



**TUGAS AKHIR - KS184822**

**PEMODELAN JUMLAH KEMATIAN BAYI  
DI SULAWESI BARAT MENGGUNAKAN  
METODE *ZERO INFLATED POISSON REGRESSION***

**ROCHMATUN ANNISA  
NRP 062115 4000 0052**

**Dosen Pembimbing  
Erma Oktania Permatasari, S.Si., M.Si  
Dr. Agnes Tuti Rumiati, M.Sc**

**PROGRAM STUDI SARJANA  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2019**





**TUGAS AKHIR - KS184822**

**PEMODELAN JUMLAH KEMATIAN BAYI  
DI SULAWESI BARAT MENGGUNAKAN  
METODE *ZERO INFLATED POISSON REGRESSION***

**ROCHMATUN ANNISA  
NRP 062115 4000 0052**

**Dosen Pembimbing  
Erma Oktania Permatasari, S.Si., M.Si  
Dr. Agnes Tuti Rumiati, M.Sc**

**PROGRAM STUDI SARJANA  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2019**





**FINAL PROJECT - KS184822**

**MODELLING OF INFANT MORTALITY IN WEST SULAWESI  
USING ZERO INFLATED POISSON REGRESSION METHOD**

**ROCHMATUN ANNISA  
NRP 062115 4000 0052**

**Supervisors**

**Erma Oktania Permatasari, S.Si., M.Si  
Dr. Agnes Tuti Rumiati, M.Sc**

**UNDERGRADUATE PROGRAMME  
DEPARTMENT OF STATISTICS  
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2019**

*(Halaman ini sengaja dikosongkan)*

## LEMBAR PENGESAHAN

### PEMODELAN JUMLAH KEMATIAN BAYI DI SULAWESI BARAT MENGGUNAKAN METODE ZERO INFLATED POISSON REGRESSION

#### TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat  
Memperoleh Gelar Sarjana Statistika  
pada

Program Studi Sarjana Departemen Statistika  
Fakultas Matematika, Komputasi, dan Sains Data  
Institut Teknologi Sepuluh Nopember

Oleh :

**Rochmatun Annisa**  
NRP. 062115 4000 0052

Disetujui oleh Pembimbing:

**Erma Oktania Permatasari, S.Si., M.Si.**

NIP. 19881007 201404 2 002

**Dr. Agnes Tuti Rumiati, M.Sc.**

NIP. 19570724 198503 2 002



Mengetahui,  
Kepala Departemen Statistika

**Dr. Suhartono**

NIP. 19710929 199512 1 001

SURABAYA, JULI 2019

*(Halaman ini sengaja dikosongkan)*



# PEMODELAN JUMLAH KEMATIAN BAYI MENGUNAKAN METODE *ZERO INFLATED POISSON REGRESSION*

**Nama Mahasiswa** : Rochmatun Annisa  
**NRP** : 062115 4000 0052  
**Departemen** : Statistika  
**Dosen Pembimbing** : Erma Oktania P., S.Si., M.Si.  
Dr. Agnes Tuti Rumiati, M.Sc.

## **Abstrak**

*Jumlah kematian bayi yang salah satunya dapat dicerminkan pada Angka Kematian Bayi (AKB) merupakan salah satu indikator untuk mengetahui perkembangan pembangunan sebuah negara. Target AKB pada Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2015-2019 yaitu sebesar 24 kematian per 1.000 kelahiran hidup. Berdasarkan SDKI 2017 nilai AKB di Sulawesi Barat sebesar 42 kematian per 1.000 kelahiran hidup. Faktor yang diduga mempengaruhi kematian bayi di Sulawesi Barat tidak hanya dalam bidang kesehatan tapi juga sosial ekonomi. Data jumlah kematian bayi di Sulawesi Barat merupakan data count yang memuat nilai 0. Oleh karena itu, analisis yang digunakan adalah Zero Inflated Poisson (ZIP) Regression. Hasil analisis ZIP regression didapatkan model terbaik dengan AIC dan AICc terkecil yaitu 324,3407 dan 326,7407 adalah model dengan kombinasi variabel persentase posyandu aktif, persentase penduduk terhadap akses air minum layak dan persentase penduduk terhadap akses sanitasi layak. Variabel yang berpengaruh signifikan pada model log adalah persentase posyandu aktif, persentase penduduk terhadap akses air minum layak dan persentase penduduk terhadap akses sanitasi layak. Pada model logit variabel yang berpengaruh adalah persentase posyandu aktif.*

***Kata kunci: Data Count, Jumlah Kematian Bayi, Sulawesi Barat, Zero Inflated Poisson Regression***

*(Halaman ini sengaja dikosongkan)*

# MODELLING OF INFANT MORTALITY IN WEST SULAWESI USING ZERO INFLATED POISSON REGRESSION METHOD

**Name** : Rochmatun Annisa  
**Student Number** : 062115 4000 0052  
**Department** : Statistics  
**Supervisors** : Erma Oktania P., S.Si., M.Si.  
Dr. Agnes Tuti Rumiati, M.Sc.

## Abstract

*The number of infant deaths, one of which can be reflected in the Infant Mortality Rate (IMR) is one indicator to determine the development of a country's development. The AKB target in the 2015-2019 Rencana Pembangunan Jangka Menengah Nasional (RPJMN) is 24 deaths per 1,000 live births. Based on the SDKI 2017, the IMR value in West Sulawesi is 42 deaths per 1,000 live births. Factors suspected of affecting infant mortality in West Sulawesi are not only in the health but also socio-economic fields. Data on the number of infant mortality in West Sulawesi is a data count that contains a value of 0. Therefore, the analysis used is Zero Inflated Poisson (ZIP) Regression to determine that factors influence infant mortality in West Sulawesi. The results of the ZIP regression analysis showed that the best model with the smallest AIC and AICc is 324,3407 and 326,7407 were models with a combination of the variables of the percentage of active posyandu, the percentage of residents to access to safe drinking water and the percentage of residents to access to proper sanitation. The variable that has a significant effect on the log model is the percentage of active posyandu, the percentage of the population to access to safe drinking water and the percentage of the population to access to proper sanitation. In the variable logit model that is influential is the percentage of active posyandu.*

**Keywords:** *Data Count, Number of Infant Mortality, West Sulawesi, Zero Inflated Poisson Regression.*

*(Halaman ini sengaja dikosongkan)*

## KATA PENGANTAR

Puji syukur penulis panjatkan atas rahmat dan hidayah yang diberikan Allah SWT sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul “**Pemodelan Jumlah Kematian Bayi di Sulawesi Barat Menggunakan Metode *Zero Inflated Poisson Regression***” dengan lancar. Sholawat serta salam semoga tetap tercurahkan kepada junjungan Nabi Muhammad SAW beserta keluarga dan para sahabat. Penulis menyadari bahwa Tugas Akhir ini dapat terselesaikan tidak terlepas dari bantuan dan dukungan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Kedua orang tua yaitu bapak Wito Al Abdul Chalim dan ibu Nurul Sulaimah. Kedua adik Salsabila Nur Halimah dan Yurida Salma Nasibah serta keluarga lainnya atas segala do'a, nasehat, kasih sayang, dan dukungan yang diberikan kepada penulis demi kesuksesan dan kebahagiaan penulis.
2. Dr. Suhartono selaku Kepala Departemen Statistika dan Santi Wulan P. M.Si, Ph.D selaku Ketua Program Studi Sarjana yang telah memberikan fasilitas, sarana, dan prasarana.
3. Dr. Vita Ratnasari S.Si., M.Si selaku dosen wali selama masa studi yang telah banyak memberikan saran dan arahan dalam proses belajar di Departemen Statistika.
4. Erma Oktania Permatasari, S.Si., M.Si. dan Dr. Agnes Tuti Rumiati, M.Sc selaku dosen pembimbing yang telah meluangkan waktu dan dengan sangat sabar memberikan bimbingan, saran, dukungan serta motivasi selama penyusunan Tugas Akhir.
5. Dr. Puhadi, M.Sc dan Dr. Vita Ratnasari S.Si., M.Si selaku dosen penguji yang selalu sabar dalam mengomentari serta memberikan masukan dan saran dalam penyelesaian Tugas Akhir.
6. Keluarga besar Departemen Statistika ITS, segenap dosen pengajar, staf dan karyawan yang telah memberikan ilmu

dan pengetahuan yang tak ternilai harganya, serta segenap karyawan Departemen Statistika ITS.

7. Teman-teman Statistika ITS  $\Sigma$ 26 angkatan 2015, yang selalu memberikan dukungan kepada penulis selama ini.
8. Achmad Choirul Anam, Nina Restu Nur Ariyanti, Nur'aini Fidaris Salamah, Nur Anisa dan Zamrudi Indra Husada yang selalu mendukung, menyemangati dan mendoakan.
9. Teman-teman BPH BEM ITS Kolaborasi yaitu Muhammad Luthfi Hardiawan, Dadang dan Atikah S. Ishlah selaku presiden BEM, sekretaris jenderal dan sekretaris kabinet 2, BEM ITS Gelora Aksi, BEM ITS Wahana Juang, BSO IECC Wahana Juang dan asrama mutiara 7 yang selalu mendukung dan mendoakan.
10. Semua teman, relasi dan berbagai pihak yang tidak bisa penulis sebutkan namanya satu persatu yang telah membantu dalam penulisan laporan ini.

Besar harapan penulis untuk mendapatkan kritik dan saran yang membangun sehingga Tugas Akhir ini dapat memberikan manfaat bagi semua pihak yang terkait.

Surabaya, Juli 2019

Penulis

## DAFTAR ISI

	Halaman
<b>LEMBAR PENGESAHAN</b> .....	iii
<b>PEMODELAN JUMLAH KEMATIAN BAYI MENGUNAKAN METODE <i>ZERO INFLATED POISSON REGRESSION</i></b> .....	v
<b>MODELLING OF INFANT MORTALITY IN WEST SULAWESI USING ZERO INFLATED POISSON REGRESSION METHOD</b> .....	vii
<b>KATA PENGANTAR</b> .....	ix
<b>DAFTAR ISI</b> .....	xi
<b>DAFTAR GAMBAR</b> .....	xiii
<b>DAFTAR TABEL</b> .....	xv
<b>DAFTAR LAMPIRAN</b> .....	xvii
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	3
1.5 Batasan Masalah.....	4
<b>BAB II TINJAUAN PUSTAKA</b> .....	5
2.1 Kematian Bayi.....	5
2.2 Statistika Deskriptif.....	10
2.3 Asumsi Multikolinearitas.....	11
2.4 Distribusi Poisson.....	12
2.5 Regresi Poisson.....	14
2.6 Overdispersi.....	15
2.7 <i>Zero Inflation</i> .....	16
2.8 Model <i>Zero Inflated Poisson Regression</i> .....	16
2.8.1 Estimasi Parameter Model <i>Zero Inflated Poisson Regression</i> .....	19

2.8.2	Pengujian Parameter Model <i>Zero Inflated Poisson Regression</i> .....	25
2.9	Pemilihan Model Terbaik .....	27
<b>BAB III</b>	<b>METODOLOGI PENELITIAN</b> .....	<b>29</b>
3.1	Sumber Data .....	29
3.2	Variabel Penelitian .....	29
3.3	Struktur Data .....	31
3.4	Langkah Analisis .....	31
<b>BAB IV</b>	<b>ANALISIS DAN PEMBAHASAN</b> .....	<b>33</b>
4.1	Eksplorasi Jumlah Kematian Bayi dan Faktor-faktor yang Diduga Mempengaruhi Kematian Bayi di Sulawesi Barat .....	33
4.2	Pengujian Distribusi Poisson pada Data Jumlah Kematian Bayi di Sulawesi Barat.....	39
4.3	Pengujian Overdispersi pada Data Jumlah Kematian Bayi di Sulawesi Barat .....	40
4.4	Pengujian Asumsi Multikolinearitas pada Variabel Prediktor yang Diduga Mempengaruhi Kematian Bayi di Sulawesi Barat.....	41
4.5	Model <i>Zero Inflated Poisson (ZIP) Regression</i> pada Data Jumlah Kematian Bayi di Sulawesi Barat .....	42
4.6	Model Terbaik pada Data Jumlah Kematian Bayi di Sulawesi Barat .....	45
<b>BAB V</b>	<b>KESIMPULAN DAN SARAN</b> .....	<b>51</b>
5.1	Kesimpulan.....	51
5.2	Saran .....	52
<b>DAFTAR PUSTAKA</b>	.....	<b>53</b>
<b>LAMPIRAN</b>	.....	<b>57</b>



## DAFTAR GAMBAR

	Halaman
<b>Gambar 2.1</b> Konsep Determinan Sosial Ekonomi.....	7
<b>Gambar 2.2</b> Faktor-faktor Penyebab Kematian Bayi .....	8
<b>Gambar 2.3</b> Kerangka Teori Variabel Penelitian .....	9
<b>Gambar 3.1</b> Diagram Alir Langkah Analisis.....	32
<b>Gambar 4.1</b> Diagram Batang Jumlah Kematian Bayi di Sulawesi Barat .....	34
<b>Gambar 4.2</b> Grafik ln (y) dengan $X_1$ .....	49
<b>Gambar 4.3</b> Grafik ln (y) dengan $X_5$ .....	50
<b>Gambar 4.4</b> Grafik ln (y) dengan $X_6$ .....	50

*(Halaman ini sengaja dikosongkan)*

## DAFTAR TABEL

	Halaman
<b>Tabel 3.1</b> Variabel Penelitian.....	29
<b>Tabel 3.2</b> Struktur Data .....	31
<b>Tabel 4.1</b> Variabel Prediktor Kecamatan Pasangkayu .....	35
<b>Tabel 4.2</b> Variabel Prediktor Kecamatan Banggae .....	35
<b>Tabel 4.3</b> Variabel Prediktor 16 Kecamatan .....	35
<b>Tabel 4.4</b> Statistika Deskriptif Faktor-Faktor yang Diduga Mempengaruhi Jumlah Kematian Bayi di Sulawesi Barat .....	36
<b>Tabel 4.5</b> Pengujian Distribusi Poisson pada Data Jumlah Kematian Bayi di Sulawesi Barat .....	39
<b>Tabel 4.6</b> Pengujian Distribusi Data Menggunakan <i>Software</i> <i>Easy Fit</i> .....	39
<b>Tabel 4.7</b> Perbandingan Nilai <i>Mean</i> dan Varians Data Jumlah Kematian Bayi .....	40
<b>Tabel 4.8</b> Pengujian Overdispersi pada data Jumlah Kematian Bayi di Sulawesi Barat .....	40
<b>Tabel 4.9</b> Nilai Koefisien Korelasi Faktor-Faktor yang Diduga Mempengaruhi Jumlah Kematian Bayi di Sulawesi Barat .....	41
<b>Tabel 4.10</b> Nilai VIF Faktor-Faktor yang Diduga Mempengaruhi Jumlah Kematian Bayi di Sulawesi Barat .....	41
<b>Tabel 4.11</b> Hasil Estimasi Parameter Model ZIP <i>Regression</i> .....	42
<b>Tabel 4.12</b> Nilai AIC pada kombinasi variabel.....	45
<b>Tabel 4.13</b> Hasil Estimasi Parameter Model ZIP <i>Regression</i> .....	46

*(Halaman ini sengaja dikosongkan)*

## DAFTAR LAMPIRAN

	Halaman
<b>Lampiran 1.</b> Data Kematian Bayi dan Faktor-Faktor yang Diduga Mempengaruhi.....	57
<b>Lampiran 2.</b> Output Statistika Deskriptif.....	62
<b>Lampiran 3.</b> Output Pengujian Distribusi Poisson.....	62
<b>Lampiran 4.</b> Output Pengujian Overdispersi.....	64
<b>Lampiran 5.</b> Output Pengujian Asumsi Multikolinearitas.....	65
<b>Lampiran 6.</b> Syntax <i>Zero Inflated Poisson Regression</i> .....	73
<b>Lampiran 7.</b> Output <i>Zero Inflated Poisson Regression</i> .....	74
<b>Lampiran 8.</b> Output AIC dan AICc.....	76
<b>Lampiran 9.</b> Output Model Terbaik.....	79
<b>Lampiran 10.</b> Surat Pernyataan Keaslian Data.....	81

*(Halaman ini sengaja dikosongkan)*

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Pembangunan kesehatan merupakan salah satu bagian yang tak terpisahkan dalam pembangunan sebuah negara yang terus diupayakan oleh pemerintah. Salah satu indikator dari dampak pembangunan kesehatan dapat dilihat dari Angka Kematian Bayi (AKB). AKB adalah angka yang menunjukkan banyaknya kematian bayi usia 0 sampai 1 tahun dari setiap 1.000 kelahiran hidup pada tahun tertentu atau dapat dikatakan juga sebagai probabilitas bayi meninggal sebelum mencapai usia satu tahun (BPS, 2019). Selain dapat digunakan sebagai indikator untuk mengetahui status kesehatan anak dan penduduk secara keseluruhan, AKB juga dapat digunakan untuk mengetahui kondisi sosial ekonomi penduduk tersebut.

Berdasarkan data Survei Demografi dan Kesehatan Indonesia (SDKI) 2012, AKB nasional mencapai 32 kematian per 1.000 kelahiran hidup. Oleh karena itu, salah satu sasaran pokok pembangunan nasional dalam Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2015-2019 adalah meningkatnya status kesehatan dan gizi masyarakat dengan ditargetkannya AKB sebesar 24 kematian per 1.000 kelahiran hidup. Upaya pemerintah dalam rangka penurunan AKB nasional cukup berhasil, hal tersebut dibuktikan nilai AKB nasional berdasarkan data SDKI 2017 adalah 24 kematian per 1.000 kelahiran hidup. Namun, nilai AKB di beberapa provinsi di Indonesia masih jauh dari AKB nasional. Daerah-daerah tersebut terutama daerah di Indonesia Timur seperti Gorontalo, Papua Barat, Maluku Utara, Sulawesi Utara dan Sulawesi Barat.

Sulawesi Barat merupakan salah satu provinsi dengan jumlah AKB tertinggi di Indonesia. Berdasarkan data Survei Demografi dan Kesehatan Indonesia 2017, AKB di Sulawesi Barat mencapai 42 kematian per 1.000 kelahiran hidup. Angka tersebut masih sangat jauh dari target yang ditetapkan pemerintah dalam

RPJMN 2015-2019. Salah satu faktor yang diduga mempengaruhi angka kematian bayi adalah faktor sosial ekonomi. Hal ini dapat dilihat dari masih terdapatnya kesenjangan kematian bayi di tingkat pendidikan, sosial ekonomi, antar perkotaan dan pedesaan.

Penelitian terdahulu yang pernah dilakukan tentang kematian bayi adalah Handayani & Pratiwi (2016). Pada penelitian tersebut dilakukan analisis mengenai kebiasaan persalinan pada suku Kaili Da'a di Kabupaten Mamuju Sulawesi Barat. Masyarakat suku Kaili Da'a masih banyak yang melahirkan di rumah dengan dibantu keluarga atau dukun. Hal tersebut membuat kematian bayi di Sulawesi Barat meningkat. Penelitian lain yang membahas tentang faktor-faktor yang mempengaruhi jumlah kematian ibu dan bayi di Jawa Tengah menggunakan metode *Bivariate Generalized Poisson Regression* dilakukan oleh Putri & Purhadi (2017). Hasil yang diperoleh yaitu variabel yang berpengaruh signifikan terhadap jumlah kematian ibu dan bayi adalah persentase persalinan oleh tenaga kesehatan, persentase ibu hamil mendapatkan tablet Fe<sub>3</sub>, persentase komplikasi kebidanan yang ditangani, persentase rumah tangga ber-PHBS dan persentase ibu hamil melaksanakan program K4. Penelitian tentang AKB di Jawa Timur dilakukan oleh Suryani (2018) yaitu menganalisis faktor-faktor yang mempengaruhi jumlah kematian bayi dan ibu dengan metode *Geographically Weighted Bivariate Generalized Poisson Regression*. Pada penelitian tersebut memperhatikan faktor wilayah pada masing-masing unit observasi. Hasil dari penelitian didapatkan variabel yang signifikan yaitu persentase pelayanan kunjungan ibu hamil dengan K4, ibu hamil mendapatkan tablet Fe<sub>3</sub>, komplikasi kebidanan yang ditangani, rumah tangga yang ber-PHBS serta wanita kawin dengan umur perkawinan pertama dibawah 18 tahun.

