteksi risiko kehamilan tetapi juga menjadi sarana edukasi bagi ibu tentang pentingnya perawatan bayi setelah lahir. Ibu yang rutin melakukan ANC cenderung lebih siap dalam memberikan ASI eksklusif, menjaga kebersihan, dan melakukan imunisasi yang sangat penting untuk mencegah kematian pasca-neonatal.

d. Paritas (Jumlah Kelahiran Sebelumnya)

Ibu dengan jumlah anak yang banyak (misalnya >4 anak) akan mengalami kelelahan fisik dan keterbatasan waktu serta perhatian dalam mengasuh bayi baru. Pengasuhan yang tidak optimal dapat menyebabkan bayi lebih rentan terhadap infeksi, kekurangan gizi, dan keterlambatan penanganan penyakit yang dapat menyebabkan kematian setelah masa neonatal.

e. Jarak Kehamilan

Jarak kehamilan yang terlalu dekat (<2 tahun) dapat menyebabkan ibu belum pulih secara fisik dan mental dari kehamilan sebelumnya. Hal ini mengganggu kemampuan ibu dalam merawat bayi secara maksimal termasuk dalam pemberian ASI, perhatian terhadap tumbuh kembang, dan deteksi dini penyakit. Akibatnya, bayi lebih berisiko mengalami gangguan kesehatan serius di masa pasca-neonatal.



Bayi itu sendiri juga mempunyai peran penting dalam kematian post neonatal. Kondisi biologis dan fisiologis bayi yang lahir dapat meningkatkan kerentanan terhadap berbagai gangguan kesehatan. Hal ini disebabkan oleh beberapa faktor utama dari bayi yang dapat meningkatkan risiko kematian (Mahmudah et al., 2015):

Ibu yang melahirkan cenderung lebih rentan terhadap berbagai gangguan kesehatan bayi perempuan, termasuk pada periode pasca neonatal. Hal ini

berkaitan dengan faktor biologis di mana bayi perempuan memiliki dua kromosom X (XX) yang memungkinkan kompensasi jika salah satu kromosom mengalami kerusakan. Bayi laki-laki hanya memiliki satu kromosom X dan satu Y (XY) sehingga tidak memiliki cadangan. Kerentanan ini membuat bayi laki-laki lebih berisiko mengalami penyakit infeksi dan komplikasi jangka panjang yang dapat berdampak hingga usia pasca neonatal.

b. Sepsis Neonatal

Meski sepsis biasanya terjadi di awal kehidupan, dampaknya bisa berlanjut hingga pasca neonatal. Bayi yang selamat dari sepsis bisa mengalami gangguan organ atau penurunan imunitas yang menyebabkan rentan terhadap infeksi berulang. Sepsis juga lebih banyak terjadi pada bayi dengan jenis kelamin laki-laki, bayi kembar, dan bayi dengan kelahiran prematur atau berat badan lahir rendah. Komplikasi seperti infeksi sistemik dapat berdampak jangka panjang jika tidak ditangani optimal.

c. Berat Badan Lahir Rendah (BBLR)

BBLR adalah kondisi dimana bayi lahir dengan berat badan kurang dari 2.500 gram terlepas dari usia kehamilan. BBLR tidak hanya berisiko pada masa neonatal, tetapi juga berdampak besar pada periode pasca neonatal. Bayi dengan BBLR memiliki organ yang belum matang sepenuhnya termasuk sistem imun dan pernapasan, sehingga rentan terhadap pneumonia, diare, dan gagal tumbuh. Kondisi ini meningkatkan risiko infeksi berulang, rawat inap, hingga kematian setelah periode neonatal.

d. Asfiksia Neonatorum

Bayi yang mengalami asfiksia saat lahir dapat mengalami gangguan neurologis jangka panjang seperti kejang, keterlambatan tumbuh kembang, atau kerusakan otak permanen. Komplikasi ini tidak selalu menyebabkan kematian segera, tetapi dapat meningkatkan kerentanan terhadap infeksi dan komplikasi sistemik pada masa pasca neonatal yang pada akhirnya meningkatkan risiko kematian bayi setelah 28 hari kehidupan.

e. Kelainan Kongenital

Kelainan kongenital terutama yang tidak terdeteksi atau tidak tertangani sejak dini dapat menimbulkan komplikasi berat setelah masa neonatal. Misalnya, kelainan jantung bawaan, kelainan saluran pencernaan, atau sistem saraf pusat bisa menjadi penyebab gagal tumbuh, gangguan fungsi organ, dan kematian pasca neonatal. Deteksi dini melalui pemeriksaan antenatal seperti USG dan skrining metabolik sangat penting untuk mengurangi risiko kematian ini.

