

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Angka Kematian Bayi (AKB) adalah salah satu indikator penting yang digunakan untuk mengukur kualitas kesehatan di suatu wilayah. AKB mencerminkan tingkat keberhasilan pelayanan kesehatan, terutama terkait dengan kesehatan ibu dan bayi. Provinsi Sulawesi Selatan, masih menghadapi tantangan besar dalam menurunkan AKB, yang menjadi perhatian utama bagi pemerintah dan para pemangku kebijakan kesehatan. Dalam upaya memahami faktor-faktor yang mempengaruhi AKB, pendekatan statistika sering digunakan untuk memodelkan hubungan antara variabel respon, yaitu AKB, dengan variabel-variabel penjelas seperti akses kesehatan, tingkat pendidikan ibu, atau status ekonomi. Salah satu pendekatan yang umum digunakan adalah pemodelan regresi, di mana hubungan antara variabel-variabel ini dapat dianalisis secara mendalam. Analisis regresi adalah salah satu pemodelan dalam statistika yang digunakan untuk mengetahui hubungan antara satu variabel respon (terikat) dengan satu atau lebih variabel penjelas (bebas). Asumsi mendasar untuk analisis regresi klasik adalah memiliki variabel respon kontinu yang berdistribusi normal (Tiro,dkk 2010). Asumsi ini sulit dipertahankan, berhubung banyaknya skala data yang berupa cacahan (*count data*) yang umumnya tidak menyebar normal. Data AKB umumnya berupa data cacahan (*count data*) yaitu data yang menggambarkan jumlah kejadian sehingga pendekatan regresi linier klasik kurang cocok karena asumsi distribusi normalnya tidak terpenuhi. Oleh karena itu, model yang lebih tepat digunakan untuk data cacahan adalah regresi Poisson, yang secara khusus dirancang untuk memodelkan suatu data.

Regresi Poisson mengasumsikan bahwa variansi data sama dengan nilai rata-ratanya, yang dikenal sebagai *equidispersi*. Regresi poisson merupakan salah satu alternatif yang baik pada variabel respon diskrit (Amin, dkk 2020; Agresti 2015). Namun, dalam praktik, asumsi ini jarang terpenuhi karena sering dijumpai kasus *overdispersi* atau *underdispersi*. Data AKB sering menunjukkan *overdispersi*, yaitu kondisi di mana variansi lebih besar daripada rata-rata. Mengatasi masalah *overdispersi* ini, model alternatif seperti Poisson Tweedie dapat digunakan. Regresi poisson mengalami pengembangan melalui distribusi Tweedie dan menghasilkan model regresi Poisson Tweedie yang mampu membantu menangani kejadian *over-under disperse* tersebut (Abonazel et al., 2024). Model Poisson Tweedie menawarkan fleksibilitas dalam memodelkan data cacahan dengan distribusi yang lebih bervariasi, memungkinkan penanganan *overdispersi* lebih baik dibandingkan regresi Poisson.

Regresi Poisson Tweedie merupakan kelompok dari analisis regresi, sehingga dapat juga digunakan untuk memodelkan hubungan sebab akibat antara variabel independen dengan variabel dependen data diskrit (Aswi, dkk 2022; Sundari and Sihombing 2021; Winata 2023). Angka Kematian Bayi (AKB) merupakan jumlah

kematian bayi usia di bawah 1 tahun (0-11 bulan) per 1000 kelahiran hidup, sehingga merupakan data diskrit (Oktaviana, 2024; Pemerintah Provinsi Sulawesi Selatan, 2023).

Data AKB memungkinkan terjadinya pelanggaran asumsi *equidispersi*. Beberapa peneliti telah mencoba memberikan alternatif solusi untuk mengatasi masalah tidak dipenuhinya *equidispersi* ini, seperti menggunakan regresi binomial negatif (Fitri et al., 2021); penanganan *overdispersi* Regresi Zero Inflated Poisson (ZIP) (Putri et al., 2022); *generalized Poisson regression* (Arisandi et al., 2019); regresi Poisson Invers Gaussian (PIG) (Damayanti CR & Yanti, 2022). Keluarga Poisson Tweedie dapat beradaptasi dengan data yang memiliki nol berlebih tanpa perlu menggunakan regresi ZIP dan sangat berekor ke kanan (Bonat et al., 2017). Oleh karena itu, Regresi Poisson Tweedie (RPT) ini dapat dijadikan salah satu solusi alternatif dalam memodelkan data count seperti AKB.

Beberapa metode yang telah digunakan untuk data kematian bayi adalah regresi Conway-Maxwell Poisson (Oktaviana, 2024), regresi Poisson Bayes berhierarki (Hajarisman, 2013) pada jumlah kematian bayi, analisis jalur (Sanusi, dkk 2020; Yasril, dkk 2022) pada kesehatan balita, dan regresi nonparametrik (Sanusi, dkk 2020) pada AKB yang menemukan bahwa Berat Badan Lahir Rendah (BBLR) mempengaruhi AKB.

Keunggulan Poisson Tweedie (PTw) telah dikaji beberapa peneliti sebelumnya dan diperoleh bahwa PTw mampu mengatasi permasalahan *under* maupun *overdispersi* (Vinogradov and Paris 2017; Nasution, dkk 2022; Bonat et al. 2017), PTw memberikan kinerja lebih baik pada kasus kecelakaan lalu lintas yang memiliki data *overdispersi* (Debrabant et al., 2018; Saha et al., 2020). PTw memiliki kelebihan mengatasi variasi dispersi, fleksibilitas, dan mengatasi masalah non-homogenitas (Novani et al., 2014), serta mampu memprediksi kasus asuransi (Putra et al., 2021; Rahmawati et al., 2019). PTw pada bidang asuransi dan keuangan untuk memprediksi kerugian agregat dan tingkat keparahan klaim dalam asuransi dan keuangan (Putra et al., 2021), di bidang biostatistik untuk memprediksi kasus biostatistik, seperti kejadian dan prevalensi penyakit (Nasution et al., 2022). Lalu, penelitian ini akan mencoba mengevaluasi kinerja PTw pada kasus AKB dengan memperhatikan nilai AIC terkecil. Penelitian menggunakan Regresi PTw telah banyak dilakukan, namun masih belum menyentuh faktor signifikan yang mempengaruhi AKB dengan menggunakan Maximum Likelihood Estimation (MLE) dengan pendekatan Algoritma Fisher scoring. Kesesuaian dan kecocokan model dengan keunggulan-keunggulan regresi PTw ini diharapkan dapat memberikan informasi yang signifikan mengenai faktor-faktor yang mempengaruhi AKB, yang selanjutnya dapat digunakan untuk mendukung rekomendasi kebijakan kesehatan yang lebih efektif.