Data jumlah kematian bayi merupakan data diskrit (*count*) yang mengikuti distribusi *Poisson*. Oleh karena itu, untuk memodelkan data jumlah kematian bayi di Sulawesi Barat dapat menggunakan regresi *Poisson*. Asumsi dalam regresi *Poisson* yang harus dipenuhi adalah equidispersi. Equidispersi merupakan



kesamaan nilai *mean* dan varians pada data. Namun, pada kenyataannya data jumlah kematian bayi di Sulawesi Barat tidak memenuhi asumsi tersebut. Pada data jumlah kematian bayi Sulawesi Barat terjadi overdispersi yaitu nilai varians yang lebih besar daripada nilai *mean* karena terdapat data yang bernilai nol. Pada kasus yang terjadi overdispersi dan adanya data nol, maka model *Zero Inflated Poisson (ZIP) regression* lebih disarankan (Lambert, 1992). Oleh karena itu, pada penelitian ini menganalisis jumlah kematian bayi di Sulawesi Barat dengan menggunakan pendekatan *Zero Inflated Poisson regression* karena diduga terjadi kasus overdispersi. Variabel respon yang digunakan adalah jumlah kematian bayi di Sulawesi Barat serta variabel prediktor merupakan faktor-faktor yang diduga mempengaruhi jumlah kematian bayi di Sulawesi Barat.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang telah dijelaskan, permasalahan yang dibahas dalam penelitian ini adalah faktor-faktor apa saja yang mempengaruhi jumlah kematian bayi kelahiran di Sulawesi Barat menggunakan metode *Zero Inflated Poisson Regression*.

## **1.3 Tujuan Penelitian**

Berdasarkan rumusan masalah yang telah diuraikan di atas, tujuan yang ingin dicapai dalam penelitian ini adalah:

1. Mendeskripsikan karakteristik faktor-faktor yang diduga mempengaruhi jumlah kematian bayi di Sulawesi Barat.
2. Mendapatkan faktor-faktor yang mempengaruhi jumlah kematian bayi di Sulawesi Barat menggunakan metode *Zero Inflated Poisson Regression*.

## **1.4 Manfaat Penelitian**

Pada penelitian ini diharapkan dapat memberikan informasi kepada Dinas Kesehatan Provinsi Sulawesi Barat dalam upaya penurunan jumlah kematian bayi dan membantu dalam perencanaan program preventif jumlah kematian bayi di Sulawesi Barat berdasarkan faktor-faktor yang berpengaruh berdasarkan

hasil penelitian ini. Disamping itu penelitian ini dapat menambah wawasan bagi peneliti dalam pembangunan dan penerapan ilmu statistika dalam fokus bidang sosial kependudukan.

### **1.5 Batasan Masalah**

Batasan masalah yang diterapkan dalam penelitian ini adalah variabel yang digunakan bersumber dari data profil kesehatan seluruh kabupaten di Provinsi Sulawesi Barat tahun 2016 dengan unit penelitiannya adalah kecamatan.

## **BAB II**

### **TINJAUAN PUSTAKA**

Bab ini membahas mengenai distribusi Poisson, overdispersi, asumsi multikolinearitas, overdispersi, *zero inflated poisson regression*, *Akaike's Information Criterion (AIC)* dan *Akaike's Information Corected Criterion (AIC<sub>c</sub>)*.

#### **2.1 Kematian Bayi**

Kematian bayi merupakan kematian individu yang berusia kurang dari 1 tahun. Kematian bayi salah satunya dapat dicerminkan dalam Angka Kematian Bayi (AKB). AKB adalah angka yang menunjukkan banyaknya kematian bayi usia 0 sampai 1 tahun dari setiap 1.000 kelahiran hidup pada tahun tertentu atau dapat dikatakan juga sebagai probabilitas bayi meninggal sebelum mencapai usia satu tahun. AKB merupakan indikator yang penting untuk mencerminkan keadaan derajat kesehatan di suatu masyarakat, karena bayi yang baru lahir sangat sensitif terhadap keadaan lingkungan tempat orang tua bayi tersebut tinggal dan sangat erat kaitannya dengan status sosial orang tua. Kemajuan yang dicapai dalam bidang pencegahan dan pemberantasan berbagai penyakit penyebab kematian akan tercermin secara jelas dengan menurunnya tingkat AKB. Dengan demikian AKB merupakan tolok ukur yang sensitif dari semua upaya intervensi yang dilakukan oleh pemerintah khususnya di bidang kesehatan (Badan Pusat Statistik, 2019).

Kematian bayi disebabkan oleh banyak faktor yang mempengaruhinya. Secara garis besar dari sisi peyebabnya, kematian bayi ada dua macam yaitu endogen dan eksogen, kematian bayi endogen atau yang umum disebut dengan kematian neonatal adalah kematian bayi yang terjadi pada bulan pertama setelah dilahirkan, dan umumnya disebabkan oleh faktor-faktor yang dibawa anak sejak lahir, yang diperoleh dari orangtuanya pada saat konsepsi atau didapat selama kehamilan. Kematian bayi eksogen atau kematian bayi post neonatal, adalah kematian bayi yang terjadi setelah usia satu bulan sampai menjelang usia satu

tahun yang disebabkan oleh faktor-faktor yang bertalian dengan pengaruh lingkungan luar (Kementrian Kesehatan Republik Indonesia, 2015).

Tinggi rendahnya kematian bayi sangat dipengaruhi oleh beberapa faktor yaitu (Badan Perencanaan Pembangunan Nasional, 2009):

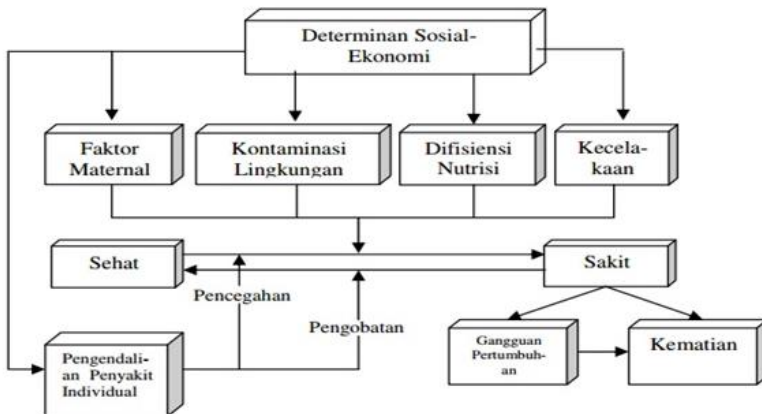
1. Persalinan dengan tenaga non medis, kejadian komplikasi pada ibu dan bayi baru lahir sebagian besar terjadi pada masa sekitar persalinan sehingga pemeriksaan kesehatan pada saat hamil dan kehadiran serta pertolongan tenaga kesehatan yang terampil pada masa persalinan sangat penting.
2. Semakin banyak wanita yang berumah tangga di bawah umur 18 tahun, semakin banyak bayi yang rentan terhadap segala penyakit dan gangguan lain karena ketidakpastian ibu.
3. Kurangnya kesadaran akan pentingnya ASI, bayi yang tidak diberikan ASI lebih mudah terserang penyakit daripada bayi yang diberi ASI, karena pemberian ASI pada bayi sangat berpengaruh dalam kekebalan terhadap penyakit.
4. Tingkat pendidikan wanita, semakin tinggi tingkat pendidikan wanita maka kesadaran terhadap kesehatan semakin tinggi sehingga perawatan bayi akan semakin baik. Tingkat pendidikan ibu memiliki korelasi kuat dengan tingkat kematian anak, AKB pada penduduk yang tidak berpendidikan tiga kali lipat lebih besar dibandingkan dengan yang berpendidikan tinggi.
5. Faktor sosial ekonomi, AKB pada tingkat sosial ekonomi rendah ditaksirkan lebih besar dibandingkan dengan tingkat sosial ekonomi tinggi.

Mosley dan Chen (1984), membagi variabel-variabel yang berpengaruh terhadap kelangsungan hidup anak menjadi dua, yaitu:

1. Variabel yang dianggap eksogeneus atau sosial ekonomi seperti budaya, sosial, ekonomi, masyarakat dan faktor regional. Kematian bayi dipengaruhi oleh faktor sosial ekonomi yang secara tidak langsung terbagi dalam lima faktor utama:
  - a. Faktor maternal : umur, paritas, dan jarak kelahiran.

- b. Kontaminasi lingkungan : pencemaran lingkungan berkaitan dengan penularan kepada anak (dan ibu).
- c. Difisiensi nutrisi : kelangsungan hidup anak tidak hanya dipengaruhi oleh tersedianya gizi bagi anak melainkan bagi ibu juga.
- d. Kecelakaan: kecelakaan ini meliputi kecelakaan fisik, contoh nyata adalah pembunuhan bayi,
- e. Pengendalian penyakit individu: salah satu komponen dalam pengendalian penyakit perorangan adalah tindakan preventif yang diambil oleh orang sehat untuk mencegah penyakit.

Kelima variabel diatas dirangkum dalam Gambar 2.1 berikut ini.

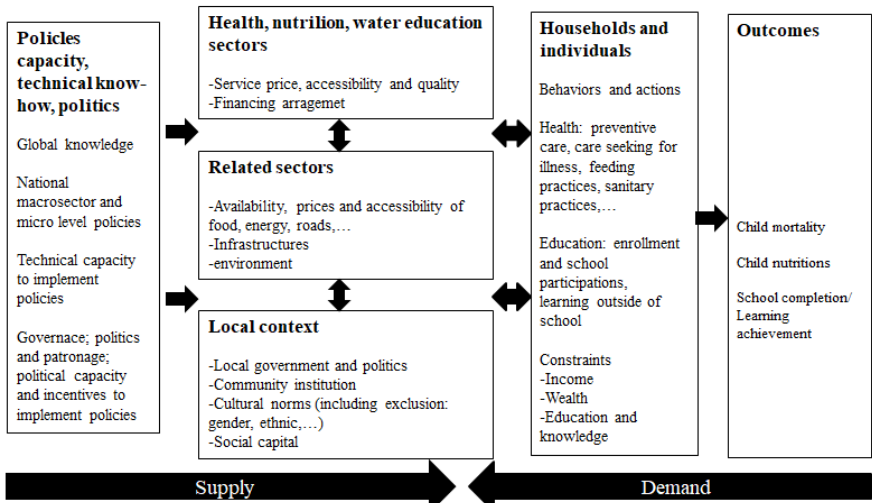


**Gambar 2.1** Konsep Determinan Sosial Ekonomi

2. Variabel endogeneous atau faktor *biomedical* seperti pola pemberian ASI, kebersihan, sanitasi dan nutrisi.
  - A. Tingkat individu: produktivitas rumah tangga yang meliputi pendidikan, kesehatan dan waktu, dan tradisi/norma/sikap dalam lingkungan.
  - B. Tingkat rumah tangga: efek pendapatan/kekayaan.
  - C. Tingkat wilayah: lingkungan ekologi, ekonomi politik, sistem kesehatan.

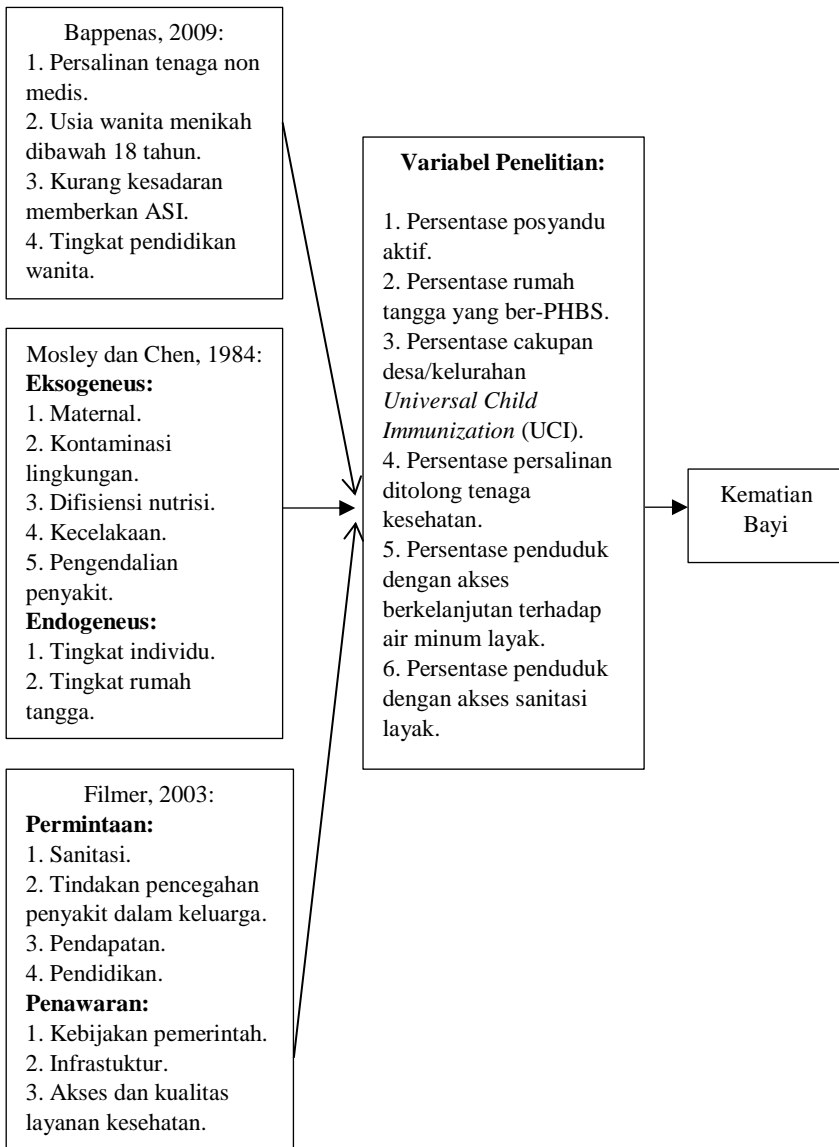
Menurut Filmer (2003), faktor-faktor penyebab kematian anak ditunjukkan pada teori Gambar 2.2. Kematian anak yang meliputi

kematian bayi dipengaruhi oleh sisi permintaan dan sisi penawaran. Sisi permintaan adalah perilaku atau karakteristik rumah tangga dan individual seperti sanitasi, tindakan pencegahan penyakit dalam keluarga, pendapatan, pendidikan dan pengetahuan orang tua. Dari sisi penawaran adalah kebijakan pemerintah di tingkat mikro maupun makro sekaligus implementasi kebijakannya, kapabilitas dari pemerintah daerah, dan infrastruktur serta akses dan kualitas layanan kesehatan.



**Gambar 2.2** Faktor-faktor Penyebab Kematian Bayi Menurut Filmer (2003)

Berdasarkan teori-teori penyebab kematian dari Bappenas, Mosley & Chen, serta Filmer, maka kerangka konsep variabel penelitian yang digunakan terdapat pada Gambar 2.3.



**Gambar 2.3** Kerangka Teori Variabel Penelitian

Pemilihan variabel penelitian didasarkan pada beberapa teori dan penelitian yang telah dijelaskan sebelumnya. Variabel persentase posyandu aktif diambil berdasarkan teori Filmer (2003) yaitu pada akses dan kualitas layanan kesehatan. Posyandu merupakan fasilitas layanan kesehatan seperti penyediaan layanan imunisasi untuk bayi. Variabel persentase rumah tangga yang ber-PHBS dipilih berdasarkan teori Mosley dan Chen (1984) yaitu pada penyebab kematian bayi dari sisi endogenous pada tingkat rumah tangga. Pengambilan variabel persentase cakupan desa/kelurahan UCI didasarkan pada teori Filmer (2003) yaitu pada kebijakan pemerintah. Menurut Kementerian Kesehatan menargetkan pada tahun 2014 seluruh desa/ kelurahan mencapai 100% UCI (Universal Child Immunization) atau 90% dari seluruh bayi di desa/ kelurahan tersebut memperoleh imunisasi dasar lengkap yang terdiri dari BCG, Hepatitis B, DPT-HB, Polio dan campak. Pengambilan variabel persentase persalinan ditolong tenaga kesehatan didasarkan pada penelitian Bappenas (2009) yaitu dari variabel persalinan tenaga non medis. Berdasarkan teori Mosley dan Chen (1984) pada penyebab eksogeneus yaitu kontaminasi lingkungan maka variabel persentase penduduk dengan akses berkelanjutan terhadap air minum layak bukan jaringan perpipaan dipilih untuk dianalisis apakah berpengaruh signifikan terhadap jumlah kematian bayi. Variabel persentase penduduk dengan akses sanitasi layak dipilih berdasarkan teori Filmer (2003) pada sisi permintaan yaitu sanitasi.

## **2.2 Statistika Deskriptif**

Metode statistik adalah prosedur-prosedur yang digunakan dalam pengumpulan, penyajian, analisis, dan penafsiran data. Metode-metode tersebut terbagi menjadi dua yaitu statistika deskriptif dan statistika inferensia. Statistika deskriptif adalah metode yang berkaitan dengan pengumpulan dan penyajian data sehingga memberikan informasi hanya mengenai data dan tidak menarik kesimpulan (Walpole *et al*, 2007). Pada penelitian ini akan digunakan statistika deskriptif berupa nilai minimum, nilai maksimum, rata-rata dan varians.



### Rata-rata

Ukuran yang menunjukkan letak nilai tersebut berada, di pusat atau di lokasi yang lainnya. Rumus menghitung rata-rata ditunjukkan pada persamaan 2.1.

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (2.1)$$

### Varians

Varians menunjukkan ukuran keberagaman data. Persamaan 2.2 merupakan rumus menghitung varians.

$$s^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1} \quad (2.2)$$

dengan,

- $\bar{x}$  : rata-rata
- $s^2$  : varians
- $x_i$  : pengamatan ke-i
- $n$  : banyaknya pengamatan.

### Minimum

Minimum adalah nilai terkecil pada sekumpulan data.

### Maksimum

Maksimum adalah nilai terbesar pada sekumpulan data.

## 2.3 Asumsi Multikolinearitas

Istilah multikolinearitas (kolinearitas ganda) pertama kali ditemukan oleh Ragnar Frisch, yang berarti adanya hubungan linear yang sempurna atau pasti di antara beberapa atau semua variabel penjelas (bebas) dari model regresi ganda. Selanjutnya istilah multikolinearitas digunakan dalam arti yang lebih luas, yaitu untuk terjadinya korelasi linear yang tinggi di antara variabel-variabel penjelas  $(X_1, X_2, \dots, X_k)$  (Setiawan & Kusriani, 2010). Adanya kasus multikolinearitas dapat mengakibatkan hasil estimasi parameter menjadi tidak akurat. Hal ini dikarenakan *standart error* yang besar dari hasil estimasi parameter dengan signifikansi yang kecil, bahkan menjadi

tidak signifikan pada pengujian individu namun sangat signifikan pada pengujian serentak.

Cara mendeteksi kasus multikolinearitas apabila  $R^2$  yang tinggi dalam model, namun sedikit atau tidak satupun parameter regresi yang signifikan jika diuji secara individu. Apabila dalam model regresi memperoleh koefisien regresi ( $\hat{\beta}_j$ ) dengan tanda yang berbeda dengan koefisien korelasi antara Y dengan  $X_j$ . Misalnya, korelasi antara Y dengan  $X_j$  bertanda positif ( $r_{YX_j} > 0$ ), tetapi koefisien regresi yang berhubungan dengan  $X_j$  bertanda negatif ( $\hat{\beta}_j < 0$ ) atau sebaliknya.

Menurut Gujarati (2004) salah satu cara mengidentifikasi ada atau tidaknya kasus multikolinearitas yaitu melalui nilai *Variance-inflating Factor* (VIF). Nilai VIF menunjukkan bagaimana variasi dari hasil estimasi parameter meningkat karena adanya kasus multikolinearitas. Nilai VIF yang lebih dari 10 menunjukkan adanya kasus multikolinearitas. Nilai VIF dirumuskan pada persamaan 2.3.

$$\text{VIF}_j = \frac{1}{1 - R_j^2} \quad (2.3)$$

dengan  $R_j^2$  merupakan koefisien determinasi antara  $X_j$  dengan variabel prediktor lainnya.

Cara mengatasi kasus multikolinearitas adalah melakukan transformasi variabel-variabel, mengeluarkan satu variabel yang memiliki korelasi tinggi terhadap variabel lain, menambah data baru dan menggunakan metode lain seperti regresi komponen utama (*Principal Component Regression*), regresi *ridge*, regresi kuadrat terkecil parsial (*Partial Least Squares Regression*).

## 2.4 Distribusi Poisson

Distribusi *Poisson* adalah suatu distribusi yang digunakan dalam peristiwa yang terjadi dalam suatu interval waktu atau wilayah tertentu. Distribusi *Poisson* adalah suatu distribusi untuk peristiwa yang

memiliki peluang kejadian kecil dan berupa variabel diskrit (Walpole *et al*, 2007).

Fungsi distribusi peluang untuk variabel random *Poisson*  $Y$  dengan parameter  $\lambda$  ditunjukkan pada persamaan 2.4.

$$P(Y_i = y_i) = \begin{cases} \frac{e^{-\lambda} \lambda^{y_i}}{y_i!}, & \text{untuk } y_i = 0, 1, 2, \dots \\ 0, & \text{untuk } y_i \text{ lain} \end{cases} \quad (2.4)$$

$$E[Y_i] = \lambda_i, \text{ var}(Y_i) = \lambda_i$$

Kesamaan *mean* dan varians pada distribusi *Poisson* disebut dengan *equidispersi*. Namun, dalam data riil pernyataan tersebut sering dilanggar. Pelanggaran yang terjadi yaitu *overdispersi* atau *underdispersi* (Cameron & Trivedi, 2013).