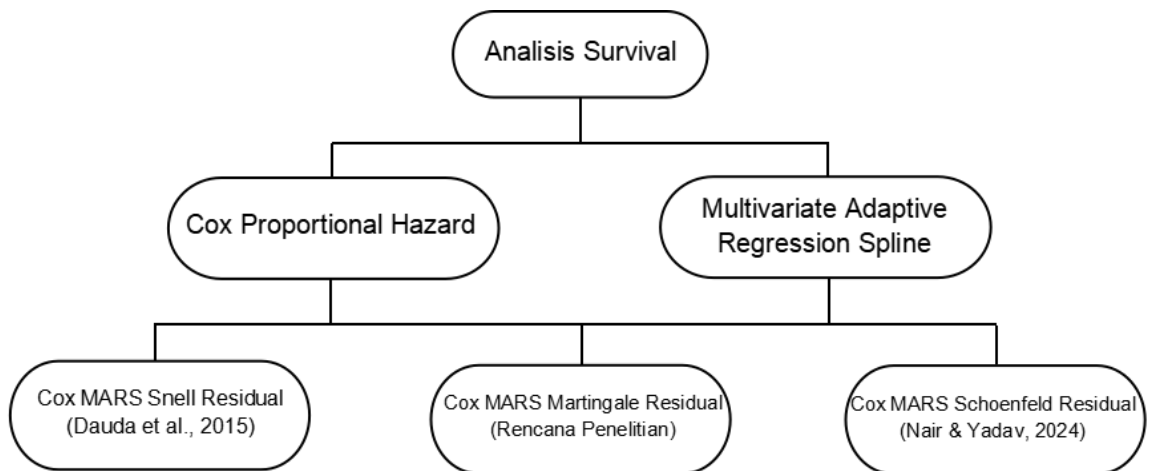
f. Status Imunisasi

Imunisasi dasar merupakan imunisasi awal yang diberikan kepada bayi sebelum berusia satu tahun. Pada kondisi ini, diharapkan sistem kekebalan tubuh dapat bekerja secara optimal. Setiap bayi (usia 0- 330 hari) diwajibkan untuk mendapatkan imunisasi dasar lengkap yang terdiri dari 1 dosis Hepatitis



3 dosis DPT-HBiB, 4 dosis polio tetes, dan 1 dosis campak imunisasi sangat berpengaruh terhadap kelangsungan hidup a pasca neonatal. Bayi yang tidak mendapatkan imunisasi lebih tinggi terkena penyakit infeksi serius seperti pneumonia, batuk rejan) dan campak yang merupakan penyebab utama negara berkembang. Imunisasi dasar lengkap sesuai jadwal banyak kematian pasca neonatal menjadikannya intervensi ya menurunkan angka kematian bayi di Indonesia

1.6.11 Kerangka Konseptual



Gambar 4. Kerangka Konseptual



BAB II METODE PENELITIAN

2.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder dari rekam medik kematian bayi pasca neonatal di RS Labuang Baji Makassar tahun 2021-2024. *Failure event* terjadi apabila bayi meninggal dalam rentang waktu 29 hari sampai 330 hari.

2.2 Variabel Penelitian

Tabel 1. Variabel Penelitian

Variabel Penelitian	Definisi Operasional	Tipe	Kategori
Waktu Bertahan Hidup (Y)	Waktu bertahan hidup bayi [hari]	Kontinu	
Jenis Kelamin(X_1)	Jenis kelamin bayi	Kategorik	0. Laki-laki 1. Perempuan
Berat Badan Lahir (X_2)	Bayi lahir dengan Berat Badan Lahir Rendah [gram]	Kontinu	
Kelainan Kongenital (X_3)	Bayi lahir dengan kelainan bawaan (kongenital)	Kategorik	0. Tidak 1. Iya
Status Imunisasi(X_4)	Status kelengkapan imunisasi dasar bayi sesuai usia	Kategorik	0. Lengkap 1. Tidak lengkap
Usia Ibu (X_5)	Usia ibu saat melahirkan bayi	Kontinu	
Jarak Kehamilan (X_6)	Jarak antara kelahiran sebelumnya dengan kelahiran bayi saat ini	Kategorik	0. < 2 Tahun 1. $\geq$ 2 Tahun

2.3 Metode Analisis Data

Adapun langkah-langkah yang dilakukan berdasarkan tujuan penelitian adalah sebagai berikut

1. Eksplorasi Awal Data

Melakukan eksplorasi deskriptif terhadap data kematian bayi pasca neonatal di RS Labuang Baji Makassar.

2. Analisis Kurva Survival

Menyusun kurva kaplan meier untuk memperoleh gambaran peluang kelangsungan hidup untuk semua variabel.

3. Analisis Distribusi Data Survival

Penggunaan distribusi data survival menggunakan uji AD untuk distribusi yang paling sesuai. Distribusi terbaik berdasarkan nilai statistik AD terkecil.



4. Pemodelan Survival (Model Cox Proportional Hazard)
 - a) Mengestimasi parameter model cox menggunakan metode Partial Maximum Likelihood Estimation (MLE) untuk mendapatkan nilai hazard ratio dan koefisien regresi.
 - b) Membangun model cox proportional hazard
 - c) Mengestimasi fungsi survival dan fungsi hazard
 - d) Melakukan uji asumsi cox proportional hazard menggunakan schoenfeld residual.
5. Analisis Residual Martingale
 - a) Menghitung residual Martingale dari model Cox PH untuk setiap individu, dan menggunakan residual tersebut sebagai variabel respon pada pemodelan lanjutan untuk mendeteksi ketidaksesuaian bentuk hubungan linier pada Cox PH.
 - b) Membuat plot martingale residual terhadap setiap prediktor untuk mendeteksi indikasi hubungan non-linier atau interaksi antar variabel.
6. Pemodelan dengan *Multivariate Adaptive Regression Splines* (MARS)
 - a) Melakukan estimasi parameter model Multivariate Adaptive Regression Splines (MARS) menggunakan metode Ordinary Least Squares (OLS) untuk menentukan basis functions dan koefisiennya.
 - b) Membangun model MARS dengan martingale residual sebagai variabel respon dan prediktor dari data survival sebagai variabel penjelas
 - c) Menentukan kombinasi terbaik parameter MARS (BF, MI dan MO)
 - d) Memilih model terbaik berdasarkan nilai minimum dari *Generalized Cross Validation* (GCV).
7. Pemodelan Survival dengan Pendekatan MARS
 Membuat dan menginterpretasi model MARS untuk mengidentifikasi bentuk hubungan antar prediktor dengan risiko kematian bayi pasca neonatal serta melihat potensi ketidaklinieran hubungan yang tidak mampu ditangkap oleh model Cox PH.