Berdasarkan hasil eksplorasi data diperoleh bahwa data AKB di Sulawesi Selatan memperlihatkan adanya *overdispersi* sehingga menggugurkan kemungkinan menggunakan regresi klasik ataupun regresi Poisson klasik itu sendiri. Selain itu tidak ditemukan adanya multikolinearitas. Kondisi AKB di Sulawesi Selatan memberikan peluang besar untuk menggunakan regresi PTw sebagai solusi

alternatif dalam memodelkan kasus AKB di Sulawesi Selatan. Penelitian ini bertujuan untuk mendeskripsikan kondisi AKB di Sulawesi Selatan, membangun model PTw, menentukan faktor berpengaruh terhadap AKB di Sulawesi Selatan dan menentukan model terbaik. Untuk mencapai tujuan tersebut dilakukan analisis deskriptif dan inferensial melalui model regresi PTw.

## 1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang, masalah yang dikaji dalam penelitian ini adalah:

1. Bagaimana formulasi model Regresi Poisson dan Regresi Poisson Tweedie pada data AKB di Sulawesi Selatan?
2. Mengidentifikasi model terbaik dari model Regresi Poisson dan Regresi Poisson Tweedie berdasarkan data AKB di Sulawesi Selatan menggunakan AIC.

## 1.3 Batasan masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Penaksiran parameter model regresi Poisson Tweedie menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan algoritma Fisher Scoring.
2. Studi kasus yang digunakan yaitu data AKB di Sulawesi Selatan tahun 2020-2023.

## 1.4 Tujuan Penelitian

Adapun tujuan penelitian ini yaitu:

1. Memperoleh model Regresi Poisson dan Poisson Tweedie pada data AKB di Sulawesi Selatan.
2. Memperoleh model terbaik pada model Poisson dan Poisson Tweedie pada data AKB di Sulawesi Selatan.

## 1.5 Manfaat Penelitian

1. **Peningkatan Pemahaman:** Dengan model PTw, penelitian ini akan meningkatkan pemahaman tentang faktor-faktor yang mempengaruhi AKB. Hal ini akan membantu dalam mengidentifikasi faktor risiko yang paling berpengaruh dan mengarah pada strategi intervensi yang lebih efektif.
2. **Kontribusi Terhadap Literatur Ilmiah:** Penelitian ini diharapkan dapat memberikan kontribusi kepada literatur ilmiah tentang metode pendekatan statistik dalam analisis AKB, sehingga dapat digunakan sebagai referensi oleh peneliti dan praktisi kesehatan masyarakat di masa depan.
3. **Bagi Peneliti:** Menambah wawasan dalam menganalisis apakah regresi PTw dapat digunakan untuk memodelkan kasus AKB di Sulawesi Selatan.

## 1.6 Teori

Sub-bab teori ini membahas mengenai beberapa teori yang digunakan untuk membahas pokok pembahasan terkait regresi *Poisson* dan *Poisson Tweedie*

### 1.6.1 Regresi Poisson

Regresi Poisson digunakan untuk menganalisis suatu data jumlahan (*count*) misalnya banyaknya kejadian yang terjadi dalam kurun waktu dan atau wilayah tertentu. Regresi Poisson merupakan regresi yang digunakan untuk menganalisis data yang bersifat diskrit. Regresi Poisson standar mengasumsikan bahwa nilai variansi sama dengan nilai rata-ratanya (*equidispersi*). Namun, kasus *equidispersi* jarang terjadi pada kasus nyata. Kejadian *overdispersi* dan *underdispersi* lebih sering terjadi dibandingkan *equidispersi* (Sellers & Premeaux, 2020). Model regresi poisson termasuk dalam model regresi non linier dan merupakan model standar untuk data diskrit (Colin & Pravin, 2013).

Model regresi Poisson adalah:

$$\begin{aligned}\ln \mu_i &= \eta_i \\ \ln \mu_i &= \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} \\ \mu_i &= \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})\end{aligned}\quad (1)$$

Dengan fungsi rata-rata:

$$(\mu_i) = \exp(X\beta),$$

dengan:

$\mu_i$  : rata-rata(mean) dari respons  $Y_i$ ,

$X_i$  :  $(1, X_{i1}, X_{i2}, \dots, X_{ip})$  adalah vektor prediktor untuk pengamatan ke-  $i$ ,

$\beta$  :  $(\beta_0, \beta_1, \dots, \beta_p)$  adalah vektor parameter regresi.

### 1.6.2 Distribusi Poisson

Distribusi Poisson atau Sebaran poisson merupakan sebaran diskrit. Apabila nilai  $\mu$  (mean) yang besar akan lebih mendekati distribusi normal. Sementara kasus yang jarang terjadi nilai  $\mu$  akan kecil. Distribusi Poisson adalah suatu distribusi yang paling sederhana dalam memodelkan data cacah. Distribusi Poisson adalah distribusi probabilitas untuk variabel acak yang menghitung jumlah kejadian yang terjadi dalam interval waktu atau ruang tertentu, dengan asumsi bahwa peristiwa terjadi secara acak dan independen dengan laju rata-rata yang tetap ( $\lambda$ ). Misalnya, jumlah kejadian kecelakaan lalu lintas dalam satu hari kematian bayi dan ibu dapat dimodelkan dengan distribusi Poisson. Bilangan  $Y$  menyatakan banyaknya hasil percobaan dalam suatu percobaan Poisson disebut variabel acak Poisson adapun sebaran peluangnya disebut sebaran Poisson. Sebaran peluang bagi variabel acak  $Y$ , menyatakan banyaknya hasil percobaan yang terjadi selama suatu selang waktu tertentu. (Walpole, 1995) menyatakan, sebaran Poisson memiliki fungsi peluang sebagai berikut:

$$P(Y = y) = \frac{e^{-\mu} \mu^y}{y!} \quad (2)$$

dengan:

$y$  : 0,1,2, ....

$\mu$  : rata-rata banyaknya hasil percobaan yang terjadi dalam waktu tertentu

$e$  : 2, 7182...

Sebaran Poisson dapat digunakan untuk menghampiri sebaran binomial apabila  $n$  besar dan peluang sukses  $p$  kecil atau dekat dengan nol (Walpole, 1995).