Uji *Komolgorov Smirnov* dilakukan untuk mengetahui suatu data berdistribusi *Poisson* atau tidak. Hipotesis pengujiannya adalah sebagai berikut:

$H_0$  : data berdistribusi *Poisson*

$H_1$  : data tidak berdistribusi *Poisson*

Statistik uji yang digunakan dalam pengujian hipotesis ini adalah statistik uji  $D$  yang merupakan nilai maksimum dari harga mutlak perbedaan antara distribusi frekuensi kumulatif yang dihitung dari data sampel ( $S_n(y)$ ) dengan fungsi distribusi kumulatif yang dihipotesiskan yaitu *Poisson* ( $F_0(y)$ ),

$$D = \text{maksimum} |F_0(y) - S_n(y)| \quad (2.5)$$

Hipotesis nol yang akan ditolak jika  $D > D_{(\alpha, n)}$ , dengan  $n$  adalah ukuran sampel dan  $D_{(\alpha, n)}$  adalah nilai kritis  $D$  uji *Kolmogorov Smirnov* dengan tingkat signifikansi atau  $\alpha$  pada tabel *Kolmogorov Smirnov*.

## 2.5 Regresi Poisson

Model standar untuk data *count* adalah model regresi *Poisson*, dimana model tersebut merupakan model regresi nonlinear. Model regresi ini diturunkan dari distribusi *Poisson* dengan parameter  $\lambda$  bergantung pada kovariat (regresi) (Cameron & Trivedi, 2013). Parameter  $\lambda$  sangat bergantung pada beberapa unit tertentu atau periode dari waktu, jarak, luas area, volume, dan sebagainya. Distribusi ini kemudian digunakan untuk memodelkan suatu peristiwa yang keberadaannya relatif jarang atau langka untuk terjadi pada satuan unit tertentu.

Analisis data *count* untuk peristiwa yang jarang terjadi dilakukan dengan regresi *Poisson*. Banyaknya peristiwa tersebut sering tergantung pada variabel-variabel prediktor  $x_1, x_2, \dots, x_k$ , saat pengambilan data dilakukan. Bentuk matriks dari variabel respon, variabel-variabel prediktor, dan parameter regresi *Poisson* dapat ditulis sesuai persamaan 2.4.

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \quad \text{dan} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad (2.6)$$

*Generalized Linier Model* standar untuk data *count* adalah model regresi *Poisson* dengan *link function*  $\ln$ . *Mean* dari distribusi *Poisson* yang merupakan fungsi dari variabel-variabel prediktornya ditunjukkan oleh regresi *Poisson*. Maka model regresi *Poisson* dapat ditulis berdasarkan persamaan 2.6 (Lestari, 2008).

$$\lambda_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \quad (2.7)$$

dengan,

$\lambda_i$  : rata-rata jumlah peristiwa yang terjadi pada periode waktu tertentu observasi ke- $i$

$\mathbf{x}_i$  : vektor dari variabel prediktor, dengan

$$\mathbf{x}_i^T = [1 \quad x_{i1} \quad x_{i2} \quad \cdots \quad x_{ij}]$$

$\beta$  : vektor dari parameter regresi *Poisson*,

dan *link function*  $g(\cdot) = \ln(\lambda_i)$

dimana:

$$E(Y_i) = \lambda_i$$

$$\ln(\lambda_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_j x_{ij}$$

$$\ln(\lambda_i) = \beta_0 + \sum_{j=1}^k \beta_j x_{ij}; \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, k \quad (2.8)$$

$$\ln(\lambda_i) = \mathbf{x}_i^T \boldsymbol{\beta}$$

$$\lambda_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$$

## 2.6 Overdispersi

Analisis regresi *Poisson* adalah analisis regresi yang termasuk bagian dari *Generalized Linear Model* (GLM). Regresi *Poisson* digunakan untuk data dengan variabel respon yang mengikuti distribusi *Poisson* ( $Y \sim \text{Poisson}$ ). Asumsi ini sangat penting pada analisis tersebut yaitu varians harus sama dengan rata-rata atau dapat disebut equidispersi. Pada data yang riil kondisi tersebut sering tidak dapat terpenuhi yaitu sering ditemukan data *count* (data hitung) yang memiliki nilai varians lebih besar dari rata-rata atau disebut overdispersi. Namun, jika nilai varians kurang dari rata-rata maka disebut underdispersi (Hilbe, 2011).

Menurut Hinde & Demetrio (2007), terdapat kemungkinan tidak terpenuhi equidispersi pada suatu pemodelan, antara lain adalah adanya keragaman hasil pengamatan (keragaman antar individu sebagai komponen yang tidak dijelaskan model), korelasi antar respon individu, terjadi pengelompokkan dalam populasi dan variabel teramat yang dihilangkan. Akibat dari tidak terpenuhinya equidispersi adakah regresi *Poisson* tidak sesuai untuk memodelkan data karena model yang terbentuk akan menghasilkan estimasi parameter yang bias. Selain itu, overdispersi juga mengakibatkan nilai *standart error* menjadi lebih kecil (*underestimates*) dari nilai yang seharusnya, sehingga menghasilkan kesimpulan yang tidak sesuai.

Pemeriksaan overdispersi dapat menggunakan statistik uji *Pearson's Chi-Square* sebagai berikut:

$$\sum \frac{(y_i - \hat{\mu}_i)^2}{\hat{\sigma}_i^2} \sim \chi^2_{(n-p)}; \text{ dengan } p=k+1 \quad (2.9)$$

Data mengalami overdispersi jika statistik uji *Pearson Chi-Square* atau nilai devians dibagi dengan derajat bebas menghasilkan nilai lebih dari 1.

## 2.7 Zero Inflation

Nilai nol yang sangat banyak pada variabel respon (*zero inflation*) sering ditemukan pada analisis regresi *Poisson* baik untuk data diskrit maupun data *count*. Apabila nilai nol memiliki arti penting dalam penelitian, maka data tersebut tidak dapat langsung dihilangkan namun harus dimasukkan dalam proses analisis. Pada beberapa penelitian dapat dijumpai kondisi terlalu banyak nilai nol pada variabel respon yang lebih dari 50%. Menurut Famoye & Singh (2006), besarnya proporsi data yang bernilai nol dapat berakibat pada ketepatan dari inferensia. Selain itu, regresi *Poisson* juga menjadi tidak tepat lagi memodelkan data yang sebenarnya.

## 2.8 Model Zero Inflated Poisson Regression

Seperti yang telah dijelaskan pada subbab 2.4 mengenai regresi *Poisson*, ada asumsi kesamaan *mean* dan varians (*equidispersi*). Namun, pada kasus dimana data dengan variabel respon mengandung nilai nol dalam proporsi besar (*zero inflation*) maka model *Zero Inflated Poisson (ZIP) regression* lebih disarankan (Lambert, 1992).

Model regresi ZIP adalah salah satu alternatif metode untuk menganalisis data dengan banyak nilai nol yang terkandung pada variabel respon. Banyaknya nilai nol pada data dapat mengakibatkan pelanggaran pada asumsi kesamaan *mean* dan varians pada distribusi *Poisson*. Untuk setiap pengamatan pada variabel respon  $Y_1, Y_2, \dots, Y_n$  yang saling bebas dan  $Y$  mengikuti distribusi yang ditunjukkan persamaan 2.10.

$$Y_i \sim \begin{cases} 0 & , \text{ dengan peluang } \pi_i \\ \text{Poisson}(\lambda_i) & , \text{ dengan peluang } (1 - \pi_i) \end{cases} \quad (2.10)$$

Maka fungsi probabilitas untuk  $\mathbf{Y}_i$  ditunjukkan pada persamaan 2.10.

$$P(Y_i = y_i) = \begin{cases} \pi_i + (1 - \pi_i) e^{-\lambda_i}, & \text{ untuk } y_i = 0 \\ \frac{(1 - \pi_i) e^{-\lambda_i} \lambda_i^{y_i}}{y_i!}, & \text{ untuk } y_i > 0 \end{cases} \quad (2.10)$$

dengan parameter  $\boldsymbol{\lambda} = (\lambda_1 \lambda_2 \dots \lambda_n)^T$  dan  $\boldsymbol{\pi} = (\pi_1 \pi_2 \dots \pi_n)^T$  yang memenuhi persamaan 2.10.

$$\boldsymbol{\lambda}_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$$

$$\boldsymbol{\pi}_i = \frac{e^{\mathbf{x}_i^T \boldsymbol{\gamma}}}{1 + e^{\mathbf{x}_i^T \boldsymbol{\gamma}}} \quad \text{dan} \quad (1 - \boldsymbol{\pi}_i) = \frac{1}{1 + e^{\mathbf{x}_i^T \boldsymbol{\gamma}}} \quad (2.11)$$

dengan,

$$\begin{aligned} \log(\boldsymbol{\lambda}) &= \begin{bmatrix} \log \eta_1 \\ \log \eta_2 \\ \vdots \\ \log \eta_n \end{bmatrix} \\ &= \begin{bmatrix} \beta_0 + \beta_1 x_{11} + \dots + \beta_k x_{1k} \\ \beta_0 + \beta_1 x_{21} + \dots + \beta_k x_{2k} \\ \vdots \\ \beta_0 + \beta_1 x_{n1} + \dots + \beta_k x_{nk} \end{bmatrix} \\ &= \begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \\ &= \mathbf{X}\boldsymbol{\beta} \end{aligned} \quad (2.12)$$

dan

$$\begin{aligned}
\text{logit}(\boldsymbol{\pi}) &= \begin{bmatrix} \text{logit } \phi_1 \\ \text{logit } \phi_2 \\ \vdots \\ \text{logit } \phi_n \end{bmatrix} \\
&= \begin{bmatrix} \gamma_0 + \gamma_1 x_{11} + \cdots + \gamma_k x_{1k} \\ \gamma_0 + \gamma_1 x_{21} + \cdots + \gamma_k x_{2k} \\ \vdots \quad \quad \quad \ddots \quad \quad \quad \vdots \\ \gamma_0 + \gamma_1 x_{n1} + \cdots + \gamma_k x_{nk} \end{bmatrix} \\
&= \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_k \end{bmatrix} \\
&= \mathbf{X}\boldsymbol{\gamma}
\end{aligned} \tag{2.13}$$

dimana,

$$\text{logit}(\boldsymbol{\pi}) = \begin{bmatrix} \text{logit } \phi_1 \\ \text{logit } \phi_2 \\ \vdots \\ \text{logit } \phi_n \end{bmatrix} = \begin{bmatrix} \log \frac{\pi_1}{1-\pi_1} \\ \log \frac{\pi_2}{1-\pi_2} \\ \vdots \\ \log \frac{\pi_n}{1-\pi_n} \end{bmatrix} \tag{2.14}$$

Model regresi ZIP dapat dituliskan sesuai dengan persamaan 2.15.

$$\begin{aligned}
\log \lambda_i &= \beta_0 + \beta_1 x_{i1} + \cdots + \beta_j x_{ij} & ; i=1,2,\dots,n; j=1,2,\dots,k \\
\text{logit } \pi_i &= \gamma_0 + \gamma_1 x_{i1} + \cdots + \gamma_j x_{ij} & ; i=1,2,\dots,n; j=1,2,\dots,k
\end{aligned} \tag{2.15}$$

dengan,

$\boldsymbol{\beta}$  : vektor dari parameter regresi yang ditaksir mewakili variabel yang mempengaruhi *Poisson state*.



$\boldsymbol{\gamma}$  : vektor dari parameter regresi yang ditaksir mewakili variabel yang mempengaruhi *zero state*.

$\mathbf{X}$  : matriks variabel yang memuat himpunan-himpunan yang berbeda dari faktor eksperimen yang berhubungan dengan peluang *zero state* dan *mean Poisson* pada *Poisson state*.

### 2.8.1 Estimasi Parameter Model Zero Inflated Poisson Regression

Penaksir parameter ZIP *Regression* digunakan metode *Maksimum Likelihood Estimation* (MLE). Metode ini biasanya digunakan untuk menaksir parameter suatu model yang diketahui fungsi densitasnya. Terlebih dahulu akan dibangun *likelihood* dari fungsi ZIP, dengan mensubstitusi *link function* dari model ZIP *Regression* ke fungsi densitasnya.

$$\begin{aligned}\lambda_i &= e^{\mathbf{x}_i^T \boldsymbol{\beta}} \\ \pi_i &= \frac{e^{\mathbf{x}_i^T \boldsymbol{\gamma}}}{1 + e^{\mathbf{x}_i^T \boldsymbol{\gamma}}} \\ (1 - \pi_i) &= \frac{1}{1 + e^{\mathbf{x}_i^T \boldsymbol{\gamma}}}\end{aligned}\tag{2.16}$$

Persamaan 2.16 disubstitusikan dalam persamaan 2.11, maka menghasilkan persamaan 2.17.

$$P(Y_i = y_i) = \begin{cases} \frac{e^{\mathbf{x}_i^T \boldsymbol{\gamma}} + \left(\exp\left(-e^{\mathbf{x}_i^T \boldsymbol{\beta}}\right)\right)}{1 + e^{\mathbf{x}_i^T \boldsymbol{\gamma}}}, & \text{untuk } y_i = 0 \\ \frac{1}{1 + e^{\mathbf{x}_i^T \boldsymbol{\gamma}}} \left( \frac{\exp\left(\left(-e^{\mathbf{x}_i^T \boldsymbol{\beta}}\right) + \left(\mathbf{x}_i^T \boldsymbol{\beta}\right) y_i\right)}{y_i!} \right), & \text{untuk } y_i > 0 \end{cases}\tag{2.17}$$

Fungsi *likelihood* dari model ZIP *regression* yang terbentuk adalah:

$$L(\gamma, \beta) = \begin{cases} \prod_{i=1}^n \frac{e^{\mathbf{x}_i^T \gamma} + \left( \exp\left(-e^{\mathbf{x}_i^T \beta}\right) \right)}{1 + e^{\mathbf{x}_i^T \gamma}}, & \text{untuk } y_i = 0 \\ \prod_{i=1}^n \frac{1}{1 + e^{\mathbf{x}_i^T \gamma}} \left( \exp\left(\left(-e^{\mathbf{x}_i^T \beta}\right) + \left(\mathbf{x}_i^T \beta\right) y_i\right) \right)}{y_i!}, & \text{untuk } y_i > 0 \end{cases} \quad (2.18)$$

dan fungsi log *likelihood*nya adalah:

$$\log L(\gamma, \beta) = \begin{cases} \sum_{i=1}^n \log\left(e^{\mathbf{x}_i^T \gamma} + e^{-e^{\mathbf{x}_i^T \beta}}\right) - \sum_{i=1}^n \log\left(1 + e^{\mathbf{x}_i^T \gamma}\right), & \text{untuk } y_i = 0 \\ \sum_{i=1}^n \left(-e^{\mathbf{x}_i^T \beta} + \left(\mathbf{x}_i^T \beta\right) y_i\right) - \sum_{i=1}^n \log\left(1 + e^{\mathbf{x}_i^T \gamma}\right) - \sum_{i=1}^n y_i!, & \text{untuk } y_i > 0 \end{cases} \quad (2.19)$$

Persamaan 2.19 disebut sebagai *incomplete likelihood*. Hal ini dikarenakan nilai nol pada suku pertama tidak diketahui mana yang berasal dari *zero state* dan mana yang berasal dari *Poisson state*, sehingga persamaan 2.19 diselesaikan dengan mendefinisikan kembali variabel  $Y_i$  dengan suatu variabel indikator  $Z_i$ , yaitu:

$$Z_i = \begin{cases} 1, & \text{jika } y_i \text{ berasal dari } zero \text{ state} \\ 0, & \text{jika } y_i \text{ berasal dari } Poisson \text{ state} \end{cases} \quad (2.20)$$

Jika  $y_i > 0$ , maka nilai  $Z_i = 0$ . Akan tetapi, jika  $y_i = 0$ , maka  $Z_i$  dapat bernilai 0 atau 1. Oleh karena itu,  $Z_i$  dianggap hilang sebagian. Permasalahan tersebut dapat diselesaikan menggunakan algoritma Ekspektasi-Maksimalisasi (EM). Algoritma EM merupakan salah satu alternatif metode iteratif untuk memaksimalkan fungsi *likelihood* yang mengandung data tidak lengkap (*missing*). Pada setiap iterasinya, algoritma EM memiliki 2 tahap yaitu tahap ekspektasi dan maksimalisasi. Tahap ekspektasi yaitu tahap perhitungan ekspektasi dari fungsi log *likelihood* dengan memperhatikan data yang tidak lengkap. Tahap maksimalisasi merupakan tahap perhitungan untuk mencari penaksir parameter yang memaksimalkan fungsi log *likelihood* hasil dari tahap ekspektasi.

Untuk mengestimasi parameter  $Z_i$  yang telah didefinisikan pada persamaan 2.18 dengan algoritma EM, berikut ini adalah langkah-langkahnya:

1. Menentukan terlebih dahulu distribusi dari variabel  $Z_i$ , yaitu:  
Berdasarkan persamaan 2.10 dan 2.20, maka:

$$\begin{aligned} P(Z_i=1) &= \pi_i \\ P(Z_i=0) &= P(Y_i \sim \text{Poisson } \lambda_i) = 1 - \pi_i \end{aligned}$$

dapat ditulis,

$$P(Z_i = z_i) = \begin{cases} \pi_i & , \text{ jika } z_i = 1 \\ 1 - \pi_i & , \text{ jika } z_i = 0 \end{cases} \quad (2.21)$$

Peluang pada saat  $Z_i=0$  tersebut memiliki nilai yang sama dengan peluang  $Y_i$  yang berdistribusi Poisson dengan parameter  $\lambda_i$  yaitu sebesar  $(1 - \pi_i)$ , seperti yang didefinisikan pada persamaan 2.10. Dengan demikian, dapat dikatakan bahwa variabel  $Z_i \sim \text{Binomial}(1, \pi_i)$  dengan  $E(Z_i) = \pi_i$  dan  $Var(Z_i) = \pi_i(1 - \pi_i)$ . Kemudian membentuk distribusi gabungan dengan data yang lengkap  $(y, z)$  menjadi:

$$\begin{aligned} f(y, z | \pi, \lambda) &= f(z) f(y | z) \\ &= f(z | 1, \pi) f(y | z, \lambda) \\ &= (\pi_i)^{z_i} (1 - \pi_i)^{1 - z_i} \left( \frac{\exp(-\lambda_i) \lambda_i^{y_i}}{y_i!} \right)^{1 - z_i} \end{aligned} \quad (2.20)$$

Substitusikan persamaan 2.16 ke persamaan 2.20.

$$f(\boldsymbol{\beta}, \boldsymbol{\gamma} | y, z) = \left( \frac{\exp(-e^{\boldsymbol{x}_i^T \boldsymbol{\beta}}) (e^{\boldsymbol{x}_i^T \boldsymbol{\beta}})^{y_i}}{y_i!} \right)^{1 - z_i} \left( \frac{1}{1 + e^{\boldsymbol{x}_i^T \boldsymbol{\gamma}}} \right) (e^{\boldsymbol{x}_i^T \boldsymbol{\gamma}})^{z_i} \quad (2.21)$$

Sehingga fungsi *likelihood*nya menjadi:

$$L(\boldsymbol{\beta}, \boldsymbol{\gamma} | \mathbf{y}, \mathbf{z}) = \prod_{i=1}^n \left[ \left( \frac{\exp(-e^{\mathbf{x}_i^T \boldsymbol{\beta}}) (e^{\mathbf{x}_i^T \boldsymbol{\beta}})^{y_i}}{y_i!} \right)^{1-z_i} \left( \frac{1}{1+e^{\mathbf{x}_i^T \boldsymbol{\gamma}}} \right)^{z_i} \right] \quad (2.22)$$

dan fungsi log natural *likelihood*nya adalah:

$$\begin{aligned} \log L(\boldsymbol{\beta}, \boldsymbol{\gamma} | \mathbf{y}, \mathbf{z}) &= \sum_{i=1}^n \log(f(y_i | z_i, \boldsymbol{\beta})) + \sum_{i=1}^n \log(f(z_i | \boldsymbol{\gamma})) \\ &= \sum_{i=1}^n (1-z_i) \left( y_i \mathbf{x}_i^T \boldsymbol{\beta} - e^{\mathbf{x}_i^T \boldsymbol{\beta}} \right) - \sum_{i=1}^n (1-z_i) \log(y_i!) \quad (2.23) \\ &\quad + \sum_{i=1}^n \left( z_i \mathbf{x}_i^T \boldsymbol{\gamma} - \log(1+e^{\mathbf{x}_i^T \boldsymbol{\gamma}}) \right) \end{aligned}$$

Fungsi log-*likelihood* ini biasa disebut *complete data likelihood*. Fungsi ini yang akan dimaksimumkan menggunakan algoritma EM, dimana vektor parameter  $\boldsymbol{\beta}$  dan  $\boldsymbol{\gamma}$  dapat diestimasi secara terpisah dengan persamaan 2.24.

$$\log L(\boldsymbol{\beta}, \boldsymbol{\gamma} ; \mathbf{y}, \mathbf{z}) = L(\boldsymbol{\beta} ; \mathbf{y}, \mathbf{z}) - \sum_{i=1}^n (1-z_i) \log(y_i!) + L(\boldsymbol{\gamma} ; \mathbf{y}, \mathbf{z}) \quad (2.24)$$

dengan,

$$\log L(\boldsymbol{\beta} ; \mathbf{y}, \mathbf{z}) = \sum_{i=1}^n (1-z_i) \left( y_i \mathbf{x}_i^T \boldsymbol{\beta} - e^{\mathbf{x}_i^T \boldsymbol{\beta}} \right) \quad (2.25)$$

dan,

$$\log L(\boldsymbol{\gamma} ; \mathbf{y}, \mathbf{z}) = \sum_{i=1}^n \left( z_i \mathbf{x}_i^T \boldsymbol{\gamma} - \log(1+e^{\mathbf{x}_i^T \boldsymbol{\gamma}}) \right)$$

Sedangkan untuk  $-\sum_{i=1}^n (1-z_i) \log(y_i!)$  dapat diabaikan karena tidak

mengandung parameter  $\boldsymbol{\beta}$  dan  $\boldsymbol{\gamma}$ .

3. Selanjutnya dilakukan tahap ekspektasi dan maksimalisasi:

a. Tahap ekspektasi yang dilakukan adalah.

Menghitung nilai ekspektasi dari  $Z_i$  dengan mengganti variabel  $Z_i^{(k)}$  dengan

$$E\left(Z_i | y_i, \boldsymbol{\gamma}^{(k)}, \boldsymbol{\beta}^{(k)}\right) = Z_i^{(k)}$$

$$Z_i^{(k)} = P\left(Z_i = 1 | y_i, \boldsymbol{\gamma}^{(k)}, \boldsymbol{\beta}^{(k)}\right)$$

$$= \begin{cases} \frac{1}{1 + \exp\left(-\mathbf{x}_i^T \boldsymbol{\gamma} - \exp\left(\mathbf{x}_i^T \boldsymbol{\beta}\right)\right)}, & \text{untuk } y_i = 0 \\ 0, & \text{untuk } y_i > 0 \end{cases}$$

Sehingga persamaan 2.25 menjadi:

$$\log L\left(\boldsymbol{\beta}; \mathbf{y}, \mathbf{z}^{(k)}\right) = \sum_{i=1}^n \left(1 - z_i^{(k)}\right) \left(y_i \mathbf{x}_i^T \boldsymbol{\beta} - e^{\mathbf{x}_i^T \boldsymbol{\beta}}\right) \quad (2.26)$$

dan

$$\log L\left(\boldsymbol{\gamma}; \mathbf{y}, \mathbf{z}^{(k)}\right) = \sum_{i=1}^n \left(z_i^{(k)} \mathbf{x}_i^T \boldsymbol{\gamma} - \log\left(1 + e^{\mathbf{x}_i^T \boldsymbol{\gamma}}\right)\right) \quad (2.27)$$

b. Tahap maksimalisasi, terdiri dari:

-Maksimalisasi untuk  $\boldsymbol{\beta}$ , dengan menghitung  $\boldsymbol{\beta}^{(k+1)}$  yang diperoleh dari memaksimalkan persamaan 2.27 dengan metode quasi-Newton.

-Maksimalisasi untuk  $\boldsymbol{\gamma}$ , dengan memaksimalkan persamaan 2.26

dimana untuk setiap  $y_i > 0$ , nilai  $z_i^{(k)}$  akan sama dengan 0, sehingga persamaan 2.26 dapat ditulis menjadi:

$$\begin{aligned} \log L\left(\boldsymbol{\gamma}; \mathbf{y}, \mathbf{z}^{(k)}\right) &= \sum_{y_i=0} z_i^{(k)} \mathbf{x}_i^T \boldsymbol{\gamma} - \sum_{y_i=0} z_i^{(k)} \log\left(1 + e^{\mathbf{x}_i^T \boldsymbol{\gamma}}\right) \\ &\quad - \sum_{i=1}^n \left(1 - z_i^{(k)}\right) \log\left(1 + e^{\mathbf{x}_i^T \boldsymbol{\gamma}}\right) \end{aligned} \quad (2.28)$$

Metode Quasi-Newton mengganti komputasi yang bersifat derivatif dengan fungsi komputasi langsung. Matriks Hessian H diganti dengan aproksimasi atau perkiraan matriks Hessian A yang merupakan

matriks definit positif yang mempunyai sifat seperti matriks Hessian. Format iteratif dari metode ini adalah sebagai berikut:

$$\boldsymbol{\gamma}^{(k+1)} = \boldsymbol{\gamma}^{(k)} + \boldsymbol{\alpha}^{(k)} \mathbf{S}^{(k)} \quad (2.29)$$

Dimana,

$\boldsymbol{\alpha}^{(k)}$  : *step length* yang dapat meminimumkan fungsi  $g(\boldsymbol{\alpha}) = f(\boldsymbol{\gamma}^{(k)} + \boldsymbol{\alpha}^{(k)} \mathbf{S}^{(k)})$

$\mathbf{S}^{(k)}$  : *search direction*.

Dengan,

$$\mathbf{S}^{(k)} = -(\mathbf{A}^{(k)})^{-1} \nabla f(\boldsymbol{\gamma}^{(k)}) \quad (2.30)$$

Metode Quasi-Newton berbeda dalam bagaimana perkiraan matriks Hessian dibentuk dan diupdate. Metode Quasi-Newton yang paling sederhana menetapkan perkiraan matriks Hessian sebagai matriks identitas. Cara update yang paling terkenal adalah dengan metode *Broyden, Fletcher, Goldfarb, and Shanno* (BFGS). Metode ini dikenal akan ketahannya (*robustness*) dan dapat mencapai konvergensi superlinear dengan baik. Rumus yang digunakan untuk update matriks Hessian adalah:

$$\mathbf{A}^{(k+1)} = \mathbf{A}^{(k)} + \frac{\mathbf{q}^{(k)} (\mathbf{q}^{(k)})^T}{(\mathbf{q}^{(k)})^T \mathbf{s}^{(k)}} + \frac{\mathbf{g}^{(k)} (\mathbf{g}^{(k)})^T}{(\boldsymbol{\delta}^{(k)})^T \mathbf{g}^{(k)}} \quad (2.31)$$

Dengan

$$\mathbf{q}^{(k)} = \nabla h(\boldsymbol{\gamma}^{(k+1)}) - \nabla h(\boldsymbol{\gamma}^{(k)})$$

$$\mathbf{g}^{(k)} = \nabla h(\boldsymbol{\gamma}^{(k)})$$

$$\boldsymbol{\delta}^{(k)} = \mathbf{S}^{(k)}$$

$$\mathbf{s}^{(k)} = \Delta \boldsymbol{\gamma} = \boldsymbol{\gamma}^{(k+1)} - \boldsymbol{\gamma}^{(k)}$$

$$h(\boldsymbol{\gamma}^{(k)}) = \log L(\boldsymbol{\gamma}; \mathbf{y}, \mathbf{z}^{(k)})$$

4. Mengganti  $\boldsymbol{\beta}^{(k)}$  dan  $\boldsymbol{\gamma}^{(k)}$  dengan  $\boldsymbol{\beta}^{(k+1)}$  dan  $\boldsymbol{\gamma}^{(k+1)}$ , kemudian lakukan kembali tahap ekspektasi.

5. Mengulangi tahap ke-3 dan tahap ke-4 hingga diperoleh penaksir parameter yang konvergen.

### 2.8.2 Pengujian Parameter Model *Zero Inflated Poisson Regression*

Pengujian kesesuaian model regresi ZIP menggunakan *Likelihood Ratio (LR) test*. Masing-masing pengujian beserta hipotesis statistik *LR test* yang akan dijelaskan lebih lanjut (Lambert, 1992).

#### 1. Pengujian Serentak

Pengujian serentak dilakukan untuk menguji parameter  $\beta$  dan  $\gamma$  secara bersama-sama. Parameter  $\beta$  merupakan parameter untuk *Poisson state* dan  $\gamma$  merupakan parameter untuk *zero state*. Hipotesis yang digunakan adalah:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1: \text{paling sedikit ada satu } \beta_j \neq 0, j = 1, 2, \dots, k$$

dan

$$H_0: \gamma_1 = \gamma_2 = \dots = \gamma_k = 0$$

$$H_1: \text{paling sedikit ada satu } \gamma_j \neq 0, j = 1, 2, \dots, k$$

Dengan  $k + 1$  merupakan jumlah parameter,  $\beta_j$  merupakan parameter model log ke-  $j$  dan  $\gamma_j$  merupakan parameter logit ke-  $j$ . Statistik uji yang digunakan adalah ukuran *likelihood ratio* (devians) yang dibentuk dengan menentukan himpunan parameter di bawah populasi ( $\Omega$ ) yaitu  $\Omega = \{\beta, \gamma\}$ . Himpunan parameter dibawah  $H_0$  benar ( $\omega$ ) yaitu  $\omega = \{\beta_0, \gamma_0\}$ . Kemudian pada himpunan parameter di bawah populasi, dibentuk fungsi *likelihood* untuk model penuh (*saturated*) yang melibatkan seluruh variabel prediktor. Sedangkan pada populasi parameter di bawah  $H_0$  benar, dibentuk fungsi *likelihood* untuk model yang tidak melibatkan variabel prediktor. Kedua fungsi *likelihood*  $L(\Omega)$  dan  $L(\omega)$  adalah:

$$L(\Omega) = \prod_{i=1}^n \left[ \left( \frac{\exp(-e^{x_i^T \beta}) (e^{x_i^T \beta})^{y_i}}{y_i!} \right)^{1-z_i} \left( \frac{1}{1+e^{x_i^T \gamma}} \right) (e^{x_i^T \gamma})^{z_i} \right] \quad (2.32)$$

dan,

$$L(\omega) = \prod_{i=1}^n \left[ \left( \frac{\exp(-e^{\beta_0}) (e^{\beta_0})^{y_i}}{y_i!} \right)^{1-z_i} \left( \frac{1}{1+e^{\gamma_0}} \right) (e^{\gamma_0})^{z_i} \right] \quad (2.33)$$

Perhitungan statistik uji untuk pengujian kesesuaian model adalah sebagai berikut (Hall & Shen, 2009).

$$G^2 = -2 \ln \left[ \frac{L(\hat{\omega})}{L(\hat{\Omega})} \right]$$

$$= -2 \ln \left( \frac{\prod_{i=1}^n \left[ \left( \frac{\exp(-e^{x_i^T \hat{\beta}}) (e^{x_i^T \hat{\beta}})^{y_i}}{y_i!} \right)^{1-z_i} \left( \frac{1}{1+e^{x_i^T \hat{\gamma}}} \right) (e^{x_i^T \hat{\gamma}})^{z_i} \right]}{\prod_{i=1}^n \left[ \left( \frac{\exp(-e^{\beta_0}) (e^{\beta_0})^{y_i}}{y_i!} \right)^{1-z_i} \left( \frac{1}{1+e^{\gamma_0}} \right) (e^{\gamma_0})^{z_i} \right]} \right) \quad (2.34)$$

dengan  $z_i$  merupakan variabel indikator ke- $i$ ,  $x_i^T$  merupakan variabel prediktor observasi ke- $i$ ,  $y_i$  merupakan variabel respon ke- $i$ ,  $\hat{\gamma}$  dan  $\hat{\beta}$  merupakan parameter yang ditaksir. Statistik  $G^2$  berdistribusi  $\chi^2_{(df)}$  sehingga pada taraf signifikansi  $\alpha$  tolak  $H_0$  jika nilai  $G^2 > \chi^2_{(df, \alpha)}$ , dimana  $df$  merupakan jumlah parameter di bawah populasi dikurangi dengan parameter di bawah  $H_0$ .



## 2. Pengujian Parsial

Pada pengujian parameter secara individu terdapat dua pengujian, yaitu pengujian parameter model log dan pengujian parameter model logit. Berikut ini adalah hipotesis pengujian parameter model log .

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0, j = 1, 2, \dots, k$$

statistik uji:

$$Z = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (2.35)$$

dan hipotesis yang digunakan dalam pengujian parameter secara parsial untuk model logit adalah:

$$H_0: \gamma_j = 0$$

$$H_1: \gamma_j \neq 0, j = 1, 2, \dots, k$$

statistik uji:

$$Z = \frac{\hat{\gamma}_j}{SE(\hat{\gamma}_j)} \quad (2.36)$$

dengan  $z_i$  merupakan variabel indikator ke-i,  $x_i^T$  merupakan variabel prediktor ke-i,  $y_i$  merupakan variabel respon ke-i,  $\hat{\gamma}$  dan  $\hat{\beta}$  merupakan parameter yang ditaksir. Daerah penolakan untuk pengujian kesesuaian model, pengujian parameter model log dan pengujian parameter model logit adalah tolak  $H_0$  apabila  $Z > Z_{\alpha/2}$  atau  $Z < -Z_{\alpha/2}$ .

## 2.9 Pemilihan Model Terbaik

Model terbaik pada *Zero Inflated Poisson regression* didapatkan dengan pemilihan model terbaik menggunakan kriteria *Akaike's Information Criterion* (AIC) dan *Akaike's Information Corrected Criterion* (AIC<sub>c</sub>). AIC dan AIC<sub>c</sub> adalah kriteria kesesuaian model dalam mengestimasi secara statistik. Kriteria statistik biasanya digunakan apabila pembentukan model regresi bertujuan untuk mendapatkan

faktor-faktor yang berpengaruh terhadap model, bukan untuk melakukan suatu prediksi.

Biasanya nilai AIC sejalan dengan nilai devians dari model. Semakin kecil nilai devians maka akan semakin kecil pula tingkat kesalahan yang dihasilkan model sehingga model yang diperoleh menjadi semakin tepat. Oleh karena itu, model terbaik adalah model dengan nilai AIC terkecil. Nilai AIC dirumuskan pada persamaan 2.37 (Bozdogan, 2000).

$$AIC = -2 \times (\log \text{likelihood}) + 2k \quad (2.37)$$

dengan,

$k$  : jumlah parameter dalam model.

Rumus  $AIC_c$  ditunjukkan pada persamaan 2.38.

$$AIC_c = -2 \times (\log \text{likelihood}) + 2k \left( \frac{n}{n - k - 1} \right) \quad (2.38)$$

dengan,

$n$  : jumlah observasi.

Model dengan nilai  $AIC_c$  kecil merupakan model yang lebih baik.

## BAB III METODOLOGI PENELITIAN

### 3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Profil Kesehatan masing-masing kabupaten di Sulawesi Barat 2016 yaitu Profil Kesehatan Kabupaten Majene 2016 (Dinas Kesehatan Kabupaten Majene, 2017), Profil Kesehatan Mamasa 2016 (Dinas Kesehatan Kabupaten Mamasa, 2017), Profil Kesehatan Kabupaten Mamuju 2016 (Dinas Kesehatan Kabupaten Mamuju, 2017), Profil Kesehatan Kabupaten Polewali Mandar 2016 (Dinas Kesehatan Kabupaten Polewali Mandar, 2017), Profil Kesehatan Kabupaten Mamuju Tengah 2016 (Dinas Kesehatan Kabupaten Mamuju Tengah, 2017), dan Profil Kesehatan Kabupaten Mamuju Utara 2016 (Dinas Kesehatan Kabupaten Mamuju Utara, 2017). Unit pengamatan sebanyak 69 unit berupa kecamatan di Provinsi Sulawesi Barat.

### 3.2 Variabel Penelitian

Variabel penelitian yang digunakan pada penelitian ini terdiri variabel respon yaitu jumlah kematian bayi dan prediktor yaitu faktor-faktor yang diduga mempengaruhi jumlah kematian bayi di Sulawesi Barat. Variabel-variabel pada penelitian ini ditunjukkan pada Tabel 3.1.

**Tabel 3.1** Variabel Penelitian

	<b>Variabel</b>	<b>Skala</b>
Y	Jumlah kematian bayi	Rasio
X <sub>1</sub>	Persentase posyandu akif	Rasio
X <sub>2</sub>	Persentase rumah tangga yang ber-PHBS	Rasio
X <sub>3</sub>	Persentase cakupan desa/kelurahan <i>Universal Child Immunization</i> (UCI)	Rasio
X <sub>4</sub>	Persentase persalinan ditolong tenaga kesehatan	Rasio
X <sub>5</sub>	Persentase penduduk dengan akses berkelanjutan terhadap air minum layak	Rasio
X <sub>6</sub>	Persentase penduduk dengan akses sanitasi layak	Rasio

Adapun definisi variabel penelitian adalah sebagai berikut:

1. Jumlah Kematian Bayi

Jumlah kematian bayi adalah banyaknya kematian bayi usia 0 sampai 1 tahun.

2. Persentase posyandu aktif

Posyandu aktif meliputi posyandu pratama, madya, purnama dan mandiri.

3. Persentase rumah tangga yang ber-PHBS

Perilaku hidup bersih dan sehat di rumah tangga merupakan upaya dalam membe rdayakan anggota keluarga rumah tangga agar tahu, sadar, mau dan mampu melakukan Perilaku Hidup Bersih dan Sehat (PHBS) dalam memelihara dan meningkatkan kesehatannya, mencegah risiko terjadinya penyakit dan melindungi diri dari ancaman penyakit serta berperan aktif dalam gerakan kesehatan masyarakat (Dinas Kesehatan Kabupaten Majene, 2017).

4. Persentase cakupan desa/kelurahan *Universal Child Immunization* (UCI)

UCI Desa merupakan indikator penting dalam program imunisasi, Desa/ Kelurahan UCI adalah Desa/ Kelurahan dimana 80% dari jumlah bayi yang ada di desa tersebut sudah mendapat imunisasi dasar lengkap dalam waktu satu tahun (Dinas Kesehatan Kabupaten Majene, 2017).

5. Persentase ibu bersalin ditolong tenaga kesehatan

Persentase ibu bersalin ditolong tenaga kesehatan merupakan persentase pertolongan persalinan adalah proses pelayanan persalinan yang dimulai pada kala I sampai dengan kala IV persalinan. Pencapaian upaya kesehatan ibu bersalin diukur melalui indikator persentase persalinan ditolong tenaga kesehatan terlatih. Indikator ini memperlihatkan diantaranya tingkat kemampuan pemerintah dalam menyediakan pelayanan.

6. Persentase penduduk dengan akses berkelanjutan terhadap air minum layak

Variabel tersebut merupakan variabel persentase penduduk yang menggunakan akses terhadap air minum melalui sumur gali terlindung, sumur gali dengan pompa, sumur bor dengan pompa, terminal air, mata air terlindung, dan penampungan air hujan serta perpipaan Perusahaan Daerah Air Minum (PDAM) dan Badan Peningkatan Sistem Penyedia Air Minum (BPSPAM) yang layak. Kriteria air minum yang layak sendiri adalah syarat syarat air minum adalah tidak berasa, tidak berbau, tidak berwarna, dan tidak mengandung logam berat.

#### 7. Persentase penduduk dengan akses sanitasi layak

Sanitasi layak dapat diartikan bahwa sistem sanitasi sederhana di masyarakat yang di nilai layak dengan kriteria rumah tangga tersebut memiliki akses ke sanitasi layak jika menggunakan fasilitas buang air besar sendiri atau bersama dan memiliki tempat pembuangan akhir seperti *septic tank* atau sistem pengolahan air limbah (SPAL).

### 3.3 Struktur Data

Struktur data secara umum yang digunakan dalam penelitian ini yang terdapat pada Tabel 3.2.