### 1.6.3 Poisson Tweedie

Poisson Tweedie adalah sebuah model statistik yang digunakan untuk memodelkan data cacah yang memiliki dispersi yang lebih besar daripada yang diharapkan oleh model Poisson biasa. Model ini dikembangkan oleh Alan G. Tweedie dan diperkenalkan dalam penelitian tahun 1984. Model Poisson Tweedie memungkinkan penggunaan parameter yang lebih fleksibel dalam mengatasi dispersi yang tidak terdistribusi secara normal, seperti yang sering terjadi dalam data cacah. Model ini mengasumsikan bahwa data mengikuti distribusi Tweedie, yang merupakan keluarga distribusi eksponensial yang mencakup distribusi Poisson dan Gamma sebagai kasus khusus. Syarat penggunaan model Poisson Tweedie meliputi beberapa aspek penting yang perlu dipertimbangkan (Hugo Bonat et al., 2017):

1. **Karakteristik Data:** Model Poisson Tweedie cocok digunakan untuk data yang bersifat count (jumlah) dan memiliki varians yang lebih besar daripada rata-rata (over-dispersion).
2. **Indeks  $p$ :** Pemilihan nilai indeks  $p$  sangat penting dalam menentukan jenis distribusi Tweedie yang akan digunakan. Nilai  $p$  yang berbeda akan menghasilkan distribusi yang berbeda dalam keluarga Tweedie, seperti distribusi Poisson untuk  $p = 1$ , distribusi Gamma untuk  $p = 2$ , dan distribusi Normal untuk  $p = 0$ . Oleh karena itu, pemilihan  $p$  harus disesuaikan dengan karakteristik data yang dianalisis.
3. **Model Regresi Generalized Linear:** Model Poisson Tweedie merupakan bagian dari pendekatan Generalized Linear Model (GLM). Oleh karena itu, penggunaannya memerlukan pemahaman tentang GLM dan bagaimana menghubungkan variabel respon dengan variabel prediktor menggunakan fungsi link yang sesuai.
4. **Estimasi Parameter:** Proses estimasi parameter dalam model Poisson-Tweedie harus dilakukan dengan metode yang tepat, seperti Maximum Likelihood Estimation (MLE), untuk memastikan bahwa model yang dihasilkan dapat diandalkan dan akurat dalam memprediksi nilai-nilai yang diharapkan.
5. **Asumsi Model:** Seperti model statistik lainnya, model Poisson Tweedie juga memiliki asumsi yang harus dipenuhi, seperti independensi observasi dan hubungan yang sesuai antara variabel respon dan prediktor.

### 1.6.4 Distribusi Tweedie

Distribusi Tweedie adalah sebuah keluarga distribusi yang merupakan bagian dari Model Dispersi Eksponensial (EDM). Distribusi Tweedie adalah keluarga distribusi probabilitas yang mencakup beberapa distribusi terkenal seperti distribusi normal, gamma, invers Gaussian, dan Poisson. Distribusi Tweedie adalah salah satu distribusi dalam GLM yang sangat berguna karena memiliki fleksibilitas dalam menangani data yang memiliki rata-rata dan variansi yang terhubung secara non-linear (Dunn & Smyth, 2005). Bentuk umum dari model distribusi Tweedie.

$$\ln \pi (y; \mu, \phi) = \left[ \frac{y}{\mu^{(1-p)}} (1 - p) - \frac{\mu^{-p}}{2-p} \right] \phi + c(y, \phi) \quad (3)$$

dengan:

- y : nilai observasi (respon)
- $\mu$  : parameter rata-rata dari distribusi
- $\phi$  : parameter disperse yang mengontrol variabilitas data di sekitar rata-ratanya
- p : parameter yang menentukan bentuk distribusi Tweedie.
- $c(y, \phi)$  : fungsi yang tidak tergantung pada  $\mu$ , konstanta normalisasi.

Kasus-kasus khusus distribusi Tweedie. Distribusi Tweedie mencakup beberapa kasus khusus, seperti:

1. Distribusi Normal ( $p = 0$ )
2. Distribusi Poisson ( $p = 1$ )
3. Distribusi Gamma ( $p = 2$ )
4. Distribusi Invers Gaussian ( $p = 3$ )

**Tabel 1.** Karakteristik Family Tweedie

Sebaran	Jenis Data	Nilai p
Neyman Type A	Over, under, zero-inflation	$p = 1$
Compound Poisson Tweedie	Over, under, zero-inflation	$1 < p < 2$
Binomial Negatif	Over, under	$p = 2$
Poisson Invers Gaussian	Over, heavy tail	$p = 3$

Sumber: Bonat et al. (2017)

Keluarga distribusi ini memiliki karakteristik (Miller et al., 2012). Sebagai berikut:

1. Rata-rata (mean) dari  $E(Y) = \mu$
2. Varians dari  $Var(Y) = \phi \mu^p$

Angka  $p$  dalam fungsi varians adalah parameter bentuk tambahan untuk distribusi. Fungsi kepadatan probabilitas untuk keluarga Tweedie sangat kompleks dan tidak dapat dinyatakan dalam bentuk tertutup yang sederhana (tetapi kadang-kadang dinyatakan sebagai serangkaian fungsi). Dalam GLM, model distribusi Tweedie (Dunn & Smyth, 2005) dinyatakan sebagai:

$$Y_i \sim \text{Tweedie}(\mu_i, \phi, p)$$

Dengan  $\mu_i$  ditentukan oleh kombinasi linear dari variabel prediktor melalui fungsi

$$\text{link log: } \mu_i = \exp(X_i, \beta)$$

### 1.6.5 Distribusi Poisson Tweedie

Distribusi Poisson Tweedie adalah sebuah distribusi probabilitas yang menggabungkan sifat dari distribusi Poisson dan distribusi Tweedie. Distribusi ini dapat mengatasi variasi yang lebih besar dalam data count, yang sering kali terjadi dalam praktek di mana data memiliki overdispersi atau underdispersi yang signifikan.

Perbedaan utama antara distribusi Poisson Tweedie dan distribusi Tweedie terletak pada sifat-sifat masing-masing distribusi dalam hal hubungan antara rata-rata dan variansi, serta kemampuan untuk memodelkan data count dengan variasi yang berbeda-beda.

Distribusi Poisson Tweedie merupakan pengembangan dari distribusi Tweedie yang dirancang khusus untuk menangani kasus-kasus di mana distribusi Poisson atau distribusi Tweedie tunggal tidak mencukupi untuk menjelaskan karakteristik data yang ditemui dalam praktik.