**Tabel 3.2** Struktur Data

No	Kecamatan	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
1	Banggae	Y <sub>1</sub>	X <sub>(1,1)</sub>	X <sub>(1,2)</sub>	X <sub>(1,3)</sub>	X <sub>(1,4)</sub>	X <sub>(1,5)</sub>	X <sub>(1,6)</sub>
2	Banggae Timur	Y <sub>2</sub>	X <sub>(2,1)</sub>	X <sub>(2,2)</sub>	X <sub>(2,3)</sub>	X <sub>(2,4)</sub>	X <sub>(2,5)</sub>	X <sub>(2,6)</sub>
3	Pamboang	Y <sub>3</sub>	X <sub>(3,1)</sub>	X <sub>(3,2)</sub>	X <sub>(3,3)</sub>	X <sub>(3,4)</sub>	X <sub>(3,5)</sub>	X <sub>(3,6)</sub>
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
68	Tapoyo	Y <sub>68</sub>	X <sub>(68,1)</sub>	X <sub>(68,2)</sub>	X <sub>(68,3)</sub>	X <sub>(68,4)</sub>	X <sub>(68,5)</sub>	X <sub>(68,6)</sub>
69	Karossa	Y <sub>69</sub>	X <sub>(69,1)</sub>	X <sub>(69,2)</sub>	X <sub>(69,3)</sub>	X <sub>(69,4)</sub>	X <sub>(69,5)</sub>	X <sub>(69,6)</sub>

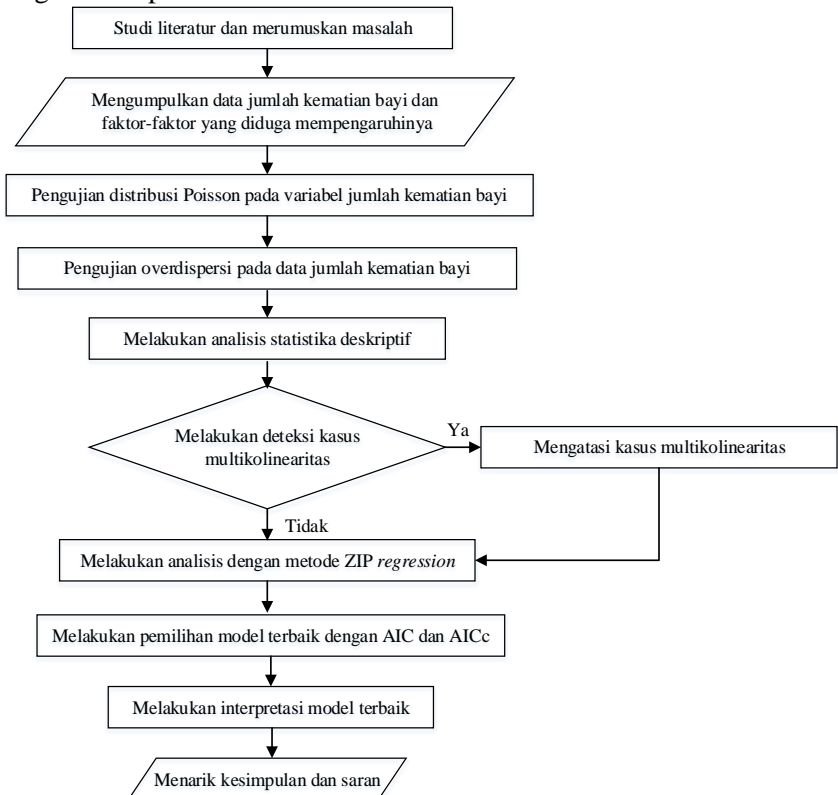
### 3.4 Langkah Analisis

Langkah analisis yang dilakukan adalah sebagai berikut:

1. Studi literatur dan merumuskan permasalahan.
2. Mengumpulkan data jumlah kematian bayi dan faktor-faktor yang diduga mempengaruhi jumlah kematian bayi di Sulawesi Barat.
3. Melakukan pengujian distribusi Poisson pada variabel respon.
4. Melakukan pengujian overdispersi pada variabel respon.

5. Melakukan analisis statistika deskriptif.
6. Melakukan deteksi multikolinearitas dengan VIF berdasarkan persamaan 2.3.
7. Melakukan analisis dengan metode regresi *Zero Inflated Poisson* berdasarkan persamaan 2.15.
8. Melakukan pemilihan model terbaik dengan AIC dan  $AIC_C$  berdasarkan persamaan 2.37 dan 2.38.
9. Melakukan interpretasi model yang terbentuk.
10. Menarik kesimpulan dan saran.

Langkah-langkah analisis secara umum digambarkan pada diagram alir pada Gambar 3.1.



**Gambar 3.1** Diagram Alir Langkah Analisis

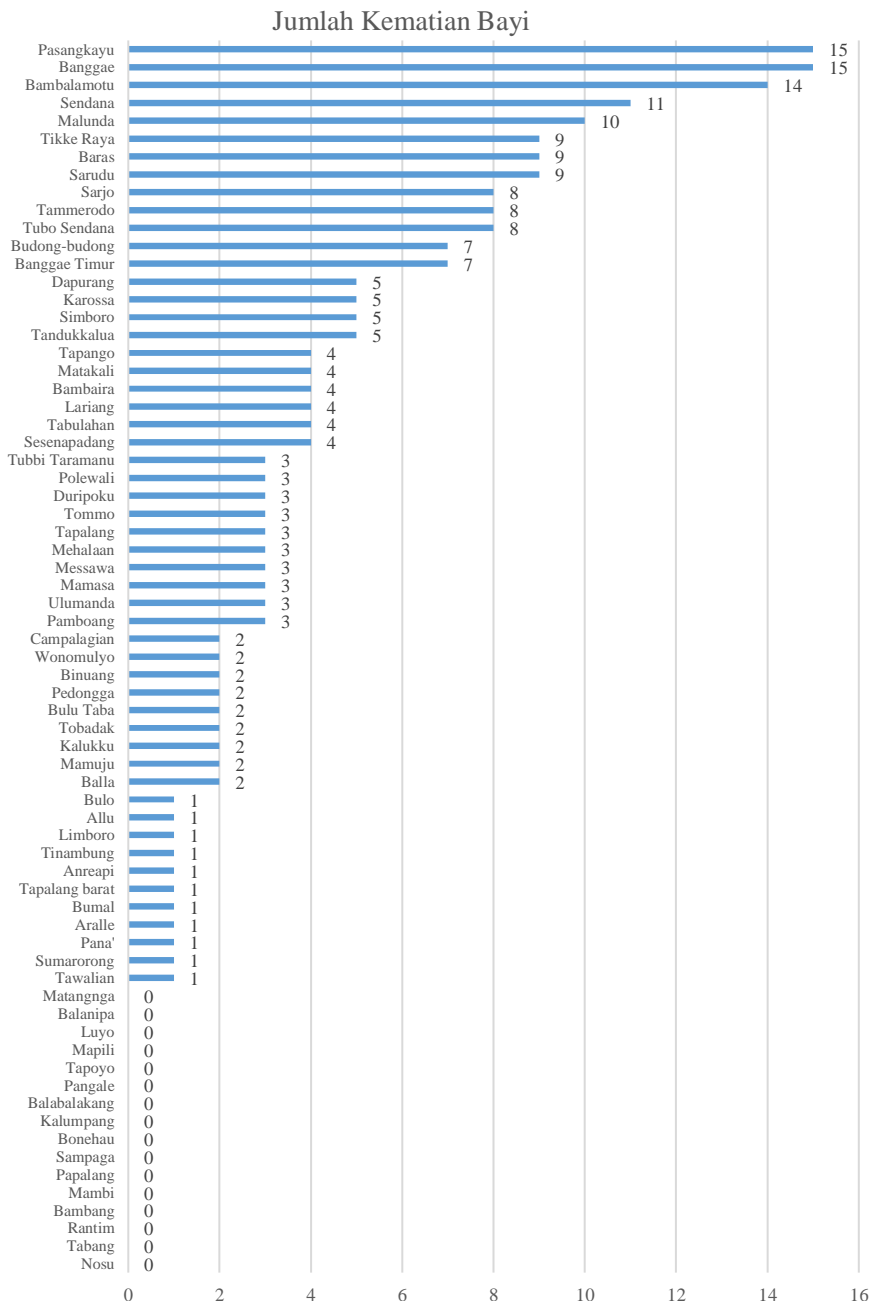
## **BAB IV**

### **ANALISIS DAN PEMBAHASAN**

Bagian ini membahas mengenai hasil penelitian, meliputi karakteristik kematian bayi dan faktor-faktor yang mempengaruhi kematian bayi di Sulawesi Barat, pembentukan model *Zero Inflated Poisson Regression* serta pemilihan model terbaik.

#### **4.1 Eksplorasi Jumlah Kematian Bayi dan Faktor-faktor yang Diduga Mempengaruhi Kematian Bayi di Sulawesi Barat**

Kematian bayi adalah kematian individu yang berusia 0-1 tahun. Kematian bayi yang salah satunya dapat direfleksikan dalam Angka Kematian Bayi (AKB) yang merupakan indikator yang penting untuk mencerminkan keadaan derajat kesehatan di suatu masyarakat, serta tolok ukur yang sensitif dari semua upaya intervensi yang dilakukan oleh pemerintah khususnya di bidang kesehatan. Gambar 4.1 menunjukkan grafik kematian bayi di Sulawesi Barat tahun 2016.



**Gambar 4.1** Diagram Batang Jumlah Kematian Bayi di Sulawesi Barat tahun 2016 per Kecamatan



Berdasarkan data profil kesehatan Kabupaten Mamuju Utara 2016, beberapa faktor yang diduga menyebabkan tingginya kematian bayi di kecamatan Pasangkayu ditunjukkan pada Tabel 4.1.

**Tabel 4.1** Variabel Prediktor Kecamatan Pasangkayu

Variabel	Persentase
Persentase cakupan desa/kelurahan UCI	37,5%
Persentase rumah tangga yang ber-PHBS	54,551%

Faktor tersebut adalah persentase cakupan desa/kelurahan *Universal Child Immunization* (UCI) dan persentase rumah tangga yang ber-PHBS juga cukup rendah. Persentase cakupan desa/kelurahan UCI hanya sebesar 37,5% tidak mencapai 50% persentase rumah tangga yang ber-PHBS sebesar 54,551% masih cukup rendah.

**Tabel 4.2** Variabel Prediktor Kecamatan Banggae

Variabel	Persentase
Persentase posyandu aktif	46,512%
Persentase desa/kelurahan UCI	62,50%

Kecamatan Banggae juga terdapat kematian bayi yang tinggi dibanding dengan kecamatan lainnya yaitu sebesar 15 kematian bayi pada tahun 2016. Berdasarkan profil kesehatan Kabupaten Majene, rendahnya keberadaan posyandu aktif dan persentase desa/kelurahan UCI diduga menjadi salah satu faktor yang berpengaruh pada kematian bayi di kecamatan Banggae. Persentase posyandu aktif di kecamatan Banggae hanya sebesar 46,512% dan persentase desa/kelurahan UCI yaitu 62,50%.

**Tabel 4.3** Variabel Prediktor 16 Kecamatan

Variabel	Persentase
Rata-rata persentase ibu bersalin ditolong tenaga kesehatan	73,90%
Rata-rata persentase desa/kelurahan UCI	65,775%
Rata-rata persentase rumah tangga yang ber-PHS	53,855%

Sedangkan ada 16 kecamatan yang tidak terjadi kematian bayi di daerahnya, yaitu kecamatan Tapoyo, Pangale, Balabalakang, Kalumpang, Bonehau, Sampaga, Papalang, Mambi, Bambang, Rantim, Tabang, Nosu, Matangnga, Balanipa, Luyo dan Mapili. Berdasarkan profil kesehatan kabupaten di Sulawesi Barat 2016, rata-rata persentase desa/kelurahan UCI, rumah tangga yang ber-PHBS dan ibu bersalin ditolong tenaga kesehatan di 16 daerah tersebut cukup tinggi dan ditunjukkan pada Tabel 4.3. Rata-rata persentase ibu bersalin ditolong tenaga kesehatan cukup tinggi yaitu sebesar 73,90% dan rata-rata persentase desa/kelurahan UCI cukup tinggi yaitu sebesar 65,775%, serta rata-rata persentase rumah tangga yang ber-PHBS yaitu 53,855%.

Karakteristik faktor-faktor yang diduga mempengaruhi kematian bayi di Sulawesi Barat tahun 2016 ditunjukkan pada Tabel 4.4.

**Tabel 4.4** Statistika Deskriptif Faktor-Faktor yang Diduga Mempengaruhi Kematian Bayi di Sulawesi Barat tahun 2016

<b>Variabel</b>	<b>Mean</b>	<b>Varians</b>	<b>Minimum</b>	<b>Maksimum</b>
X <sub>1</sub>	60,75	1578,08	0,00	100,00
X <sub>2</sub>	67,96	865,40	0,00	160,75
X <sub>3</sub>	63,18	861,97	0,00	125,00
X <sub>4</sub>	77,60	160,19	43,79	103,16
X <sub>5</sub>	38,89	1366,44	0,00	120,39
X <sub>6</sub>	49,52	1468,95	0,00	137,37

Berdasarkan Tabel 4.4 dapat diketahui bahwa variabel X<sub>1</sub> yang merupakan persentase posyandu aktif memiliki rata-rata sebesar 60,75. Nilai tersebut menunjukkan bahwa persentase posyandu yang aktif melakukan kegiatan utama dan memiliki kader dalam 6 bulan terakhir di Sulawesi Barat sebanyak 60,75%. Kecamatan yang tidak memiliki persentase posyandu aktif ditunjukkan dengan nilai minimum 0% adalah kecamatan Tubbi Taramanu, Rantim, Tapalang, Tapalang Barat, Simboro, Papalang,

Bonehau, Kalumpang, Balabalakang, dan Pangale. Sedangkan kecamatan yang semua posyandu di daerahnya aktif 100% adalah kecamatan Alu, Mamasa, Tawallan, Sesenapadang, Balla, Tandukkalua, Sumarorong, Messawa, Nosu, Pana', Tabang, Bambang, Aralle, Tabulahan, Mambi, Bumai, Mehalaan, Sarudu, Dapurang, Duripoku, Lariang, Bulu Taba, Baras, Pasangkayu, Tikke Raya, Pedongga, Bambalamotu, Bambaira, dan Sarjo atau 29 kecamatan dari 69 kecamatan yang ada di Sulawesi Barat.

Variabel persentase RT yang ber-PHBS ( $X_2$ ) memiliki rata-rata sebesar 67,47. Nilai tersebut menunjukkan bahwa sebesar 67,47% rumah tangga di Sulawesi Barat telah melakukan Perilaku Hidup Bersih dan Sehat (PHBS). Keragaman rumah tangga yang telah melakukan PHBS di Sulawesi Barat cukup tinggi yaitu sebesar 755,14. Persentase rumah tangga yang ber-PHBS minimal adalah 0%, artinya seluruh rumah tangga tidak melakukan perilaku hidup sehat. Kecamatan yang seluruh rumah tangganya tidak melakukan PHBS adalah kecamatan Rantim. Persentase rumah tangga yang ber-PHBS maksimum adalah sebesar 120,65% artinya sebesar 120,65% rumah tangga melakukan PHBS yang ada di kecamatan Tapoyo.

Variabel  $X_3$  atau variabel persentase cakupan desa/kelurahan UCI memiliki rata-rata sebesar 63,16%. Nilai tersebut menunjukkan 63,16% kecamatan di Sulawesi Barat dimana desa/kelurahan dari jumlah bayi yang ada di desa tersebut sudah mendapat imunisasi dasar lengkap dalam waktu satu tahun. Persentase desa/kelurahan UCI bernilai minimum 0% artinya kecamatan tersebut desa/kelurahannya tidak memiliki desa/kelurahan UCI. Kecamatan yang tidak memiliki desa/kelurahan UCI adalah kecamatan Tawallan dan Balabalakang. Nilai maksimum desa/kelurahan UCI sebesar 125%, artinya sebanyak 125% desa/kelurahan di Kecamatan Tikke Raya memiliki desa/kelurahan dari jumlah bayi yang ada di desa tersebut sudah mendapat imunisasi dasar lengkap dalam waktu satu tahun.

Rata-rata persentase ibu bersalin ditolong tenaga kesehatan atau  $X_4$  adalah 77,60%. Sebanyak 77,60% kecamatan yang ada di

Sulawesi Barat para ibu hamilnya telah melakukan persalinan yang dibantu tenaga kesehatan. Varians dari persentase ibu bersalin ditolong tenaga kesehatan sebesar 160,19. Nilai minimum persentase ibu bersalin ditolong tenaga kesehatan yaitu 43,79% yang terdapat pada kecamatan Bumal. Nilai maksimum sebesar 103,16%, artinya sebanyak 103,16% ibu bersalin di kecamatan Pangale ditolong tenaga kesehatan.

Variabel persentase penduduk dengan akses berkelanjutan terhadap air minum layak ( $X_5$ ) memiliki rata-rata sebesar 38,89% atau dapat dikatakan sebesar 38,89% penduduk di Sulawesi Barat mendapatkan akses berkelanjutan air minum layak. Berdasarkan Tabel 4.4 nilai varians  $X_5$  yaitu sebesar 1366,44. Nilai keragaman persentase penduduk dengan akses berkelanjutan terhadap air minum layak cukup tinggi yaitu sebesar 1366,44. Nilai minimum yaitu 0,00% yang artinya ada wilayah yang penduduknya tidak mendapat akses berkelanjutan terhadap air minum layak yaitu kecamatan Balla, Nosu, Pana', Rantim, Bambang, Bumal, Mehalaan, Pangale, Budong-budong, Tobadak, Tapoyo dan Karossa. Sedangkan nilai maksimum sebesar 120,39% menunjukkan bahwa kecamatan Banggae merupakan kecamatan yang persentase penduduknya memiliki akses terhadap air minum yang layak paling tinggi yaitu sebesar 120,39%.

$X_6$  menunjukkan variabel persentase penduduk dengan akses sanitasi layak. Nilai rata-rata yaitu 49,52% menunjukkan bahwa sebesar 49,52% penduduk di Sulawesi Barat sudah mendapatkan akses sanitasi layak. Varians sebesar 1468,95 menunjukkan bahwa keragaman persentase penduduk dengan akses sanitasi layak cukup tinggi yaitu sebesar 1468,95. Nilai minimum yaitu 0,00% menunjukkan bahwa pada kecamatan Pangale, Budong-budong, Tobadak, Tapoyo dan Karossa penduduknya tidak memiliki akses sanitasi layak. Nilai maksimum sebesar 137,37% artinya pada kecamatan Bambang penduduknya sebanyak 137,37% memiliki akses sanitasi layak.

#### 4.2 Pengujian Distribusi Poisson pada Data Jumlah Kematian Bayi di Sulawesi Barat

Distribusi *Poisson* adalah suatu distribusi yang digunakan dalam peristiwa yang terjadi dalam suatu interval waktu atau wilayah tertentu. Pengujian distribusi Poisson pada data jumlah kematian bayi di Sulawesi Barat menggunakan uji Kolmogorov Smirnov dengan hipotesisnya adalah sebagai berikut:

$H_0$  : data berdistribusi Poisson

$H_1$  : data tidak berdistribusi Poisson

**Tabel 4.5** Pengujian Distribusi Poisson pada Data Jumlah Kematian Bayi di Sulawesi Barat

Variabel	<i>Kolmogorov Smirnov</i>
Jumlah kematian bayi	0,24181

Berdasarkan hasil pengujian didapatkan nilai D yaitu sebesar 0,24181. Jika nilai tersebut dibandingkan dengan nilai  $D_{(0,05;69)}$  yaitu 0,1637 akan didapatkan keputusan tolak  $H_0$ . Kesimpulan yang dapat diambil adalah data jumlah kematian bayi di Sulawesi Barat tahun 2016 tidak berdistribusi Poisson.

Namun, dalam pengujian menggunakan *software easy fit* didapatkan hasil bahwa data memiliki distribusi Poisson pada urutan kedua. Tabel 4.6 menunjukkan hasil pengujian menggunakan bantuan *software easy fit*.

**Tabel 4.6** Pengujian Distribusi Data Menggunakan *Software Easy Fit*

Variabel	<i>Kolmogorov Smirnov</i>	
	Statistik Uji	Peringkat
Geometric	0,22848	1
Poisson	0,24181	2
Neg. Binomial	0,24625	3
D. Uniform	0,30769	4

Tabel 4.6 menunjukkan bahwa data jumlah kematian bayi di Sulawesi Barat tahun 2016 berdistribusi Poisson pada peringkat kedua, sehingga dapat dikatakan bahwa data berdistribusi Poisson.

### 4.3 Pengujian Overdispersi pada Data Jumlah Kematian Bayi di Sulawesi Barat

Overdispersi merupakan kasus yang terjadi jika sebuah variabel memiliki nilai varians yang lebih besar dari nilai rata-rata. Untuk mengetahui sebuah terjadi overdispersi atau tidak dapat dilakukan dengan membandingkan nilai rata-rata dan varians variabel tersebut. Nilai rata-rata dan varians variabel jumlah kematian bayi ditunjukkan pada Tabel 4.7.

**Tabel 4.7** Perbandingan Nilai *Mean* dan Varians Data Jumlah Kematian Bayi di Sulawesi Barat

Variabel	Mean	Varians
Jumlah kematian bayi	3,377	13,915

Berdasarkan Tabel 4.7 dapat diketahui bahwa nilai *mean* atau rata-rata dari variabel jumlah kematian bayi adalah 3,377 dan nilai varians sebesar 13,915. Nilai varians tersebut lebih besar daripada nilai rata-rata atau 13,915 lebih besar dari 3,377 sehingga dapat disimpulkan bahwa variabel jumlah kematian bayi terjadi overdispersi.

Agar asumsi tersebut lebih valid maka dilakukan pengujian overdispersi dengan membandingkan nilai devians per derajat bebas. Jika hasilnya melebihi nilai 1 maka terjadi overdispersi.

**Tabel 4.8** Pengujian Overdispersi pada data Jumlah Kematian Bayi di Sulawesi Barat

Variabel	Devians	Devians / Derajat Bebas
Jumlah kematian bayi	262,83	3,856

Tabel 4.8 menunjukkan nilai devians per derajat bebas adalah sebesar 3,856 sehingga dapat disimpulkan bahwa data jumlah kematian bayi di Sulawesi Barat tahun 2016 terjadi overdispersi.