Bentuk Umum Fungsi Massa Probabilitas PoissonTweedie:

$$f(y; \mu, \phi, \theta) = \frac{\mu^y}{y!} \left( \frac{\phi^\theta}{(\phi + (y + 1)^\theta)} \right)^{\frac{1}{\theta}} \frac{1}{B(y, \theta)} \quad (4)$$

dengan:

$y$  : variabel acak yang mewakili jumlah kejadian

$\mu$  : parameter mean (rata-rata)

$\phi$  : parameter power yang mengatur distribusi Tweedie

$\theta$  : parameter bentuk yang menentukan jenis distribusi Tweedie yang spesifik

$B(y, \theta)$  : fungsi normalisasi untuk menjaga integral fungsi massa probabilitas

### 1.6.6 Model Linear Umum

Model linear umum atau Generalized Linear Model (GLM) adalah perluasan dari model regresi linier di mana variabel respon adalah anggota keluarga eksponensial yang memungkinkan variabel respon untuk memiliki distribusi non-Gaussian. GLM mencakup berbagai model, seperti regresi data panel, yang mengkombinasikan data time series dan cross section dalam sebuah persamaan. Dengan menggunakan GLM, model regresi dapat diadaptasi untuk memodelkan berbagai jenis hasil, seperti

kategori, hitungan, atau waktu terjadinya suatu peristiwa. GLM merupakan pengembangan dari model regresi umum untuk variabel respon yang tidak mengharuskan berdistribusi normal (Sari, 2018). Pada GLM, terdapat hubungan antara rata-rata dari variabel respon dengan variabel prediktor yang dihubungkan oleh suatu fungsi linear (Cahyandari, 2014)(Rusyana et al., 2021).

Terdapat tiga komponen utama yang harus ada dalam GLM (Hugo Bonat et al., 2017; McCullagh & Nelder, 2010) diantaranya:

1. Komponen Acak

Di dalam GLM, harus terdapat  $y_1, y_2, y_3, \dots, y_n$  yang merupakan variabel respon berasal dari distribusi keluarga eksponensial.

2. Komponen Sistematis

Komponens sistematis merupakan komponen kedua dalam GLM yang menghubungkan vektor  $\eta_i$  dengan variabel prediktor menggunakan prediktor linear.  $\eta_i$  dapat dinyatakan dengan persamaan berikut:

$$\eta_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}$$

$$\eta_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij}$$

Dalam bentuk matriks:

$$\begin{bmatrix} \eta_i \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}$$

dengan  $n$  banyaknya pengamatan dan  $p$  merupakan banyaknya variabel prediktor.

3. Fungsi Penghubung

Fungsi penghubung atau *link function*  $g(\cdot)$  adalah suatu fungsi yang menghubungkan komponen acak dengan komponen sistematis.

$$g(\mu_i) = \eta_i = \mathbf{X}\boldsymbol{\beta}$$

Fungsi transformasi  $g$  disebut *link function*. Yang berfungsi dalam menentukan bagaimana nilai rata-rata berhubungan dengan variabel-variabel bebas. Adapun *link function* yang digunakan untuk distribusi Tweedie adalah fungsi logaritma (log link)

**1.6.7 Metode Maximum Likelihood Estimation**

Metode *Maximum Likelihood Estimation* (MLE) digunakan untuk mengestimasi parameter suatu model yang fungsi probabilitasnya diketahui. Misalkan  $Y_1, Y_2, \dots, Y_n$  adalah sampel random dan berdistribusi Poisson, maka fungsi probabilitas bersama dari  $Y_1, Y_2, \dots, Y_n$  dapat dinyatakan pada persamaan (5) sebagai berikut.

$$f(y_1, y_2, \dots, y_n | \boldsymbol{\theta}) = f(y_1 | \boldsymbol{\theta}) \cdot f(y_2 | \boldsymbol{\theta}) \dots f(y_n | \boldsymbol{\theta}) = f(y_i | \boldsymbol{\theta}) \tag{5}$$

dengan  $\theta$  adalah parameter yang tidak diketahui. Fungsi *likelihood* didefinisikan sebagai fungsi probabilitas bersama dari  $Y_1, Y_2, \dots, Y_n$  yang dimisalkan sebagai fungsi dari  $\theta$  yang dituliskan pada Persamaan (6) berikut (Hoog dkk., 2019).

$$L(\theta | y) = f(y_1, y_2, \dots, y_n | \theta) = f(y_1 | \theta) \cdot f(y_2 | \theta) \dots f(y_n | \theta) = \prod_{i=1}^n f(y_i | \theta) \quad (6)$$

Penaksir dari maksimum *likelihood*  $\hat{\theta}$  yakni nilai  $\theta$  yang memaksimumkan fungsi *likelihood*  $L(\theta | y)$ . Akan tetapi, lebih mudah bekerja dengan logaritma natural dari fungsi *likelihood*, yaitu  $l(\theta | y) = \ln L(\theta | y)$  sehingga dapat dituliskan pada Persamaan (7) berikut.

$$l(\theta | y) = L(\theta | y) = \ln\{\prod_{i=1}^n f(y_i | \theta)\} = \sum_{i=1}^n \ln f(y_i | \theta) \quad (7)$$

Untuk mendapatkan nilai  $\theta$  yang memaksimumkan fungsi  $l(\theta | y)$ , maka  $l(\theta | y)$  diturunkan terhadap  $\theta$  dan kemudian menyamakannya dengan nol seperti pada persamaan berikut (Hogg dkk., 2019).

$$\frac{\partial l(\theta | y)}{\partial \theta} = 0 \quad (8)$$

MLE pada model Poisson-Tweedie digunakan untuk menemukan nilai parameter yang memaksimumkan fungsi kemungkinan (*likelihood*) dari data sampel.

### 1.6.8 Metode Fisher Scoring

Model regresi Poisson Tweedie adalah model statistik yang digunakan untuk memodelkan data cacah dengan dispersi yang lebih besar daripada yang diharapkan oleh model Poisson standar. Model ini merupakan kombinasi dari model Poisson dan model Tweedie, yang memungkinkan untuk memodelkan data dengan variasi yang lebih kompleks. Berikut adalah beberapa langkah dalam menggunakan metode Fisher Scoring:

1. Menentukan persamaan Log-Likelihood: Pertama, persamaan log-likelihood untuk model Poisson dan Tweedie
2. Turunan Parsial Pertama: Selanjutnya, turunan parsial pertama dari fungsi log-likelihood terhadap parameter yang diestimasi harus ditentukan.
3. Proses Iterasi Fisher Scoring: Persamaan iterasi Fisher Scoring digunakan untuk memperbarui parameter. Persamaan ini adalah:

$$(\beta_j)^{k+1} = (\beta_j)^k - (I\beta_j^k)^{-1} \cdot \frac{\partial [L(\beta, \phi, \theta)]}{\partial \beta_j} \quad (9)$$

dengan:

$(\beta_j)^k$  : nilai parameter pada iterasi k

$(I\beta_j^k)^{-1}$  : invers dari matriks informasi Fisher pada iterasi k

$\frac{\partial [L(\beta, \phi, \theta)]}{\partial \beta_j}$  : turunan parsial pertama dari fungsi log-likelihood pada iterasi k

4. Jika belum mendapatkan penaksir parameter yang konvergen maka dilanjutkan kembali ke langkah ke-5 dan berhenti jika  $\|(\beta_j)^{k+1} - \beta_j^k\| < \varepsilon$
5. Diperoleh hasil estimasi: Hasil estimasi akhir adalah nilai parameter pada iterasi terakhir,  $(\beta_j)^{k+1}$  dengan demikian, estimasi parameter menggunakan metode Fisher Scoring dalam model regresi Poisson Tweedie melibatkan iterasi yang berulang untuk memperbarui parameter sampai konvergensi tercapai.

### 1.6.9 Asumsi Model Regresi Poisson Tweedie

Sebelum melakukan pengujian model regresi Poisson Tweedie, dilakukan pengujian asumsi yang harus dipenuhi pada model regresi Poisson terlebih dahulu.

#### Uji Multikolinearitas

Suatu model regresi yang mengalami multikolinieritas apabila terdapat hubungan linier tinggi diantara sejumlah atau semua variabel prediktor dari suatu model regresi. Kasus multikolinieritas bisa diketahui dengan melihat nilai VIF (*Variance Inflation Factors*). Statistik uji yang digunakan untuk pengujian hipotesis sebagai berikut:

$$VIF = \frac{1}{1-R^2} \quad (10)$$

dengan  $R^2$  adalah koefisien determinasi antar variabel prediktor satu dengan yang lainnya. Kriteria penolakan,  $H_0$  ditolak jika  $VIF > 10$ . (Putri dkk., 2019).

#### Uji Overdispersi

Pada model regresi Poisson terdapat sejumlah asumsi yang harus dipenuhi. Salah satunya adalah asumsi kesamaan antara mean dan variansi yang disebut equidispersi. Akan tetapi, dalam analisis data statistik sering dijumpai kondisi data dengan variansinya lebih kecil dari rata-rata ataupun sebaliknya. Kondisi ini disebut dengan underdispersi (*underdispersion*) atau overdispersi (*overdispersion*) (Kusuma dkk., 2013).

Overdispersi bisa menimbulkan perkiraan parameter yang diperoleh tidak efisien. Penggunaan yang tidak benar pada model regresi Poisson (yang mengalami overdispersi) bisa berakibat fatal dalam interpretasi model, terutama pada perkiraan parameter model karena bisa menaksir *standard error* yang terlalu rendah dan bisa menyampaikan kesimpulan yang salah tentang signifikan atau tidaknya parameter regresi yang terlibat (Darnah, 2011). Dalam mendeteksi kasus overdispersi dalam data, dapat dilihat dari hasil uji *Deviance*. Adapun hipotesis yang digunakan sebagai berikut:

$H_0 : \phi = 1$  Tidak terjadi overdispersi

$H_1 : \phi > 1$  Terjadi overdispersi

Statistik uji sebagai berikut:

$$\phi = \frac{D^2}{db} \quad (11)$$

Untuk

$$D^2 = 2 \sum_{i=1}^n \left\{ \left( y_i \ln \left( \frac{y_i}{\mu_i} \right) \right) - (y_i - \mu_i) \right\} \quad (12)$$
$$db = n - p$$

Dengan:

$\mu_i$ : penduga bagi respon rata-rata ke- $i$

$n$ : banyaknya total observasi dalam dataset

$p$ : banyaknya parameter dalam model

dengan kriteria penolakan,  $H_0$  ditolak jika nilai  $\phi > 1$  (Kusuma dkk., 2013).

#### 1.6.10 Pemilihan Model Terbaik

Model terbaik dapat ditentukan cara melihat nilai information criteria fit test dapat dilihat dari nilai Akaike Information Criteria (AIC). Nilai AIC didasarkan pada log-likelihood function. AIC merupakan kriteria untuk memilih model dalam ekonometrika. AIC mampu menunjukkan seberapa tepat model dengan data yang dimiliki secara mutlak (Haris dkk., 2016). Model yang memiliki nilai AIC, BIC dan deviance lebih kecil maka model tersebut adalah model terbaik. Selain nilai AIC dan BIC, nilai Log Likelihood juga dapat dilihat sebagai pertimbangan model terbaik. Berbeda dengan nilai AIC dan BIC, model yang memiliki nilai Log Likelihood paling besar merupakan model terbaik (Afri, 2017). Rumus AIC adalah

$$AIC = -2 \ln L + 2p \quad (13)$$

dengan:

$AIC$  : Akaike Information Criteria

$\ln L$  : nilai dari fungsi ln-likelihood

$p$  : banyaknya parameter

#### 1.6.11 Angka Kematian Bayi

Berdasarkan data Dinas Kesehatan Provinsi Sulawesi Selatan, di tahun 2021 total AKB mencapai 844 kasus (Sulselprov, 2021). Jumlah kejadian AKB adalah banyaknya kematian bayi berusia dibawah satu tahun, per 1000 kelahiran hidup pada satu tahun tertentu (Prahutama dkk., 2017; Yasril dkk., 2022). Definisi AKB juga dijelaskan dalam daftar istilah yang dikeluarkan oleh BPS dan dituangkan pada laman website BPS Sulawesi Selatan (<https://sulsel.bps.go.id/istilah.html>, halaman 4). AKB merupakan salah satu variabel yang memiliki dampak terhadap kesehatan (Yasril dkk., 2022). Sehingga, Tingkat AKB merupakan salah satu indikator derajat kesehatan suatu bangsa (Fitri dkk., 2021; Putri dkk., 2022; Yasril dkk., 2022). AKB

yang rendah mengindikasikan bahwa kesehatan bangsa tersebut tinggi atau optimal. Indonesia menempati urutan kelima dengan jumlah AKB yaitu 30 per 1000 kelahiran hidup (Alfahmi, 2023). Diantara negara ASEAN lainnya, Indonesia masih termasuk dalam negara dengan kejadian AKB tinggi jika dibandingkan dengan target Sustainable Development Goals (SDGs) yaitu 12 per 1000 kelahiran hidup. Oleh karena itu, angka kematian ibu dan bayi menjadi salah satu target yang telah ditentukan dalam tujuan pembangunan SDGs yaitu untuk menurunkan AKB menjadi 12 per 1.000 kelahiran hidup pada tahun 2030 (Kementerian Kesehatan RI Badan Penelitian dan Pengembangan, 2018) (Yasril dkk., 2022).