#### 4.4 Pengujian Asumsi Multikolinearitas pada Variabel Prediktor yang Diduga Mempengaruhi Kematian Bayi di Sulawesi Barat

Terdapat dua cara untuk mengetahui adanya kasus multikolinearitas pada variabel prediktor yang digunakan, yaitu dengan nilai koefisien korelasi antar variabel prediktor dan nilai *Variance Inflation Factor* (VIF). Berikut ini adalah nilai koefisien korelasi antar variabel prediktor Tabel 4.9.

**Tabel 4.9** Nilai Koefisien Korelasi Faktor-Faktor yang Diduga Mempengaruhi Kematian Bayi di Sulawesi Barat tahun 2016

	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$
$X_2$	0,407				
$X_3$	-0,125	-0,155			
$X_4$	-0,109	-0,117	0,275		
$X_5$	-0,415	-0,172	0,189	0,131	
$X_6$	-0,213	0,028	0,141	-0,049	0,293

Nilai koefisien korelasi antar variabel  $X_1$ ,  $X_2$ ,  $X_3$ ,  $X_4$ ,  $X_5$  dan  $X_6$  yang ditunjukkan pada Tabel 4.9 tidak melebihi nilai  $\pm 0,95$ , sehingga dapat dikatakan bahwa antar variabel prediktor tidak memiliki korelasi atau tidak terjadi multikolinearitas.

**Tabel 4.10** Nilai VIF Faktor-Faktor yang Diduga Mempengaruhi Kematian Bayi di Sulawesi Barat tahun 2016

Variabel	VIF
$X_1$	1,441
$X_2$	1,243
$X_3$	1,143
$X_4$	1,111
$X_5$	1,306
$X_6$	1,154

Tabel 4.10 menunjukkan nilai VIF pada masing-masing variabel prediktor. Nilai VIF semua variabel prediktor tidak ada yang

melebihi nilai 10, sehingga dapat disimpulkan bahwa tidak ada korelasi antara variabel prediktor atau tidak ada kasus multikolinearitas.

#### 4.5 Model Zero Inflated Poisson (ZIP) Regression pada Data Jumlah Kematian Bayi di Sulawesi Barat

Model Zero Inflated Poisson Regression merupakan model yang digunakan untuk menganalisis data *count* yang mengalami overdispersi dan mengandung nilai nol dalam proporsi besar. Hasil estimasi parameter model ZIP Regression ditunjukkan pada Tabel 4.11.

**Tabel 4.11** Hasil Estimasi Parameter Model ZIP Regression

Parameter	Estimasi	SE	Z	p-value
$\beta_0$	1,1430	0,8156	1,402	0,1611
$\beta_1$	0,0067	0,0025	2,661	0,0078
$\beta_2$	-0,0034	0,0037	-0,952	0,3411
$\beta_3$	-0,0017	0,0028	-0,597	0,5502
$\beta_4$	-0,0015	0,0075	-0,199	0,8422
$\beta_5$	0,0174	0,0025	7,067	$1,59 \times 10^{-12}$
$\beta_6$	-0,0121	0,0025	-4,837	$1,32 \times 10^{-6}$
$\gamma_0$	6,6776	3,634	1,838	0,0661
$\gamma_1$	-0,0512	0,0194	-2,634	0,0084
$\gamma_2$	0,0161	0,0201	0,802	0,4227
$\gamma_3$	0,0065	0,0207	0,313	0,7541
$\gamma_4$	-0,0755	0,0409	-1,846	0,0649
$\gamma_5$	-0,0005	0,0233	-0,021	0,9833
$\gamma_6$	-0,0496	0,0358	-1,383	0,1666

Analisis awal yang dilakukan adalah pengujian kesesuaian model untuk mengetahui apakah model sudah sesuai. Berikut ini merupakan hipotesis yang digunakan dalam pengujian serentak pada model log atau *Poisson state* dan model logit atau *zero state*:



$$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$$

(tidak ada variabel yang berpengaruh signifikan)

$$H_1: \text{paling sedikit ada satu } \beta_j \neq 0, j = 1, 2, 3, 4, 5, 6$$

(minimal ada satu variabel yang berpengaruh signifikan)

dan

$$H_0: \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = \gamma_5 = \gamma_6 = 0$$

(tidak ada variabel yang berpengaruh signifikan)

$$H_1: \text{paling sedikit ada satu } \gamma_j \neq 0, j = 1, 2, 3, 4, 5, 6$$

(minimal ada satu variabel yang berpengaruh signifikan)

Taraf signifikansi yang digunakan adalah  $\alpha = 0,05$  maka daerah penolakannya yaitu  $\chi^2_{(0,05;14)}$  sebesar 24,996. Sehingga keputusan yang dapat diambil adalah tolak  $H_0$  karena nilai  $G^2$  yaitu 303 lebih dari  $\chi^2_{(0,05;14)}$  yaitu 24,996. Kesimpulannya adalah minimal ada satu variabel yang berpengaruh signifikan atau model ZIP *regression* telah sesuai dan dapat digunakan. Setelah diketahui bahwa model telah sesuai dan dapat digunakan maka langkah selanjutnya adalah pengujian parsial. Pengujian parsial digunakan untuk mengetahui parameter mana saja yang berpengaruh signifikan terhadap model.

Hipotesis yang digunakan dalam pengujian parameter secara parsial untuk model log atau *Poisson state* adalah:

$$H_0: \beta_j = 0 \text{ (tidak berpengaruh signifikan)}$$

$$H_1: \beta_j \neq 0, j = 1, 2, 3, 4, 5, 6 \text{ (berpengaruh signifikan)}$$

Dengan taraf signifikansi  $\alpha = 0,05$  keputusan tolak  $H_0$  dapat diambil jika  $Z > Z_{(0,025)}$  atau  $Z < -Z_{(0,025)}$ . Berdasarkan Tabel 4.11, nilai  $Z$  pada model log yang lebih dari daerah kritis  $-Z_{(0,025)}$  atau -1,96 adalah  $\beta_2, \beta_3$  dan  $\beta_4$  dengan nilai -0,952; -0,597 dan -0,199.

Keputusan yang dapat diambil adalah gagal tolak  $H_0$ , sehingga kesimpulannya adalah variabel persentase rumah tangga yang ber-PHBS ( $x_2$ ), persentase cakupan desa/kelurahan UCI ( $x_3$ ) dan persentas persalinan ditolong tenaga kesehatan ( $x_4$ ) tidak berpengaruh signifikan terhadap model log. Nilai Z pada parameter  $\beta_1$  dan  $\beta_7$  yaitu 2,661 dan 7,067 lebih dari nilai  $Z_{(0,025)}$  yaitu 1,96. Keputusan yang dapat diambil adalah tolak  $H_0$  dan kesimpulannya adalah variabel  $X_1$  atau persentase posyandu aktif dan variabel  $X_5$  atau persentase penduduk yang mendapatkan akses terhadap air minum layak berpengaruh signifikan terhadap model. Sedangkan nilai Z yang kurang dari daerah kritis  $-Z_{(0,025)}$  adalah parameter  $\beta_6$  yaitu -4,837 kurang dari -1,96 didapatkan keputusan tolak  $H_0$ . Kesimpulan yang dapat diambil adalah persentase penduduk yang mendapatkan akses terhadap sanitasi layak berpengaruh signifikan. Sehingga dapat disimpulkan bahwa pada model log, variabel yang berpengaruh signifikan terhadap model adalah persentase posyandu aktif ( $x_1$ ), persentase penduduk yang mendapat akses terhadap air minum layak ( $X_5$ ) dan persentase penduduk yang mendapat akses sanitasi layak ( $x_6$ ).

Berikut ini merupakan hipotesis yang digunakan dalam pengujian parameter secara parsial untuk model logit atau *zero state*:

$H_0 : \gamma_j = 0$  (tidak berpengaruh signifikan)

$H_1 : \gamma_j \neq 0, j = 1, 2, 3, 4, 5, 6$  (berpengaruh signifikan)

Keputusan tolak  $H_0$  dapat diambil jika Z lebih dari  $Z_{(0,025)}$  atau kurang dari  $-Z_{(0,025)}$  yaitu 1,96 atau -1,96. Nilai Z pada Tabel 4.11 yang kurang dari  $Z_{(0,025)}$  atau 1,96 adalah parameter  $\gamma_2$  dan  $\gamma_3$  yaitu 0,802 dan 0,313, sehingga keputusannya adalah gagal tolak  $H_0$ . Kesimpulannya adalah variabel  $X_2$  atau persentase rumah

tangga yang ber-PHBS dan variabel  $X_3$  atau persentase cakupan desa/kelurahan UCI tidak berpengaruh signifikan. Nilai Z yang lebih dari dari  $-Z_{(0,025)}$  atau  $-1,96$  adalah parameter  $\gamma_4, \gamma_5$  dan  $\gamma_6$ . Sehingga keputusan yang dapat diambil adalah gagal tolak  $H_0$  dan kesimpulannya adalah variabel  $X_4, X_5$  dan  $X_6$  atau persentase ibu bersalin ditolong tenaga kesehatan, persentase penduduk yang mendapatkan akses terhadap air minum layak dan persentase penduduk yang mendapatkan akses sanitasi layak tidak berpengaruh signifikan. Berdasarkan Tabel 4.11, nilai Z pada parameter model logit  $\gamma_1$  yaitu  $-2,634$  kurang dari  $-Z_{(0,025)}$  yaitu  $-1,96$ . Sehingga keputusannya adalah tolak  $H_0$ . Kesimpulan yang didapatkan adalah pada model logit, variabel persentase posyandu aktif ( $X_1$ ) berpengaruh signifikan terhadap model.

#### 4.6 Model Terbaik pada Data Jumlah Kematian Bayi di Sulawesi Barat

Pemilihan model terbaik pada *Zero Inflated Poisson regression* menggunakan kriteria *Akaike's Information Criterion* (AIC) dan *Akaike's Information Corected Criterion* ( $AIC_c$ ). AIC dan  $AIC_c$  adalah kriteria kesesuaian model dalam mengestimasi secara statistik. Nilai AIC dan  $AIC_c$  pada semua kombinasi variabel ditunjukkan pada Lampiran 9. Tabel 4.12 menunjukkan nilai AIC dan  $AIC_c$  terkecil dari semua kombinasi variabel.

**Tabel 4.12** Nilai AIC pada kombinasi variabel

<b>Kombinasi Variabel</b>	<b>AIC</b>	<b><math>AIC_c</math></b>
$X_1, X_5, X_6$	324,3407	326,7407

Berdasarkan Tabel 4.12, nilai AIC dan  $AIC_c$  pada kombinasi variabel  $X_1, X_5$  dan  $X_6$  yaitu 324,3407 dan 326,7407 menunjukkan nilai AIC dan  $AIC_c$  paling kecil dibanding dengan kombinasi variabel lainnya. Sehingga dapat disimpulkan bahwa model terbaik adalah model dengan kombinasi variabel persentase posyandu aktif ( $X_1$ ), persentase penduduk terhadap akses

berkelanjutan air minum yang layak ( $X_5$ ) dan persentase penduduk dengan akses sanitasi layak ( $X_6$ ). Tabel 4.13 menunjukkan hasil estimasi parameter pada kombinasi variabel  $X_1, X_5$  dan  $X_6$ .

**Tabel 4.13** Hasil Estimasi Parameter Model Terbaik ZIP *Regression* pada Variabel  $X_1, X_5$  dan  $X_6$

Parameter	Estimasi	SE	Z	p-value
$\beta_0$	0,7344	0,2782	2,640	0,00829
$\beta_1$	0,0061	0,0024	2,491	0,01273
$\beta_5$	0,0174	0,0025	6,954	$3,54 \times 10^{-12}$
$\beta_6$	-0,0124	0,0025	-5,067	$4,03 \times 10^{-07}$
$\gamma_0$	1,4107	1,1193	1,260	0,2075
$\gamma_1$	-0,0431	0,0200	-2,155	0,0312
$\gamma_5$	0,0024	0,0237	0,099	0,9209
$\gamma_6$	-0,0384	0,0335	-1,148	0,2510

Selanjutnya dilakukan pengujian model secara serentak untuk memastikan apakah model terbaik ini telah sesuai dan dapat digunakan. Hipotesis pengujian serentak pada model log dan model logit adalah:

$H_0: \beta_1 = \beta_5 = \beta_6 = 0$  (Tidak ada variabel yang berpengaruh signifikan)

$H_1$ : paling sedikit ada satu  $\beta_j \neq 0, j = 1, 5, 6$

(Minimal ada satu variabel yang berpengaruh signifikan)

dan

$H_0: \gamma_1 = \gamma_5 = \gamma_6 = 0$  (Tidak ada variabel yang berpengaruh signifikan)

$H_1$ : paling sedikit ada satu  $\gamma_j \neq 0, j = 1, 5, 6$

(Minimal ada satu variabel yang berpengaruh signifikan)

Dengan taraf signifikansi  $\alpha=0,05$ , maka daerah penolakan yang digunakan adalah  $\chi^2_{(0,05;8)}$  yaitu 15,507 Keputusan yang dapat

diambil adalah tolak  $H_0$  karena nilai  $G^2$  yaitu 308,4 lebih dari  $\chi^2_{(0,05;8)}$  yaitu 15,507. Kesimpulan yang dapat diambil adalah model terbaik *ZIP Regression* telah sesuai dan dapat digunakan. Selanjutnya dilakukan pengujian parsial untuk mengetahui parameter mana saja yang berpengaruh signifikan terhadap model. Hipotesis yang digunakan untuk model log atau *Poisson state* adalah sebagai berikut:

$H_0: \beta_j = 0$  (Tidak ada variabel yang berpengaruh signifikan)

$H_1: \beta_j \neq 0, j = 1, 5, 6$  (Ada variabel yang berpengaruh signifikan)

Daerah penolakan yang digunakan adalah tolak  $H_0$  jika nilai  $Z > Z_{(0,025)}$  atau  $Z < -Z_{(0,025)}$  yaitu 1,96 atau -1,96. Pada Tabel 4.13 menunjukkan nilai Z parameter  $\beta_6$  yaitu sebesar -5,067 yang nilainya kurang dari  $Z < -Z_{(0,025)}$  atau -1,96. Keputusan yang dapat diambil adalah tolak  $H_0$  sehingga kesimpulannya adalah variabel  $X_6$  atau persentase penduduk terhadap akses sanitasi yang layak berpengaruh signifikan. Parameter  $\beta_1$  dan  $\beta_5$  memiliki nilai Z sebesar 2,491 dan 6,954. Nilai Z tersebut lebih dari nilai  $Z_{(0,025)}$  yaitu 1,96, sehingga keputusannya adalah tolak  $H_0$ . Sehingga dapat disimpulkan bahwa variabel persentase posyandu aktif ( $X_1$ ) dan persentase penduduk terhadap akses air minum yang layak ( $X_5$ ) berpengaruh signifikan terhadap model log. Pada model log, variabel yang berpengaruh signifikan adalah persentase posyandu aktif, persentase penduduk terhadap akses air minum yang layak dan persentase penduduk terhadap sanitasi yang layak. Hipotesis yang digunakan untuk model logit adalah:

$H_0: \gamma_j = 0$  (Tidak ada variabel yang berpengaruh signifikan)

$H_1: \gamma_j \neq 0, j = 1, 5, 6$  (Ada variabel yang berpengaruh signifikan)

Keputusan tolak  $H_0$  dapat diambil jika nilai Z lebih dari  $Z_{(0,025)}$  yaitu 1,96 atau Z kurang dari  $-Z_{(0,025)}$  yaitu -1,96.

Berdasarkan Tabel 4.13 dapat diketahui bahwa parameter  $\gamma_5$  memiliki nilai Z yaitu 0,099. Nilai tersebut kurang dari nilai  $Z_{(0,025)}$  atau 1,96, sehingga keputusannya adalah gagal tolak  $H_0$  dan kesimpulannya adalah variabel persentase penduduk terhadap akses air minum yang layak tidak berpengaruh signifikan. Parameter  $\gamma_6$  memiliki nilai Z sebesar -1,148 yang nilainya lebih dari nilai  $-Z_{(0,025)}$  yaitu -1,96, sehingga keputusannya adalah gagal tolak  $H_0$ . Kesimpulannya adalah variabel persentase penduduk dengan akses sanitasi layak atau  $X_6$  tidak berpengaruh signifikan. Parameter  $\gamma_1$  memiliki nilai Z yaitu -2,155 yang kurang dari  $-Z_{(0,025)}$  atau -1,96. Keputusan yang dapat diambil adalah tolak  $H_0$  dan kesimpulannya adalah variabel  $X_1$  atau persentase posyandu aktif berpengaruh signifikan terhadap model logit.

Sehingga model ZIP *regression* yang terbentuk adalah:

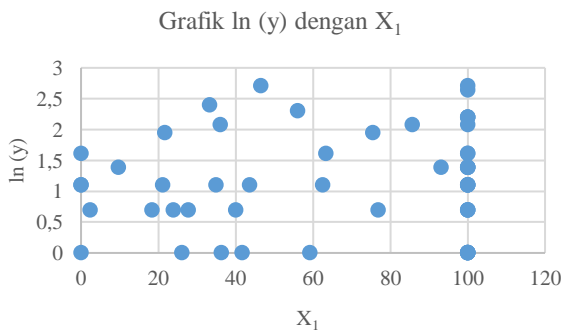
$$\log(\lambda_i) = 0,7344 + 0,0061x_{1i} + 0,0174x_{5i} - 0,0124x_{6i}$$

dan

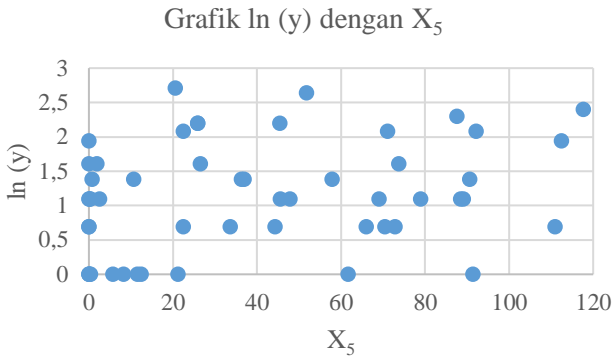
$$\text{logit}(\pi_i) = 1,411 - 0,0431x_{1i} + 0,0024x_{5i} - 0,0384x_{6i}$$

Dimana  $X_1$  menunjukkan persentase posyandu aktif,  $X_5$  adalah persentase penduduk terhadap akses air minum yang layak dan  $X_6$  atau persentase penduduk terhadap akses sanitasi layak.

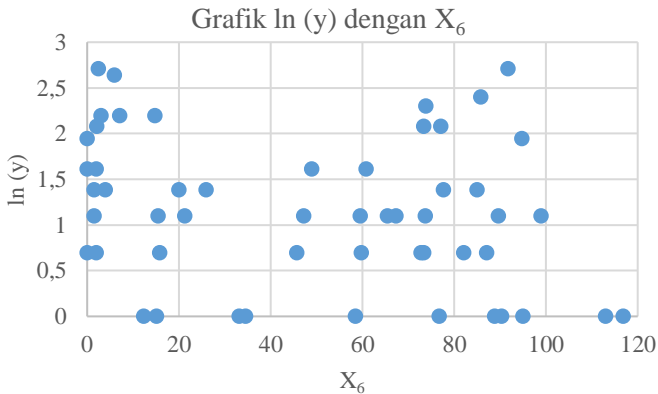
Model log menunjukkan model yang berpengaruh pada *Poisson state* dengan variabel persentase posyandu aktif, persentase penduduk terhadap akses air minum yang layak dan persentase penduduk terhadap akses sanitasi layak. Semakin meningkat persentase posyandu aktif satu satuan maka peluang kematian bayi akan meningkat sebesar  $e^{0,0061}$  atau 1,006 kali. Jika persentase penduduk terhadap akses air minum yang layak naik satu satuan maka peluang kematian bayi akan meningkat sebesar  $e^{0,0174}$  atau 1,0176 kali. Jika persentase penduduk terhadap akses sanitasi yang layak meningkat satu satuan maka akan menurunkan peluang kematian bayi sebesar  $e^{-0,0124}$  atau 0,9877 kali. Model logit menunjukkan model yang berpengaruh pada *zero state* dimana variabel yang berpengaruh signifikan adalah variabel persentase posyandu aktif. Resiko adanya kematian bayi yang berasal dari *zero state* yang memiliki posyandu aktif adalah sebesar  $e^{-0,0431}$  atau 0,9578 kali lipat dibanding dengan tidak adanya kematian bayi dari *Poisson state*. Ketidaksesuaian tanda pada model log dan logit yang terbentuk dikarenakan tidak adanya hubungan yang linier antara variabel jumlah kematian bayi dengan variabel prediktor yang berpengaruh signifikan. Berikut ini adalah grafik yang menunjukkan bahwa tidak adanya hubungan yang linier antar variabel respon dan prediktornya.