Menurut laporan BPS tahun 2020, tingkat AKB di Sulawesi Selatan sebesar 17 per 1000 kelahiran hidup, lebih tinggi dari rata-rata nasional yaitu 11 per 1000 kelahiran hidup. Selain itu, Sulawesi Selatan merupakan satu dari lima provinsi dengan Angka Kematian Ibu (AKI) dan Angka Kematian bayi (AKB) tertinggi di Indonesia, di samping DKI Jakarta, Banten, Jawa Timur, dan Sumatera Utara.

#### **1.6.12 Faktor-Faktor Kematian Bayi**

Berdasarkan data pada profil dinas kesehatan Indonesia, penyebab kematian bayi terbanyak adalah kondisi dengan Berat Badan Lahir Rendah (BBLR). Sementara itu, penyebab kematian lainnya diantaranya Asfiksia, Pneumonia, Diare, dan lainnya (Dinkes-Makassar, 2022). Indikator yang menggambarkan upaya kesehatan yang dapat dilakukan untuk mengurangi risiko kematian pada bayi (periode neonatal) yaitu 6-48 jam setelah lahir adalah cakupan Kunjungan Neonatal Pertama atau KN1. Pelayanan dalam kunjungan ini (Manajemen Terpadu Balita Muda) antara lain meliputi termasuk konseling perawatan bayi baru lahir, ASI eksklusif dan lainnya (Dinkes-Makassar, 2022). Pemberian ASI eksklusif pada bayi usia sampai dengan 6 bulan merupakan upaya pencegahan serta penanganan masalah gizi pada bayi yang berujung kematian. Berikut penjelasan terkait BBLR, Pemberian ASI eksklusif, Asfiksia, Diare, dan pneumonia yang merupakan faktor kematian bayi.

#### **BBLR**

BBLR sebagai indikator berat badan kelahiran sering dijadikan perhatian di berbagai negara terutama pada negara berkembang atau negara dengan sosio- ekonomi rendah. WHO (World Health Organization) membutuhkan definisi pada BBLR sebagai bayi yang lahir dengan berat  $\leq 2500$  gr (Sitepu et al., 2023). WHO memberikan pengelompokan pada BBLR menjadi 3, yaitu BBLR: 1500–2499 gram, BBLSR (Bayi Berat Lahir Sangat Rendah): 1000- 1499 gram, BBLER (Bayi Berat Lahir Ekstrem Rendah):  $< 1000$  gram (Putri et al., 2022; Sitepu et al., 2023). Salah satu penyebab tingginya angka kematian bayi (AKB) adalah berat badan lahir rendah (BBLR). Bayi dengan berat lahir rendah (BBLR) merupakan bayi yang rentan terhadap penyakit. Dengan demikian, ketahanan hidupnya rendah (Dinkes-Makassar, 2022). Bayi dengan BBLR adalah bayi dengan berat lahir kurang dari 2500 gram tanpa memandang masa gestasi yang ditimbang dalam 1 (satu) jam setelah lahir. Kondisi bayi BBLR disebabkan karena kondisi ibu saat hamil

(kehamilan remaja, malnutrisi, dan komplikasi kehamilan), bayi kembar, janin memiliki kelainan atau kondisi bawaan, dan gangguan pada plasenta yang menghambat pertumbuhan bayi (Dinkes, 2021).

Secara statistik menunjukkan 90% kejadian BBLR didapatkan di negara berkembang dan angka kematiannya 35 kali lebih tinggi dibanding pada bayi dengan berat lahir normal. Oleh karena itu, BBLR termasuk faktor utama dalam peningkatan mortalitas, morbiditas dan disabilitas neonatus, bayi dan anak, serta memberikan dampak jangka panjang terhadap kehidupannya dimasa depan (Dinkes, 2021). Persentase kejadian BBLR di Sulawesi Selatan Tahun 2020 sebesar 4,24% dari 154.733 kelahiran hidup. Angka terendah kasus BBLR terjadi di Kabupaten Takalar sebesar 1,22% dari 5.817 kelahiran hidup, dan Kabupaten Tana Toraja sebesar 1,86% dari 3.503 kelahiran hidup. Kasus BBLR tertinggi terjadi di Kabupaten Bulukumba sebesar 13,12% dari 6.770 kelahiran hidup dan Kabupaten Luwu sebesar 10,80% dari 6.219 kelahiran hidup (Dinkes-Sulsel, 2021). Menurut BPS (2021), sebesar 35,2% kematian bayi disebabkan karena BBLR. Bayi dengan kondisi BBLR memiliki risiko 20 kali untuk mengalami kematian dibandingkan dengan bayi normal. Selain itu

### **Pemberian ASI eksklusif**

Air Susu Ibu (ASI) merupakan sumber makanan terbaik yang mengandung enzim, hormon, kandungan imunologik dan anti infeksi yang dibutuhkan bayi terutama pada usia 0-6 bulan (Sitepu et al., 2023). Pemberian ASI eksklusif sangat berperan dalam menurunkan angka kesakitan dan kematian anak (Sitepu et al., 2023). Salah satu upaya dalam pencegahan serta penanganan masalah gizi pada bayi adalah dengan melakukan Pemberian ASI eksklusif pada bayi usia sampai dengan 6 bulan. Adanya kandungan kolostrum pada ASI akan menjadi antibodi untuk bayi, karena mengandung protein untuk daya tahan tubuh dan bermanfaat untuk mematikan kuman dalam jumlah tinggi. Dengan demikian, pemberian ASI secara eksklusif dapat mengurangi risiko kematian pada bayi (Dinkes, 2021).

### **Asfeksia**

Asfiksia termasuk penyebab utama kematian bayi di Sulawesi Selatan berdasarkan data BPS (2021) yang ditampilkan pada profil dinas kesehatan Makassar (Dinkes-Makassar, 2022). Asfiksia merupakan suatu keadaan yang dialami pada bayi berupa kegagalan bernafas secara spontan setelah lahir dan sangat beresiko untuk terjadinya kematian. Keadaan bayi tidak spontan bernafas dan teratur dapat menyebabkan terhambatnya suplai oksigen ke tubuh pada saat sebelum, selama dan setelah proses persalinan yang menimbulkan akibat buruk bagi kehidupan bayi (Irwan dkk., 2019; Mardiyanti & Hardianti, 2023). Bayi yang lahir prematur atau usia kehamilan < 37 minggu juga dapat menyebabkan asfiksia diakibatkan belum maksimalnya tingkat kematangan fungsi sistem organ tubuh sehingga bayi sulit untuk beradaptasi dengan kehidupan transisi diluar rahim atau ektrauterin (Irwan dkk., 2019). Asfiksia merupakan suatu keadaan pada bayi baru lahir yang mengalami gagal bernafas secara spontan dan teratur segera setelah lahir, sehingga bayi tidak

dapat memasukkan oksigen dan tidak dapat mengeluarkan zat asam arang dari tubuhnya, sehingga dapat menurunkan O<sup>2</sup> (oksigen) dan mungkin meningkatkan CO<sup>2</sup> (karbondioksida) yang menimbulkan akibat buruk dalam kehidupan lebih 4 lanjut. Asfiksia dapat dibagi menjadi 3 yaitu, asfiksia ringan, asfiksia sedang, dan berat.

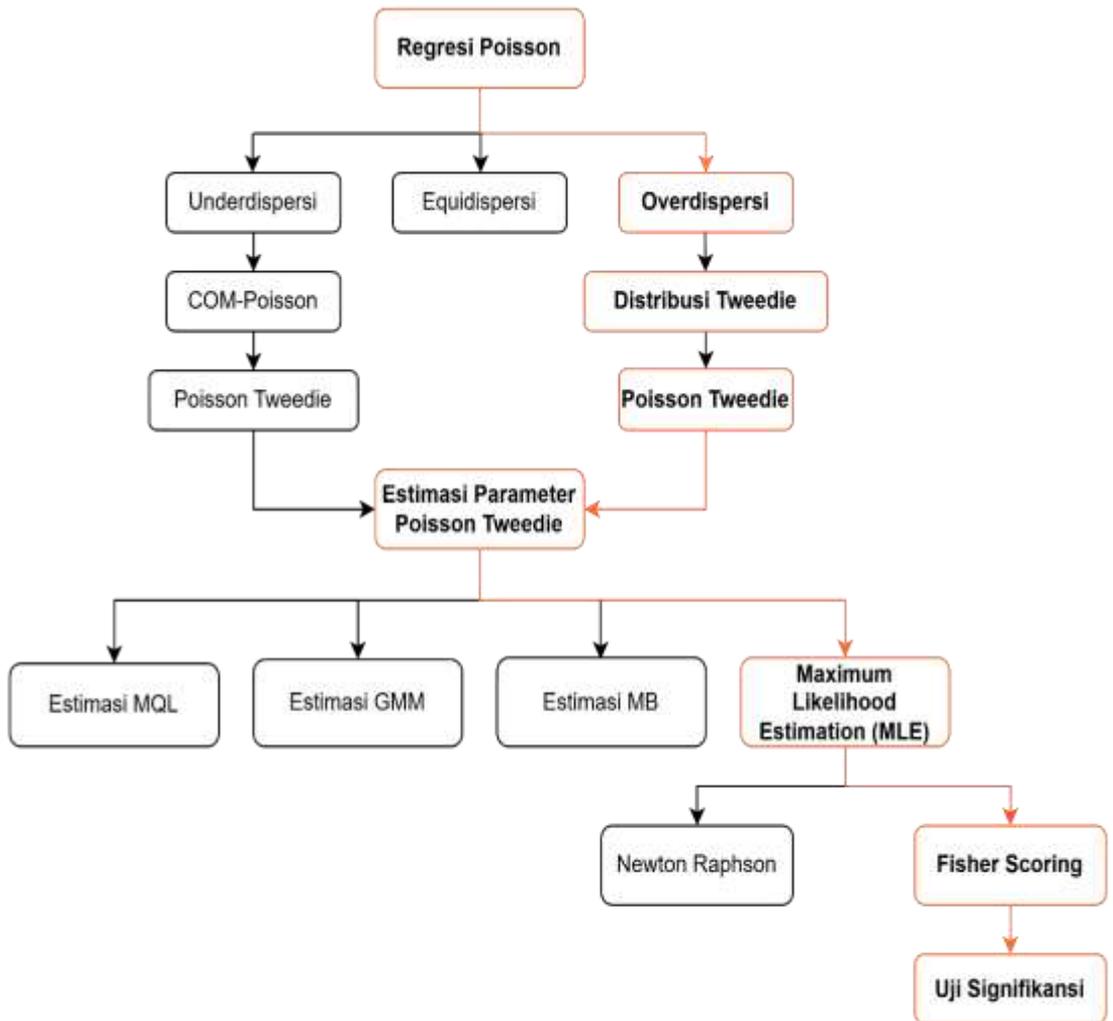
### **Pneumonia**

Pneumonia adalah infeksi akut pada jaringan paru-paru yang meliputi alveolus dan jaringan interstitial disebabkan mikroorganisme seperti virus, bakteri dan jamur (Dinkes, 2021; Hane, 2023). Pada dasarnya, kantong udara pada paru-paru yang seharusnya berisi udara menjadi berisi cairan atau nanah (Dinkes-Makassar, 2022). Diagnosa pneumonia didasarkan pada riwayat medis pasien, pemeriksaan fisik, rontgen dada, dan pemeriksaan laboratorium (Hermayani & Maran, 2023). Gejala umum pneumonia termasuk batuk berdahak, demam, nyeri otot, sakit kepala, nyeri dada, dan kesulitan bernafas yang menyebabkan tubuh kekurangan oksigen sehingga berujung pada kematian (Dinkes-Makassar, 2022; Hermayani & Maran, 2023). Pneumonia menjadi salah satu penyebab terbesar morbiditas dan mortalitas pada bayi dibawah lima tahun pada tahun 2021 (Dalimunthe & Susilawati, 2023; Hermayani & Maran, 2023). Morbiditas dan mortalitas merupakan indikator yang menggambarkan tingkat derajat kesehatan suatu wilayah dengan indikator yang digunakan adalah angka kematian bayi (AKB) dan angka kesakitan bayi. Pneumonia karena sistem kekebalan tubuh yang masih rendah.

### **Diare**

Penyakit diare disebabkan oleh bakteri, virus dan parasit organisme, dan masih menjadi masalah kesehatan dinegara berkembang, termasuk di Indonesia. Angka kesakitan dan kematian akibat penyakit diare masih tinggi, sehingga menyebabkan penyakit diare menjadi masalah kesehatan. Tindakan dalam pencegahan diare ini antara lain dengan perbaikan keadaan lingkungan, seperti penyediaan sumber air minum yang bersih, penggunaan jamban, pembuangansampah pada tempatnya, sanitasi perumahan dan penyediaan tempat pembuangan air limbah yang layak. Perbaikan perilaku ibu terhadap balita seperti pemberian ASI sampai anak berumur 2 tahun, perbaikan cara menyapih, kebiasaan mencuci tangan sebelum dan sesudah beraktivitas, membuang tinja anak pada tempat yang tepat, memberikan imunisasi morbilitas (Dinkes-Makassar, 2022).