**Gambar 4.2** Grafik  $\ln(y)$  dengan  $X_1$



**Gambar 4.3** Grafik  $\ln(y)$  dengan  $X_5$



**Gambar 4.4** Grafik  $\ln(y)$  dengan  $X_6$

Gambar 4.2, 4.3 dan 4.4 menunjukkan hubungan antara variabel  $\ln(y)$  dengan ketiga variabel prediktor yang berpengaruh signifikan menunjukkan tidak ada hubungan yang linier. Hal tersebut dikarenakan sebaran data yang diwakili titik-titik biru tidak membentuk pola tertentu. Sehingga tanda pada estimasi parameternya tidak sesuai dengan teori yang ada.



## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Berdasarkan analisis dan pembahasan yang telah dilakukan pada bab 4, maka diperoleh kesimpulan sebagai berikut.

1. Karakteristik faktor-faktor yang diduga mempengaruhi jumlah kematian bayi di Sulawesi Barat adalah rata-rata persentase posyandu aktif sebesar 60,75%. Variabel persentase RT yang ber-PHBS memiliki rata-rata sebesar 67,47%, persentase desa/kelurahan UCI yaitu sebesar 63,16% dan rata-rata persentase ibu bersalin ditolong tenaga kesehatan sebesar 77,60%. Persentase penduduk yang memiliki akses terhadap air minum yang layak yaitu 38,89% dan persentase penduduk yang memiliki akses terhadap sanitasi layak sebesar 49,52%. Ada dua kecamatan yang memiliki jumlah kematian bayi tertinggi sebanyak 15 kematian bayi yaitu kecamatan Banggae dan Pasangkayu. 16 kecamatan tidak terjadi kematian bayi selama tahun 2016 yaitu Tapoyo, Pangale, Balabalakang, Kalumpang, Bonehau, Sampaga, Papalang, Mambi, Bambang, Rantim, Tabang, Nosu, Matangnga, Balanipa, Luyo dan Mapili.
2. Hasil ZIP *regression* dengan metode terbaik adalah pada kombinasi variabel  $X_1$ ,  $X_5$  dan  $X_6$  atau persentase posyandu aktif, persentase penduduk dengan akses terhadap air minum yang layak dan persentase penduduk dengan akses sanitasi layak. Pada model log atau *Poisson state* didapatkan variabel yang berpengaruh signifikan adalah persentase posyandu aktif, persentase penduduk terhadap akses air minum yang layak dan persentase penduduk terhadap sanitasi layak. Pada model logit atau *zero state* didapatkan variabel yang berpengaruh signifikan adalah persentase posyandu aktif.

## 5.2 Saran

Berdasarkan kesimpulan yang diperoleh, dapat dirumuskan saran sebagai pertimbangan penelitian selanjutnya adalah sebagai berikut.

1. Sebaiknya dinas kesehatan provinsi Sulawesi Barat menambah variabel prediktor dalam segi sosial ekonomi, karena variabel tersebut sulit diperoleh padahal variabel tersebut juga mempengaruhi jumlah kematian bayi.
2. Sebaiknya dinas kesehatan provinsi Sulawesi Barat lebih fokus pada perbaikan fasilitas pelayanan posyandu aktif dan meningkatkan kualitas persalinan dengan tenaga kesehatan.

## DAFTAR PUSTAKA

- Badan Perencanaan Pembangunan Nasional. (2009). *Faktor-Faktor yang Mempengaruhi Kelangsungan Hidup Anak*, Jakarta: Badan Perencanaan Pembangunan Nasional.
- Badan Pusat Statistik. (2019). *Sistem Informasi Rujukan Statistik*. Retrieved Februari 12, 2019, from <https://sirusa.bps.go.id/index.php?r=indikator/view&id=79>.
- Bozdogan, H. (2000). Akaike's Information Criterion and Recent Developments in Information Complexity. *Journal of Mathematical Psychology*, Vol. 44, 62-91.
- Cameron, A. C., & Trivedi, P. K. (2013). *Regression Analysis of Count Data Second Edition*. United States of America: Cambridge University Press.
- Dinas Kesehatan Kabupaten Mamasa. (2017). *Profil Kesehatan Kabupaten Mamasa 2016*. Kabupaten Mamasa: Dinas Kesehatan Kabupaten Mamasa.
- Dinas Kesehatan Kabupaten Mamuju. (2017). *Profil Kesehatan Kabupaten Mamuju 2016*. Kabupaten Mamuju: Dinas Kesehatan Kabupaten Mamuju.
- Dinas Kesehatan Mamuju Tengah. (2017). *Profil Kesehatan Kabupaten Mamuju Tengah 2016*. Kabupaten Mamuju Tengah: Dinas Kesehatan Kabupaten Mamuju Tengah.
- Dinas Kesehatan Mamuju Utara. (2017). *Profil Kesehatan Kabupaten Mamuju Utara*. Kabupaten Mamuju Utara: Dinas Kesehatan Kabupaten Mamuju Utara.
- Dinas Kesehatan Kabupaten Majene. (2017). *Profil Kesehatan Kabupaten Majene Tahun 2016*. Majene: Dinas Kesehatan Kabupaten Majene.
- Dinas Kesehatan Kabupaten Polewali Mandar. (2017). *Profil Kesehatan Polewali Mandar 2016*. Kabupaten Polewali Mandar: Dinas Kesehatan Kabupaten Polewali Mandar.
- Famoye, F., & Singh, K. P. (2006). Zero Inflated Poisson Regression Model with an Applications Domestic

- Violence to Accident Data. *Journal of Data Science*, Vol. 4, 117-130.
- Filmer, D. (2003). *Determinants of Health & Education Outcomes Background Note for World Development Report 2004: Making Service Work for Poor People*. World Bank.
- Hall, D. B., & Shen, J. (2009). Robust Estimation for Zero Inflated Poisson Regression. *Scandinavian Journal of Statistics*, Vol. 44, 62-91.
- Handayani, S., & Pratiwi, N. L. (2016). Health Seeking Behaviour pada Persalinan Suku Kaili Da'a di Desa Wulai Kabupaten Mamuju Utara Sulawesi Barat, *Buletin Penelitian Sistem Kesehatan*, Vol.19, 91-98.
- Hilbe, J. M. (2011). *Negative Binomial Regression*, New York: Cambridge University Press.
- Hinde, J., & Demetrio, C. B. (2007). *Overdispersi: Model and Elimination*, Exeter: Departement of Laver Building.
- Kementrian Kesehatan Indonesia. (2015). *Profil Kesehatan Indonesia 2014*. Jakarta: Kementrian Kesehatan Republik Indonesia.
- Lambert, D. (1992). Zero-Inflated Poisson Regression, With an Application to Defect in Manufacturing, *American Statistical Association and the American Society for Quality Control*, Vol. 34, 1-14.
- Lestari, A. (2008). *Pemodelan Regresi Zero Inflated Poisson (Aplikasi pada Data Pekerja Seks Komersial di Klinik Reproduksi Putat Jaya Surabaya)*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Mosley, W. H., & Chen, L. C. (2003). An Analytical Framework for the Study of Child Survival in Developing Countries. *Population and Development Review*, Vol. 80, 140-145.
- Putri, M. P., & Purhadi. (2017). Analisis Faktor-Faktor yang Berpengaruh Terhadap Jumlah Kematian Ibu dan Jumlah Kematian Bayi di Provinsi Jawa Tengah. *Jurnal Sains dan Seni ITS*, Vol. 06, 108-114.

- Setiawan, & Kusriani, D. E. (2010). *Ekonometrika*. Yogyakarta: C.V Andi Offset.
- Suryani, L. E. (2018). *Analisis Faktor-Faktor yang Mempengaruhi Jumlah Kematian Bayi dan Ibu di Jawa Timur dengan Metode Bivariate Generalized Poisson Regression*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Walpole, R. E., Myres, R. H., Myres, S. L., & Ye, K. (2007). *Probability & Statistics for Engineers & Scientists 8th Edition*. United States of America: Prentice Hall.

*(Halaman ini sengaja dikosongkan)*

## LAMPIRAN

**Lampiran 1.** Data Kematian Bayi dan Faktor-Faktor yang Diduga Mempengaruhi

No	Kecamatan	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
1	Banggae	15	46,512	81,991	62,500	78,28	120,39	91,72
2	Banggae Timur	7	75,556	85,238	33,333	65,19	112,50	94,80
3	Pamboang	3	62,500	80,500	60,000	79,52	89,05	73,68
4	Sendana	11	33,300	71,600	62,500	77,07	117,65	85,84
5	Tubo Sendana	8	85,700	67,900	28,600	93,59	71,10	73,32
6	Tammerodo	8	36,100	71,900	85,700	81,33	92,14	77,11
7	Malunda	10	56,100	67,900	75,000	86,83	87,53	73,80
8	Ulumanda	3	21,212	64,758	37,500	91,1	88,49	59,58
9	Mamasa	3	100,000	100,000	25,000	84,27	0,15	21,23
10	Tawalian	1	100,000	100,000	0,000	60,23	0,41	116,84
11	Sesenapadang	4	100,000	100,000	20,000	60,96	0,70	20,02
12	Balla	2	100,000	100,000	100,000	69,33	0,00	73,35

No	Kecamatan	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
13	Tandukkalua	5	100,000	100,000	66,700	77,04	1,89	48,98
14	Sumarorong	1	100,000	100,000	80,000	73,44	5,70	58,52
15	Messawa	3	100,000	100,000	11,100	59,65	0,34	89,67
16	Nosu	0	100,000	100,000	85,700	69,52	0,00	2,43
17	Pana'	1	100,000	100,000	61,500	71,5	0,00	15,16
18	Tabang	0	100,000	100,000	28,600	59,72	1,12	23,73
19	Rantim	0	0,000	0,000	12,500	46,21	0,00	7,09
20	Bambang	0	100,000	100,000	65,000	60,63	0,00	137,37
21	Aralle	1	100,000	100,000	25,000	64,38	0,10	34,59
22	Tabulahan	4	100,000	100,000	28,600	55,74	10,65	25,88
23	Mambi	0	100,000	100,000	38,500	80,52	0,49	99,43
24	Bumal	1	100,000	100,000	54,500	43,79	0,00	12,30
25	Mehalaan	3	100,000	100,000	54,500	90,82	0,00	15,51
26	Tapalang	3	0,000	86,923	100,000	80,76	45,60	47,24
27	Tapalang barat	1	0,000	77,143	42,857	88,21	91,38	33,13



No	Kecamatan	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
28	Simboro	5	0,000	31,613	37,500	82,45	73,78	60,84
29	Mamuju	2	2,326	27,671	87,500	69,84	72,86	59,80
30	Kalukku	2	18,391	56,766	57,143	72,61	65,95	45,70
31	Papalang	0	0,000	73,016	77,778	91,28	80,45	58,90
32	Sampaga	0	41,176	54,762	100,000	74,57	79,27	57,66
33	Tommo	3	43,590	53,761	71,429	71,67	69,08	67,33
34	Bonehau	0	0,000	29,876	33,333	46,59	83,24	42,69
35	Kalumpang	0	0,000	11,180	100,000	71,01	61,21	32,97
36	Balabalakang	0	0,000	33,636	0,000	56,52	55,11	14,06
37	Pangale	0	0,000	43,026	22,222	103,16	0,00	0,00
38	Budong-budong	7	21,739	16,738	90,909	93,4	0,00	0,00
39	Tobadak	2	76,923	52,342	87,500	85,57	0,00	0,00
40	Tapoyo	0	13,333	160,750	66,667	87,29	0,00	0,00
41	Karossa	5	63,415	71,111	30,769	86,4	0,00	0,00
42	Sarudu	9	100,000	58,214	83,333	70,51	45,42	7,14

No	Kecamatan	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
43	Dapurang	5	100,000	56,637	40,000	90,07	26,55	1,99
44	Duripoku	3	100,000	77,583	66,667	81,72	47,88	1,47
45	Lariang	4	100,000	25,565	85,714	100,65	36,90	3,91
46	Bulu Taba	2	100,000	26,075	57,143	89,92	33,67	15,80
47	Baras	9	100,000	60,985	50,000	78,93	25,94	14,81
48	Pasangkayu	15	100,000	54,551	37,500	91,51	20,56	2,50
49	Tikke Raya	9	100,000	40,490	125,000	71,51	25,93	3,08
50	Pedongga	2	100,000	74,060	33,333	82,35	22,47	1,97
51	Bambalamotu	14	100,000	93,548	75,000	77,35	51,83	5,99
52	Bambaira	4	100,000	67,776	100,000	86,49	36,34	1,52
53	Sarjo	8	100,000	57,958	66,667	76,54	22,43	2,13
54	Binuang	2	23,913	81,650	90,000	82,21	44,23	72,78
55	Polewali	3	34,884	76,792	100,000	84,48	78,88	65,43
56	Anreapi	1	36,364	67,380	100,000	93,42	8,19	95,02
57	Matakali	4	93,103	50,876	100,000	86,18	90,64	84,99

No	Kecamatan	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
58	Tapango	4	9,756	50,371	71,429	74,29	57,89	77,72
59	Wonomulyo	2	40,000	79,347	100,000	77,57	110,96	87,03
60	Mapili	0	30,612	87,521	83,333	80,34	60,48	91,87
61	Luyo	0	86,842	77,349	36,364	81,94	42,54	78,15
62	Campalagian	2	27,778	92,161	94,444	85,74	70,42	82,13
63	Balanipa	0	33,333	54,369	54,545	98,08	30,37	90,39
64	Tinambung	1	59,259	35,338	75,000	84,77	61,72	76,75
65	Limboro	1	41,667	49,196	90,909	65,67	21,13	90,35
66	Allu	1	100,000	43,228	100,000	77,16	11,60	88,87
67	Tubbi Taramanu	3	0,000	20,819	100,000	84	2,52	98,91
68	Matangnga	0	50,000	26,911	57,143	75	5,11	108,60
69	Bulo	1	26,087	30,284	77,778	84,19	12,41	112,97

## Lampiran 2. Output Statistika Deskriptif

```
MTB > Describe 'x1' - 'x6';
SUBC> Mean;
SUBC> Variance;
SUBC> Minimum;
SUBC> Maximum.
```

### Descriptive Statistics: x1; x2; x3; x4; x5; x6

Variable	Mean	Variance	Minimum	Maximum
x1	60,75	1578,07	0,00	100,00
x2	67,96	865,38	0,00	160,75
x3	63,18	861,98	0,00	125,00
x4	77,60	160,19	43,79	103,16
x5	38,89	1366,44	0,00	120,39
x6	49,52	1468,95	0,00	137,37

## Lampiran 3. Output Pengujian Distribusi Poisson

### One-Sample Kolmogorov-Smirnov Test

		y
N		69
Poisson Parameter <sup>a,b</sup>	Mean	3.3768
Most Extreme Differences	Absolute	.242
	Positive	.242
	Negative	-.137
Kolmogorov-Smirnov Z		2.009
Asymp. Sig. (2-tailed)		.001

a. Test distribution is Poisson.

b. Calculated from data.

#	<u>Distribution</u>	<u>Kolmogorov Smirnov</u>		<u>Anderson Darling</u>	
		Statistic	Rank	Statistic	Rank
1	<u>D. Uniform</u>	0,30769	4	25,67	4
2	<u>Geometric</u>	0,22848	1	3,3195	1
3	<u>Neg. Binomial</u>	0,24625	3	4,6568	2
4	<u>Poisson</u>	0,24181	2	11,833	3
5	Bernoulli	No fit (data max > 1)			
6	Binomial	No fit			
7	Hypergeometric	No fit			
8	Logarithmic	No fit (data min < 1)			

## Lampiran 4. Output Pengecekan Overdispersi

```
MTB > Describe 'y';  
SUBC> Mean;  
SUBC> Variance.
```

### Descriptive Statistics: y

Variable	Mean	Variance
y	3,377	13,915

```
> rp<-glm(formula = y~ x1+x2+x3+x4+x5+x6, data=data, fa  
family=poisson)  
> summary(rp)
```

Call:

```
glm(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6, family =  
poisson,  
data = data)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.7877	-1.6496	-0.3561	0.9915	3.2304

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-0.8790286	0.6252370	-1.406	0.159750
x1	0.0127826	0.0022496	5.682	1.33e-08
x2	-0.0009223	0.0030321	-0.304	0.760993
x3	-0.0006796	0.0025307	-0.269	0.788285
x4	0.0117314	0.0063068	1.860	0.062871
x5	0.0184049	0.0024130	7.628	2.39e-14
x6	-0.0087651	0.0023929	-3.663	0.000249

(Intercept)

x1 \*\*\*  
x2  
x3  
x4 .  
x5 \*\*\*  
x6 \*\*\*

Signif. codes:

0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 262.83 on 68 degrees of freedom  
Residual deviance: 166.52 on 62 degrees of freedom  
AIC: 343.08

Number of Fisher Scoring iterations: 6

### Lampiran 5. Output Pengecekan Asumsi Multikolinearitas

MTB > Correlation 'x1'-'x6'.

#### Correlation: x1; x2; x3; x4; x5; x6

	x1	x2	x3	x4	x5
x2	0,407 0,001				
x3	-0,125 0,306	-0,155 0,202			
x4	-0,109 0,371	-0,117 0,339	0,275 0,022		
x5	-0,415 0,000	-0,172 0,159	0,189 0,121	0,131 0,285	
x6	-0,213 0,079	0,028 0,819	0,141 0,249	-0,049 0,691	0,293 0,015

Cell Contents: Pearson correlation  
P-Value

```

MTB > Regress;
SUBC> Response 'x1';
SUBC> Nodefault;
SUBC> Continuous 'x2' - 'x6';
SUBC> Terms x2 x3 x4 x5 x6;
SUBC> Constant;
SUBC> Unstandardized;
SUBC> Tmethod;
SUBC> Tanova;
SUBC> Tsummary;
SUBC> Tcoefficients;
SUBC> Tequation;
SUBC> Tdiagnostics 0.

```

### Regression Analysis: x1 versus x2; x3; x4; x5; x6

#### Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	5	32812	6562,3	5,55	0,000
x2	1	12819	12819,3	10,84	0,002
x3	1	35	34,6	0,03	0,865
x4	1	147	146,5	0,12	0,726
x5	1	9019	9019,2	7,63	0,008
x6	1	1762	1762,1	1,49	0,227
Error	63	74497	1182,5		
Total	68	107309			

#### Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
34,3875	30,58%	25,07%	10,16%

#### Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	56,0	29,7	1,89	0,064	
x2	0,481	0,146	3,29	0,002	1,06
x3	0,026	0,152	0,17	0,865	1,14
x4	-0,122	0,347	-0,35	0,726	1,11
x5	-0,336	0,122	-2,76	0,008	1,17
x6	-0,141	0,116	-1,22	0,227	1,13

#### Regression Equation

$$x1 = 56.0 + 0.481 x2 + 0.026 x3 - 0.122 x4 - 0.336 x5 - 0.141 x6$$

#### Fits and Diagnostics for Unusual Observations

Obs	x1	Fit	Resid	Std Resid	
26	0,0	68,5	-68,5	-2,04	R
37	0,0	64,6	-64,6	-2,06	R
40	13,3	124,3	-111,0	-3,64	R

R Large residual



```

MTB > Regress;
SUBC> Response 'x2';
SUBC> Nodefault;
SUBC> Continuous 'x1' 'x3' - 'x6';
SUBC> Terms x1 x3 x4 x5 x6;
SUBC> Constant;
SUBC> Unstandardized;
SUBC> Tmethod;
SUBC> Tanova;
SUBC> Tsummary;
SUBC> Tcoefficients;
SUBC> Tequation;
SUBC> TDiagnosics 0.

```

### Regression Analysis: x2 versus x1; x3; x4; x5; x6

#### Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	5	11493,5	2298,70	3,06	0,016
x1	1	8148,2	8148,20	10,84	0,002
x3	1	643,6	643,55	0,86	0,358
x4	1	56,6	56,56	0,08	0,785
x5	1	9,2	9,20	0,01	0,912
x6	1	935,7	935,73	1,24	0,269
Error	63	47352,1	751,62		
Total	68	58845,6			

#### Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
27,4157	19,53%	13,15%	0,26%

#### Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	57,7	23,2	2,48	0,016	
x1	0,3055	0,0928	3,29	0,002	1,23
x3	-0,111	0,120	-0,93	0,358	1,13
x4	-0,076	0,277	-0,27	0,785	1,11
x5	-0,011	0,103	-0,11	0,912	1,31
x6	0,1030	0,0923	1,12	0,269	1,13

#### Regression Equation

$$x2 = 57.7 + 0.3055 x1 - 0.111 x3 - 0.076 x4 - 0.011 x5 + 0.1030 x6$$

#### Fits and Diagnostics for Unusual Observations

Obs	x2	Fit	Resid	Std Resid	
19	0,00	53,50	-53,50	-2,22	R
40	160,75	47,69	113,06	4,36	R

R Large residual