### 1.6.13 Kerangka Konseptual



## BAB II

### METODOLOGI PENELITIAN

#### 2.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang merupakan data kasus kematian bayi di Sulawesi Selatan tahun 2020-2023. Unit pengamatan yang diambil adalah data di setiap kabupaten dan kota di Provinsi Sulawesi Selatan. Data tersebut diperoleh dari pengambilan data di kantor Dinkes Provinsi Sulawesi Selatan. Pada penelitian ini menggunakan bantuan *software R-Studio* dan *IBM SPSS Statistics*.

#### 2.2 Waktu dan Tempat Penelitian

Penelitian ini dilaksanakan pada bulan Mei 2024 di Dinas Kesehatan Provinsi Sulawesi Selatan berupa pengambilan data penelitian, pascasarjana Universitas Hasanuddin Makassar berupa administrasi dan perumusan model Poisson, Poisson Tweedie yang sesuai dan perpustakaan Pusat Universitas Hasanuddin berupa penyusunan proposal penelitian.

#### 2.3 Populasi dan Sampel Penelitian

Populasi dalam penelitian ini adalah seluruh data AKB, jumlah bayi yang diberi ASI eksklusif, pneumonia, asfeksia, Berat Badan Lahir Rendah (BBLR) dan diare pada bayi di seluruh Kabupaten/Kota Provinsi Sulawesi Selatan. Adapun sampel yang digunakan adalah data AKB tahun 2020 sampai 2023.

#### 2.4 Identifikasi Variabel

Variabel respon ( $Y$ ) yang digunakan dalam penelitian ini yakni Angka Kematian Bayi (AKB). Variabel predictor ( $X$ ) dalam penelitian ini terdiri dari empat variabel diantaranya: asi eksklusif, pnemonia, asfeksia, dan Berat Badan Lahir Rendah (BBLR). Deskripsi variabel penelitian di sajikan dalam Tabel. 2.1.

**Tabel 2.** Variabel Penelitian

Variabel	Keterangan	Deskripsi	Skala
$Y$	Angka Kematian Bayi	Jumlah kematian bayi di Setiap kabupaten di Sulawes - Selatan	Numerik, Ratio
$X_1$	ASI Eksklusif	Jumlah bayi yang diberi ASI Eksklusif di setiap kabupaten di Sulawesi-Selatan	Numerik, Ratio
$X_2$	Pneumonia	Jumlah bayi dengan infeksi akut paru-paru (pneumonia) di setiap kabupaten di Sulawesi Selatan.	Numerik, Ratio

Variabel	Keterangan	Deskripsi	Skala
$X_3$	Asfeksia	Jumlah bayi yang mengalami kondisi medis saat kadar oksigen berkurang di dalam tubuh di setiap kabupaten di Sulawesi Selatan.	Numerik, Ratio
$X_4$	Berat Badan Lahir Rendah (BBLR)	Jumlah bayi dengan kondisi berat badan lahir rendah di setiap kabupaten di Sulawesi Selatan.	Numerik, Ratio
$X_5$	Diare	Jumlah bayi akibat diare di setiap kabupaten di Sulawesi Selatan.	Numerik, Ratio

## 2.6 Definisi Operasional Variabel Penelitian

Angka Kematian Bayi (AKB) merupakan variabel terikat disimbolkan ( $Y$ ) yaitu banyaknya kematian bayi usia di bawah 1 tahun (0-11 bulan) per 1000 kelahiran hidup pada satu tahun tertentu (Prahutama et al., 2017). Adapun variabel independen yaitu variabel bebas disimbolkan ( $X$ ) diantaranya: Pemberian ASI Eksklusif ( $X_1$ ) yaitu pemberian hanya ASI selama 6 bulan dan dapat terus dilanjutkan sampai dengan dua tahun (Lestari, 2018), pneumonia ( $X_2$ ) yaitu penyakit infeksi saluran pernafasan akut yang mengenai jaringan paru-paru (alveolus) (Afriani & Oktavia, 2021), asfeksia ( $X_3$ ) yaitu kegagalan nafas secara spontan dan teratur pada saat lahir atau beberapa saat setelah lahir (Dinkes-Sulsel, 2021; Florencia et al., 2022; Mardiyanti & Hardianti, 2023), Berat Badan Lahir Rendah (BBLR) ( $X_4$ ) yaitu kondisi berat lahir bayi kurang dari 2500 gram (Dinkes-Sulsel, 2021) dan kondisi baik yang diare ( $X_1$ ) (Dinkes-Makassar, 2022).

## 2.7 Prosedur Penelitian

Model PTw yang merupakan salah satu model yang dapat menangani kondisi overdispersi suatu data count misalnya data AKB di Sulawesi Selatan. Tahapan analisis berupa langkah-langkah yang dilakukan untuk mencapai tujuan penelitian ini sebagai berikut:

1. Eksplorasi data AKB dan mendeteksi data pada setiap variabel (variabel) penelitian yang bertujuan untuk melihat sebaran data.
2. Membangun formulasi model data AKB menggunakan regresi Poisson. Langkah-langkah yang dilakukan adalah sebagai berikut:
  - a. Melakukan uji multikolinearitas pada persamaan (10)
  - b. Melakukan uji overdispersi pada data (uji deviance) menggunakan persamaan (11)
3. Mengidentifikasi terkait asumsi equidispersi pada regresi poisson apakah terjadi overdispersi atau underdispersi.
4. Memilih model sebagai solusi dari penanganan overdispersi menggunakan model PTw.

5. Mengestimasi parameter model PTw menggunakan *Maximum Likelihood* untuk memaksimalkan fungsi *Likelihood* dengan pendekatan Algoritma Fisher Scoring.
  - b. Melakukan pemodelan berdasarkan asumsi khusus pada distribusi Tweedie
  - c. Memilih nilai  $p$  pada distribusi Tweedie dengan AIC terkecil
6. Melakukan pemodelan data menggunakan model PTw
7. menentukan model terbaik berdasarkan nilai galat baku parameter dan nilai AIC (13)
8. Menginterpretasi hasil dari suatu model regresi PTw yang telah diperoleh.

## 2.8 Diagram Alir