```

MTB > Regress;
SUBC> Response 'x3';
SUBC> Nodefault;
SUBC> Continuous 'x1' 'x2' 'x4' 'x5' 'x6';
SUBC> Terms x1 x2 x4 x5 x6;
SUBC> Constant;
SUBC> Unstandardized;
SUBC> Tmethod;
SUBC> Tanova;
SUBC> Tsummary;
SUBC> Tcoefficients;
SUBC> Tequation;
SUBC> Tdiagnostics 0.

```

### Regression Analysis: x3 versus x1; x2; x4; x5; x6

#### Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	5	7314,5	1462,91	1,80	0,126
x1	1	23,9	23,86	0,03	0,865
x2	1	697,2	697,21	0,86	0,358
x4	1	3722,9	3722,87	4,57	0,036
x5	1	511,3	511,30	0,63	0,431
x6	1	880,8	880,76	1,08	0,302
Error	63	51299,8	814,28		
Total	68	58614,4			

#### Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
28,5356	12,48%	5,53%	0,00%

#### Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	15,9	25,2	0,63	0,532	
x1	0,018	0,105	0,17	0,865	1,44
x2	-0,121	0,130	-0,93	0,358	1,23
x4	0,595	0,278	2,14	0,036	1,04
x5	0,084	0,106	0,79	0,431	1,29
x6	0,1000	0,0962	1,04	0,302	1,13

#### Regression Equation

$x3 = 15.9 + 0.018 x1 - 0.121 x2 + 0.595 x4 + 0.084 x5 + 0.1000 x6$

#### Fits and Diagnostics for Unusual Observations

Obs	x3	Fit	Resid	Std Resid	
19	12,5	44,1	-31,6	-1,30	X
40	66,7	48,7	18,0	0,80	X
49	125,0	57,8	67,2	2,45	R

R Large residual  
X Unusual X

```

MTB > Regress;
SUBC> Response 'x4';
SUBC> Nodefault;
SUBC> Continuous 'x1' - 'x3' 'x5' - 'x6';
SUBC> Terms x1 x2 x3 x5 x6;
SUBC> Constant;
SUBC> Unstandardized;
SUBC> Tmethod;
SUBC> Tanova;
SUBC> Tsummary;
SUBC> Tcoefficients;
SUBC> Tequation;
SUBC> TDiagnostics 0.

```

### Regression Analysis: x4 versus x1; x2; x3; x5; x6

#### Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	5	1089,4	217,87	1,40	0,236
x1	1	19,3	19,28	0,12	0,726
x2	1	11,7	11,71	0,08	0,785
x3	1	711,4	711,44	4,57	0,036
x5	1	66,7	66,71	0,43	0,515
x6	1	141,6	141,64	0,91	0,344
Error	63	9803,4	155,61		
Total	68	10892,7			

#### Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
12,4743	10,00%	2,86%	0,00%

#### Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	73,26	6,10	12,00	0,000	
x1	-0,0161	0,0457	-0,35	0,726	1,44
x2	-0,0157	0,0573	-0,27	0,785	1,24
x3	0,1137	0,0532	2,14	0,036	1,07
x5	0,0305	0,0466	0,65	0,515	1,30
x6	-0,0402	0,0421	-0,95	0,344	1,14

#### Regression Equation

$$x4 = 73.26 - 0.0161 x1 - 0.0157 x2 + 0.1137 x3 + 0.0305 x5 - 0.0402 x6$$

#### Fits and Diagnostics for Unusual Observations

Obs	x4	Fit	Resid	Std Resid	
19	46,21	74,40	-28,19	-2,54	R
24	43,79	75,79	-32,00	-2,64	R
34	46,59	77,41	-30,82	-2,60	R
37	103,16	75,12	28,04	2,43	R
40	87,29	78,11	9,18	0,93	X

R Large residual

X Unusual X

```

MTB > Regress;
SUBC> Response 'x5';
SUBC> Nodefault;
SUBC> Continuous 'x1' - 'x4' 'x6';
SUBC> Terms x1 x2 x3 x4 x6;
SUBC> Constant;
SUBC> Unstandardized;
SUBC> Tmethod;
SUBC> Tanova;
SUBC> Tsummary;
SUBC> Tcoefficients;
SUBC> Tequation;
SUBC> Tdiagnostics 0.

```

### Regression Analysis: x5 versus x1; x2; x3; x4; x6

#### Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	5	21788,7	4357,75	3,86	0,004
x1	1	8611,5	8611,50	7,63	0,008
x2	1	13,8	13,82	0,01	0,912
x3	1	708,9	708,94	0,63	0,431
x4	1	484,1	484,05	0,43	0,515
x6	1	3738,1	3738,08	3,31	0,074
Error	63	71129,5	1129,04		
Total	68	92918,2			

#### Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
33,6012	23,45%	17,37%	5,87%

#### Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	24,9	29,6	0,84	0,403	
x1	-0,321	0,116	-2,76	0,008	1,28
x2	-0,017	0,154	-0,11	0,912	1,24
x3	0,117	0,148	0,79	0,431	1,13
x4	0,221	0,338	0,65	0,515	1,10
x6	0,203	0,111	1,82	0,074	1,10

#### Regression Equation

$$x5 = 24.9 - 0.321 x1 - 0.017 x2 + 0.117 x3 + 0.221 x4 + 0.203 x6$$

#### Fits and Diagnostics for Unusual Observations

Obs	x5	Fit	Resid	Std Resid	
1	120,4	51,8	68,5	2,08	R
2	112,5	36,8	75,7	2,33	R
19	0,0	38,1	-38,1	-1,33	X
40	0,0	45,1	-45,1	-1,66	X
67	2,5	74,9	-72,4	-2,26	R

R Large residual  
X Unusual X

```

MTB > Regress;
SUBC> Response 'x6';
SUBC> Nodefault;
SUBC> Continuous 'x1' - 'x5';
SUBC> Terms x1 x2 x3 x4 x5;
SUBC> Constant;
SUBC> Unstandardized;
SUBC> Tmethod;
SUBC> Tanova;
SUBC> Tsummary;
SUBC> Tcoefficients;
SUBC> Tequation;
SUBC> TDiagnosics 0.

```

### Regression Analysis: x6 versus x1; x2; x3; x4; x5

#### Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	5	13350	2670	1,94	0,100
x1	1	2047	2047	1,49	0,227
x2	1	1710	1710	1,24	0,269
x3	1	1486	1486	1,08	0,302
x4	1	1250	1250	0,91	0,344
x5	1	4548	4548	3,31	0,074
Error	63	86538	1374		
Total	68	99889			

#### Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
37,0624	13,37%	6,49%	0,00%

#### Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	53,9	32,2	1,68	0,098	
x1	-0,164	0,134	-1,22	0,227	1,41
x2	0,188	0,169	1,12	0,269	1,22
x3	0,169	0,162	1,04	0,302	1,12
x4	-0,355	0,372	-0,95	0,344	1,10
x5	0,246	0,135	1,82	0,074	1,24

#### Regression Equation

$$x6 = 53.9 - 0.164 x1 + 0.188 x2 + 0.169 x3 - 0.355 x4 + 0.246 x5$$

#### Fits and Diagnostics for Unusual Observations

Obs	x6	Fit	Resid	Std Resid	
10	116,8	35,1	81,7	2,33	R
19	7,1	39,7	-32,6	-1,03	X
20	137,4	45,9	91,5	2,56	R
40	0,0	62,3	-62,3	-2,06	R X
68	108,6	35,1	73,5	2,05	R

R Large residual  
X Unusual X

Perhitungan Nilai VIF

<b>Variabel</b>	<b>Nilai VIF</b>
$X_1$	1,440507
$X_2$	1,242699
$X_3$	1,142596
$X_4$	1,111111
$X_5$	1,306336
$X_6$	1,154335

**Lampiran 6. Syntax Zero Inflated Poisson Regression**

```
data=read.csv("E:/x1-x6.csv", header = TRUE, sep =
";")
data
library(ggplot2)
library(pscl)
library(boot)

output<- zeroinfl(y ~ x1+x2+x3+x4+x5+x6,
data=data,
                    dist = "poisson",
                    link = "logit",
                    control =
zeroinfl.control(method="BFGS", maxit = 10000,
                    trace
= FALSE, EM = TRUE, start = NULL),
                    model=TRUE, y=TRUE, x=TRUE)

summary (output)

logLik(output)
library(AICcmodavg)
AIC(output)
AICc(output)
a=summary(output)
a$coefficients
```

## Lampiran 7. Output Zero Inflated Poisson Regression

```
> output<- zeroinfl(y ~ x1+x2+x3+x4+x5+x6, data=data,
+                   dist = "poisson",
+                   link = "logit",
+                   control = zeroinfl.control(method
="BFGS", maxit = 10000,
+                                           trace
= FALSE, EM = TRUE, start = NULL),
+                   model=TRUE, y=TRUE, x=TRUE)
> summary (output)
```

Call:

```
zeroinfl(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6, data = data, dist = "poisson", link = "logit", control = zeroinfl.control(method = "BFGS", maxit = 10000, trace = FALSE, EM = TRUE, start = NULL), model = TRUE, y = TRUE, x = TRUE)
```

Pearson residuals:

	Min	1Q	Median	3Q	Max
	-1.8236	-0.9280	-0.3042	0.6590	3.7580

Count model coefficients (poisson with log link):

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	1.143034	0.815560	1.402	0.16106	
x1	0.006714	0.002523	2.661	0.00779	**
x2	-0.003480	0.003655	-0.952	0.34111	
x3	-0.001653	0.002767	-0.597	0.55021	
x4	-0.001497	0.007519	-0.199	0.84222	
x5	0.017433	0.002467	7.067	1.59e-12	***
x6	-0.012110	0.002503	-4.837	1.32e-06	***

Zero-inflation model coefficients (binomial with logit link):

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	6.6776366	3.6339900	1.838	0.06613	.
x1	-0.0512201	0.0194456	-2.634	0.00844	**
x2	0.0161480	0.0201419	0.802	0.42272	
x3	0.0064817	0.0206941	0.313	0.75412	
x4	-0.0754716	0.0408851	-1.846	0.06490	.
x5	-0.0004884	0.0233262	-0.021	0.98330	
x6	-0.0495579	0.0358248	-1.383	0.16656	

---



### Lampiran 7. Output Zero Inflated Poisson Regression (lanjutan)

```

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Number of iterations in BFGS optimization: 1
Log-likelihood: -151.5 on 14 Df
> logLik(output)
'log Lik.' -151.4828 (df=14)
> library(AICcmoavg)
> AIC(output)
[1] 330.9656
> AICc(output)
[1] 338.7434
> a=summary(output)
> a$coefficients
$count
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept)  1.143033504  0.815560222  1.4015317 1.610551e-01
x1            0.006714366  0.002523041  2.6612192 7.785824e-03
x2           -0.003479808  0.003655379 -0.9519691 3.411126e-01
x3           -0.001653157  0.002767050 -0.5974439 5.502110e-01
x4           -0.001496706  0.007518958 -0.1990576 8.422177e-01
x5            0.017432749  0.002466937  7.0665574 1.588245e-12
x6           -0.012109576  0.002503415 -4.8372236 1.316653e-06

$zero
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept)  6.6776366117  3.63398995  1.83754955 0.066128808
x1          -0.0512201141  0.01944556 -2.63402563 0.008437912
x2           0.0161480475  0.02014188  0.80171486 0.422717917
x3           0.0064817086  0.02069413  0.31321489 0.754117408
x4          -0.0754716195  0.04088510 -1.84594453 0.064900261
x5          -0.0004883825  0.02332620 -0.02093708 0.983295849
x6          -0.0495578580  0.03582481 -1.38333922 0.166560876

```

### Lampiran 8. Output AIC dan AICc

Kombinasi Variabel	AIC	AICc
$X_1$	375,8579	376,4829
$X_2$	382,9717	383,5967
$X_3$	381,359	381,984
$X_4$	378,1397	378,7647
$X_5$	366,1204	366,7454
$X_6$	377,1297	377,7547
$X_1, X_2$	377,2105	378,5654
$X_1, X_3$	376,6595	378,0143
$X_1, X_4$	371,8804	373,2352
$X_1, X_5$	343,2513	344,7061
$X_1, X_6$	374,8477	376,2025
$X_2, X_3$	383,1728	384,5276
$X_2, X_4$	381,9379	383,2928
$X_2, X_5$	369,1775	370,5324
$X_2, X_6$	380,323	381,6778
$X_3, X_4$	379,7854	381,1403
$X_3, X_5$	367,4594	368,8142
$X_3, X_6$	378,7339	380,0887
$X_4, X_5$	366,6967	368,0515
$X_4, X_6$	376,8416	378,1965
$X_5, X_6$	340,433	341,7879
$X_1, X_2, X_3$	377,386	379,786

<b>Kombinasi Variabel</b>	<b>AIC</b>	<b>AIC<sub>c</sub></b>
$X_1, X_2, X_4$	374,9669	377,3669
$X_1, X_2, X_5$	343,9317	346,3317
$X_1, X_2, X_6$	377,4691	379,8691
$X_1, X_3, X_4$	373,5285	375,9285
$X_1, X_3, X_5$	344,9873	347,3873
$X_1, X_3, X_6$	376,0521	378,4521
$X_1, X_4, X_5$	341,6211	344,0211
$X_1, X_4, X_6$	373,5583	375,9583
$X_1, X_5, X_6$	324,3407	326,7407
$X_2, X_3, X_4$	383,3092	385,7092
$X_2, X_3, X_5$	368,2778	370,6778
$X_2, X_3, X_6$	381,0381	383,4381
$X_2, X_4, X_5$	370,0657	372,4657
$X_2, X_4, X_6$	380,5989	382,9989
$X_2, X_5, X_6$	342,9201	346,3201
$X_3, X_4, X_5$	368,0719	370,4719
$X_3, X_4, X_6$	379,0845	381,4845
$X_3, X_5, X_6$	342,6853	345,0853
$X_4, X_5, X_6$	341,4995	343,8995
$X_1, X_2, X_3, X_4$	376,509	380,3021
$X_1, X_2, X_3, X_5$	344,438	348,2312
$X_1, X_2, X_3, X_6$	378,2293	382,0224
$X_1, X_2, X_4, X_5$	343,9268	347,7199

<b>Kombinasi Variabel</b>	<b>AIC</b>	<b>AIC<sub>c</sub></b>
$X_1, X_2, X_4, X_6$	376,7797	380,5728
$X_1, X_2, X_5, X_6$	327,1651	330,9582
$X_1, X_3, X_4, X_5$	343,9376	347,7307
$X_1, X_3, X_4, X_6$	375,4525	379,2456
$X_1, X_3, X_5, X_6$	328,1157	331,9088
$X_1, X_4, X_5, X_6$	324,9738	328,7669
$X_2, X_3, X_4, X_5$	370,1952	373,9883
$X_2, X_3, X_4, X_6$	382,5422	386,3353
$X_2, X_3, X_5, X_6$	344,2192	348,0123
$X_2, X_4, X_5, X_6$	344,4237	348,2168
$X_3, X_4, X_5, X_6$	345,0175	348,8106
$X_1, X_2, X_3, X_4, X_5$	345,8872	351,4587
$X_1, X_2, X_3, X_4, X_6$	378,6671	384,2386
$X_1, X_2, X_3, X_5, X_6$	330,6722	336,2436
$X_1, X_2, X_3, X_4, X_6$	378,6671	384,6671
$X_1, X_3, X_4, X_5, X_6$	328,8206	334,3921
$X_2, X_3, X_4, X_5, X_6$	346,8712	352,4426
$X_1, X_2, X_3, X_4, X_5, X_6$	330,9656	338,7434

### Lampiran 9. Output Model Terbaik

```
> output<- zeroinfl(y ~ x1+x5+x6, data=data,
+                   dist = "poisson",
+                   link = "logit",
+                   control = zeroinfl.control(method=
"BFSG", maxit = 10000,
+                                           trace =
FALSE, EM = TRUE, start = NULL),
+                   model=TRUE, y=TRUE, x=TRUE)
> summary (output)
```

Call:

```
zeroinfl(formula = y ~ x1 + x5 + x6, data = data, dist
= "poisson", link = "logit", control = zeroinfl.contr
ol(method = "BFSG",
maxit = 10000, trace = FALSE, EM = TRUE, start = N
ULL), model = TRUE, y = TRUE, x = TRUE)
```

Pearson residuals:

	Min	1Q	Median	3Q	Max
	-1.8227	-0.8890	-0.4055	0.8719	4.0559

Count model coefficients (poisson with log link):

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	0.734412	0.278171	2.640	0.00829	**
x1	0.006088	0.002444	2.491	0.01273	*
x5	0.017419	0.002505	6.954	3.54e-12	***
x6	-0.012444	0.002456	-5.067	4.03e-07	***

Zero-inflation model coefficients (binomial with logit link):

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	1.410689	1.119298	1.260	0.2075	
x1	-0.043123	0.020012	-2.155	0.0312	*
x5	0.002353	0.023695	0.099	0.9209	
x6	-0.038447	0.033490	-1.148	0.2510	

---

## Lampiran 10. Output Model Terbaik (lanjutan)

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Number of iterations in BFGS optimization: 1

Log-likelihood: -154.2 on 8 Df

```
> LogLik(output)
```

```
'log Lik.' -154.1704 (df=8)
```

```
> library(AICcmodavg)
```

```
> AIC(output)
```

```
[1] 324.3407
```

```
> AICc(output)
```

```
[1] 326.7407
```

```
> a=summary(output)
```

```
> a$coefficients
```

```
$count
```

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	0.734411739	0.278171191	2.640143	8.287104e-03
x1	0.006088063	0.002443839	2.491188	1.273168e-02
x5	0.017419093	0.002504750	6.954423	3.540088e-12
x6	-0.012443627	0.002455632	-5.067383	4.033228e-07

```
$zero
```

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	1.410688745	1.11929830	1.2603331	0.20754921
x1	-0.043122654	0.02001185	-2.1548562	0.03117309
x5	0.002352948	0.02369544	0.0992996	0.92090039
x6	-0.038446578	0.03349001	-1.1480014	0.25096800

**Lampiran II. Surat Pernyataan Keaslian Data****SURAT PERNYATAAN**

Saya yang bertanda tangan di bawah ini, mahasiswa Departemen Statistika FMKSD ITS:

Nama : Rochmatun Annisa

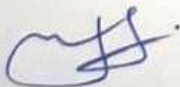
NRP : 06211540000052

menyatakan bahwa data yang digunakan dalam Tugas Akhir/ Thesis ini merupakan data sekunder yang diambil dari penelitian publikasi yaitu:

Sumber : Profil Kesehatan Kabupaten Majene 2016, Profil Kesehatan Kabupaten Mamasa 2016, Profil Kesehatan Kabupaten Mamuju 2016, Profil Kesehatan Kabupaten Mamuju Utara 2016, Profil Kesehatan Kabupaten Mamuju Tengah 2016, dan Profil Kesehatan Kabupaten Polewali Mandar 2016

Surat Pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui,  
Pembimbing I Tugas Akhir



Erma Oktania Permatasari, S.Si., M.Si  
NIP. 19881007 201404 2 002

Surabaya, 09 Juli 2019



Rochmatun Annisa  
NRP. 062115 4000 0052

*(Halaman ini sengaja dikosongkan)*



## BIODATA PENULIS



Penulis memiliki nama lengkap Rochmatun Annisa atau biasa dipanggil Annisa atau RA. Penulis dilahirkan di Nganjuk, 19 Oktober 1996 yang merupakan anak pertama dari 3 bersaudara. Penulis menempuh pendidikan formal di SDN Kramat 1, SMPN 1 Nganjuk, dan SMAN 2 Nganjuk. Kemudian penulis diterima sebagai mahasiswa Departemen Statistika ITS pada tahun 2015 melalui jalur SNMPTN. Selama masa perkuliahan, penulis aktif di beberapa organisasi yaitu Staff BPU JMMI ITS 2017, LDJ FORSIS, staf kemudian menjadi pengganti sekretaris direktur di BSO ITS *Education Care Center* BEM ITS Wahana Juang 2016-2018, asisten sekretaris eksekutif BEM ITS Gelora Aksi 2018-2019 dan sekretaris kabinet 1 BEM ITS Kolaborasi BEM ITS 2019. Penulis juga pernah mengikuti beberapa kepanitiaan yaitu OC acara GMAIL 2017, OC RDK 37, Pengajar Tangguh (PT 3) ITS Mengajar for Indonesia (IFI) 2016 dan menjadi sekretaris acara IFI 2017. Selain itu penulis juga aktif di komunitas daerah yaitu Muda Anjuk Ladang dalam divisi Pengabdian Masyarakat kemudian menjadi sekretaris 2 MAL. Capaian prestasi penulis antara lain finalis LKTI Nasional ESFEST 2018 di Universitas Pattimura Ambon dan mengikuti *Study Excursion ITS International Office Goes to Beyond* 2018 di Singapura. Bagi pembaca yang ingin berdiskusi, memberikan saran, dan kritik mengenai Tugas Akhir ini dapat disampaikan melalui email [rocmatun.annisa@gmail.com](mailto:rocmatun.annisa@gmail.com) atau melalui nomor 085655625037.